DISTRIBUTIVITY, MAXIMALITY, AND FLOATING QUANTIFIERS

by

CHRISTINE M. BRISSON

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Linguistics

written under the direction of

Veneeta Dayal

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 1998

ABSTRACT OF THE DISSERTATION

Distributivity, Maximality, and Floating Quantifiers

by CHRISTINE BRISSON

Dissertation Director:

Veneeta Dayal

This dissertation argues that the widely-accepted analysis of *all* and *both* as universal quantifiers is incorrect, and proposes instead that *all/both* are modifiers that place a boundary on the range of otherwise contextually-available interpretations allowed with definite plurals. It is argued that this proposal offers wider empirical coverage of the semantic and distributional properties of *all/both* than previous accounts.

The range of meanings of sentences with definite plurals is examined, and it is argued that a theory of distributivity that assigns universal force to distributed plural noun phrases is empirically inadequate. A change to the theory of distributivity is proposed in which the context-sensitive variable in the domain of the distributivity operator (the D operator) can be assigned a value that weakens its universal force. It is further proposed that *all/both* interact with the D operator by restricting the range of values that can be assigned to the resource domain variable, ensuring that the universal force of the D operator is not weakened. This accounts for the "strengthening" effect of *all/both* on the quantification associated with definite plurals.

Because *all* interacts with distributivity, it has the same "scope" as distributivity. This explains some differences between definite descriptions with *all* and quantified NP's with *every* that are unexplained if *all* is analysed as a universal quantifier. In particular, the different scopal possibilities of *all* and *every* in direct object position, and their differences in distribution with

collectivizing adverbial phrases, are explained. In addition, the limited distribution of *all* with collective predicates is accounted for by the proposal that lexically collective predicates belonging to certain aktionsart classes can contain a "hidden" D operator.

Finally, a theory of "floating" quantifiers in which movement does not play a central role is proposed, and argued to provide a better account for the distribution of floated *all/both* and for the relationship between *all/both* in their prenominal and floated positions.

ACKNOWLEDGMENTS

While in the midst of writing this dissertation, I imagined it would be great fun to finally get to the stage of writing the acknowledgments because not only would it mean that I was near the end, it would mean the chance to write about my mentors and friends and how much they mean to me. Now that I am in that stage, I am surprised to find out how hard it is to do justice to the debt I feel to the people who helped shape my work and my life as a graduate student.

I could not have asked for a better dissertation advisor than Veneeta Dayal. From the first day that she saw that there might be a generals paper topic lurking in my interest in *both*, she always seemed to know what I wanted to say before I knew it myself. Her linguistic intuitions are unerring and I found that whenever I ignored or set aside her suggestions I did so at my own peril. In addition to her linguistic expertise, her guidance in navigating the intellectual journey of writing a dissertation is something every graduate student should have, and I am grateful to her.

I always came away from my discussions with Viviane Déprez invigorated and excited about doing linguistics. Arguing with her was a lot of fun. Of course, most of the time she was right, but it seemed I always figured that out much later, and going through the process helped me immeasurably. I am grateful to her for improving this dissertation, and hopefully my future work, by among other things never getting tired of telling me to look at the idea behind the formalism or the rule.

One of the things I liked most about working with Roger Schwarzschild is that whenever I handed him a draft, he always zeroed in on the core of it, and reframed what I was trying to get across in such a way that I could turn it over and examine it from a new perspective to see if it was what I really wanted to say. Needless to say, many times the answer was no. In addition, I wish to

thank him for his insights and suggestions thared so many good suggestions on so many aspects of the dissertation that what I couldn't manage to incorporate in the dissertation will give me hearty food for thought in the years to come.

I am also grateful for having had the opportunity to work with Angelika Kratzer. She generously shared with me some of her insights into event semantics, and provoked me with her questions, and the dissertation emerged much the better for it.

For discussion of various parts of this dissertation, in various stages of its evolution, I wish to thank Eric Baković, Olga Babko-Malaya, Maria Bittner, Tom Ernst, Bill Ladusaw, Fred Landman, Peter Lasersohn, Jason Merchant, Susan Rothstein, Yael Sharvit, Colin Wilson,Yoad Winter, my fellow students at Rutgers, and audiences at SALT VII, the CUNY syntax lunch, and Johns Hopkins.

The graduate program in linguistics at Rutgers began the year my classmates and I arrived. From the beginning, the faculty were committed to doing their part to make it a challenging, bustling, and fun place to be. In particular, I want to thank Jane Grimshaw and Alan Prince for their constant attention to, and understanding of, the many intellectual, administrative, financial, etc. etc. needs of graduate students, and for always bringing good beer.

I want to thank my friends and fellow students in the linguistics department for helping make Rutgers such a good place to be a grad student. Special thanks to Strang Burton and Vieri Samek-Lodovici for going first, toYael Sharvit for coffee talk on linguistics and life, to Ed Keer and Kris Jaeger for great dinner parties, and to Eric Baković for making everything so much fun.

While I was working on my linguistics degree, all the while I had a secret other life as a residence counselor at Douglass College. Although at times it was difficult, I'm glad I had the chance to work with the great women who work for, and attend, Douglass. I think it helped keep

me grounded to work so closely with smart, interesting people who don't know or care what a D operator is. In particular I want to thank Kim Owens for always understanding exactly how things are.

Finally, I wish to thank my family. First my parents, who gave me innumerable gifts but perhaps the most relevant here being determination and self confidence. My brothers, Mike and Danny, and my sister Nancy, make this world a better place. And I thank Marc Meola for sharing the journey with me. He always told me things would be ok, even when I didn't want him to, and he was right.

Table of Contents

Chapter 1: Floating Quantifiers

Chapter 2: Plurality and Nonmaximality

Chapter 1

Floating Quantifiers

1 Introduction

The English words *all*, *both*, and *each* are often referred to as 'floating quantifiers' because they behave in some ways like determiner-quantifiers, but when related to subject position, they can also appear after the noun and even after auxiliary verbs. Thus they are said to have "floated" away from the canonical determiner position of the subject NP.

(1) All the girls went to the gym
The girls all went to the gym

(2) Both boys have been feeling ill
The boys have both been feeling ill

(3) Each of the students will receive a certificate of appreciation
The students will each receive a certificate of appreciation

Because of this special distribution, floating quantifiers have been used by several authors as a probe into syntactic structure (cf., Kayne 1975, Sportiche 1988, Déprez 1989, and Baltin 1995). In this work, I will use floating quantifiers as a probe into the semantics of plural noun phrases and verbal event structure. Ultimately I will argue that floating quantifiers are not quantifiers, but modifiers; and that they do not "float" (that is, their postnominal distribution is not due to movement). Hence the term "floating quantifiers" is perhaps particularly infelicitous, but I will continue to use it in a theory-neutral way because it is so widely known.

I will pay particular attention to *all* and *both* and will not have much to say about *each*; in the next section I will argue that there is good evidence for treating *all* and *both*

as comprising a natural subclass of the floating quantifiers.

1.1 The Distribution of all, both, *and* each.

The evidence for treating *all/both* as separate from *each* comes from their distribution. There are three separate distributional patterns which show *all/both* to be different from *each*.

First, broadly speaking, *all/both* can sometimes cooccur with collective predicates, in contrast to *each*. The pattern of *all/both* with collectives is a bit complicated and we will discuss it in detail throughout this dissertation. But for now, let us look at an outline of the facts.

There are three different subcases of "collectives" that show the difference between *all/both* and *each*. First, *all/both* can occur with some ordinary collectives, as in (4).

(4) a. The students all gathered in the hallway
b. The students both collided in the hallway
c. *The students each gathered/collided in the hallway

The predicate is changed from *gather* in (4)a to *collide* in (4)b to keep the sentence felicitous, since *both* carries a presupposition that the cardinality of the NP is two, and it is slightly odd to talk about two people gathering. With either predicate, however, *all* and *both* are felicitous while *each* is not.

In addition, *both/all* do not disambiguate predicates that allow both a distributive and a collective reading, which Link (1983) has termed "mixed-extension" predicates.

(5) a. The students all carried the piano upstairs (*distributive or collective)*
b. The students both carried the piano upstairs (*distributive or collective)*
c. The students each carried the piano upstairs (*distributive only*)

Finally, *both/all* allow an "internally reciprocal" reading with predicates like *share*

or *look alike*. In other words, (6)a-b can mean that the individual students who make up *the students* shared a cookie with other students in that group. *Each* does not allow this: (6)c can only mean that each student in the set denoted by *the students* shared a cookie with someone who is not one of *the students*.

(6) a. The students all shared a cookie (*internally reciprocal possible*)
b. The students both shared a cookie *(internally reciprocal possible)*
c. The students each shared a cookie (*internally reciprocal not possible*)

All of these examples, (4)-(6), represent subcases of the following generalization: *all* and *both* are permitted with some collective predicates, but *each* is never permitted with collective predicates.

Another phenomenon that picks out the same subclass of floating quantifiers was observed by Maling (1976). She points out two different environments that allow *all/both* but not *each*. The first is in an apparent partitive noun phrase without the preposition *of*, which she derives by a rule she calls *of*-deletion.

(7) All of the men
All the men

(8) Both of the men
Both men

(9) Each of the men
*Each the men

The second environment is in object position: *all/both* can appear after a pronoun, but *each* cannot.

(10) *I like the men all
I like them all

(11) *I like the cakes both
I like them both

(12) *I like the girls each
*I like them each

We will look at these phenomena in more detail in Chapter 5. It should be clear at this point, however, that whatever explanation is ultimately given, it will have to treat *both/all* as belonging to a separate subclass of the floating quantifiers.

Finally, there is the use of the so-called "binominal" *each*, (Safir and Stowell 1989) in which it appears that the quantifier specifies a kind of function relating the parts of one plural DP to the parts of another. *All/both* do not allow this use. (Thanks to Viviane Déprez, p.c. for pointing this out to me.) The contrast is shown in (13)-(15).

(13) The children bought three books each
(14) *The children bought three books all
(15) *The children bought three books both

I do not have an explanation for *each*'s binominal use (but see Safir and Stowell 1989 and references there). I simply wish to point out that it is another piece of evidence for treating *all/both* as belonging to a distinct subclass of the floating quantifiers, separate from *each*. From now on for convenience I will use the term 'floating quantifiers' to refer just to *all* and *both*, unless I specify otherwise.

1.2 The Category of Floating Quantifiers

The term *floating quantifier* of course suggests two things: first, that *both* and *all* are indeed quantifiers, and secondly, that the floated and the prenominal versions are structurally related occurrences of a single lexical item, since the former is derived from the latter by movement. Most work on floating quantifiers takes at least one or the other of these assumptions for granted.

Several authors have proposed that floated *all* is not moved from the subject

position, but base-generated in an adverbial position. For example, Dowty & Brodie (1984) explicitly propose that floated *all* is an adverb, and is generated in VP-adjoined position. Doetjes (1992) similarly proposes that floated *all* is an adverb that must bind an empty category (ie, a trace of an NP) inside the verb phrase. Baltin (1995) proposes that floating quantifiers belong to a small category he calls *preverbs* and that they are base generated in a position that directly precedes (c-commands) a predicative constituent. So while the syntactic origin of floating quantifiers is still being debated, the term 'floating quantifier' has stuck, and as I have mentioned, I will use it here in a theory-neutral way, to mean occurrences of *all/both* outside of prenominal position.

None of the authors mentioned specifically discusses the category of prenominal *all/both*; I assume that they would treat them simply as quantifiers. Authors who treat floating as the result of movement also, of course, treat prenominal *all* as a quantifier; so this is a point of commonality between the two types of appoaches to floating quantifiers.

The central claim of this dissertation is that *all/both* are not determiner-quantifiers[1] but modifiers. Throughout Chapters 3, 4, and 5 we will accumulate semantic and syntactic evidence for this claim.

In this section I want to argue that there is evidence against treating *all* as a determiner-quantifier. I will confine my attention to prenominal *all*; I will make the argument from prenominal *all* that *all* is not a true determiner-quantifier, and assume that

[1] For convenience, sometimes I will use the term *quantifier* for *determiner-quantifier*. Of course, modifiers can be quantificational too, and it is possible to think of the meaning I will give for *all/both* in Chapter 3 as being quantificational. The arguments I will make in these sections should be understood in the context of a comparison of *all* to determiner-quantifiers, in particular *every*. (I thank Roger Schwarzschild for help in clarifying this issue.)

the same holds true for the floated version. We will then look at how these arguments apply to *both*. I will adopt here and throughout this dissertation (with some qualifications to be noted in Chapter 5) what I take to be the null hypothesis, namely, that the prenominal and the floated versions of *all/both* have essentially the same meaning.

1.2.1 Arguments that all *is not a determiner-quantifier*

In this section we examine the evidence for the commonly-held assumption that *all* is a determiner-quantifier. I will argue that most of this evidence is not very decisive, and open to reanalysis. I will also provide arguments that *all* is not a quantifier: the evidence will come from discourse anaphora and questions with universal quantifiers.

Evidence for treating *all* as a quantifier comes mostly from its distribution (that is, its appearance in prenominal position) and from its apparent synonymy with the universal quantifier *every*. Let's consider each of these in turn.

1.2.1.1 Evidence from its distribution: prenominal all

The distribution of *all* is in some ways quite similar to determiner/quantifiers like *every*, *most*, and *the*.

(16) Every girl went to the gym
Most girls went to the gym
The girl(s) went to the gym

There are important distinctions among the determiner/quantifiers above -- for example *every* obligatorily takes a singular common noun as a complement, while *most* obligatorily takes a plural CN, and *the* can take either. Nevertheless, as a broad stroke generalization it seems fair to say that *all* is similar in distribution to the class of determiner/quantifiers, because it can appear prenominally.

But let us examine this observation more closely. If we restrict our attention to episodic sentences, we find that *all* is felicitous only when it appears before a definite plural DP. (Thanks to Veneeta Dayal for pointing this out to me. It has also been discussed by Partee 1995 and noted by Gil 1995;ftn.2.) It cannot occur with just a plural common noun (an NP). This is shown in (17).

(17) All the girls went to the gym
*All girls went to the gym

In episodic sentences, then, the prenominal distribution of *all* is actually quite different from the prenominal distribution of determiner/quantifiers. A quantifier like *every* combines with an NP (type <e,t>) to yield a generalized-quantifier-denoting DP or QP (type <<e,t>,t>). A determiner like *the* combines with an NP (type <e,t>) to yield an individual-denoting DP (type e).[2] Unlike either of these lexical items, *all* combines with a (plural) individual-denoting definite DP (ie., *the girls*).

Let us turn now to prenominal *all* in generic sentences. As I have just discussed, a generic interpretation is the only one possible for a sentence with prenominal *all*.

(18) All dogs are mammals

Of course, prenominal *all* isn't required in order for a generic interpretation to be available. The sentence in (19) is also interpreted generically.

(19) Dogs are mammals

There are two kinds of ideas in the literature about how sentences like (19) are

[2] Angelika Kratzer points out to me (p.c.) that it is not clear whether the definite description *the temperature* is type e, as in Barbara Partee's example, *The temperature is 85 and rising*. Nevertheless, it is still true that *the* combines with a property, which makes it apparently unlike *all*.

interpreted. Carlson (1977) proposed that bare plural NPs are always kind-denoting, and that the generic or episodic interpretation that a sentence receives is due to the properties of the verb phrase. Another approach is proposed by Wilkinson (1991) in which the bare plural contributes a variable that is bound by a (covert) generic operator.

Whichever analysis one chooses to adopt, the point that I have demonstrated here is that *all* does not apparently interfere with the generic interpretation. On the kinds analysis, it would appear that *all* combines with a kind-denoting expression and yields a kind-denoting expression. On the variable analysis, *all* does not prevent the binding of the variable introduced by the bare plural. This again suggests that *all* is not a quantifier.

If we call into question the idea that *all* is a quantifier, then the question arises as to what sort of thing it is. I argue that *all* is a modifier, which leads us to expect that prenominal *all* combines with a DP to form a DP (or with an NP to form an NP). This would explain the distributional facts we have just seen in episodic and generic sentences without any further stipulations. While I do not take up the issue of *all* with generics in this dissertation, we will see more detailed discussion of the syntactic position of prenominal *all* with definites in Chapter 5.

1.2.1.2 Evidence from discourse anaphora

The phenomenon of licensing discourse anaphora is one kind of construction where a definite plural with *all* behaves like a definite description, and not like a quantificational element. It is well-known that definite DPs freely license discourse anaphora, while quantified DPs (or QPs) do so only under certain conditions (namely, where the process of "abstraction" is possible; see Kamp and Reyle 1993 for discussion). Note the contrast

between (20) and (21).

(20) The girls came in. They sat down.

(21) Every girl came in. ??They/*she sat down.

If *all* were a quantifier, then we might expect that when it combines with a DP the resulting constituent would behave like a quantified DP (or QP) with respect to discourse anaphora. However, this expectation is not borne out. If we add prenominal *all* to the definite DP, a plural pronoun is licensed, just as above in (20). In at least this respect, *all the girls* behaves more like *the girls* than like *every girl*.

(22) All the girls came in. They sat down.

So it appears that prenominal *all* does not change the 'type' of a definite DP. This is expected if we treat *all* as a modifier.

1.2.1.3 Evidence from questions

In this section we will look at another difference between *all* and *every* that suggests that *all* is not a universal quantifier. The evidence comes from their contrasting behavior in licensing list answers to questions.

Questions with *every* in subject position license three different types of answers, as has been shown in the work of Engdahl (1986), Groenendijk and Stokhof (1984), and others. These can be called the functional answer (23)a, the individual answer (23)b, and the pair-list answer (23)c.

(23) Which woman did every boy kiss?
- a. His mother
- b. Judith
- c. John kissed Mary, Bill kissed Sue...

If we change *every boy* to *all the boys* the list answer is no longer possible..

(24) Which woman did all the boys kiss?
a. Their mothers
b. Judith
c. #John kissed Mary, Bill kissed Sue...

In the analysis of (23) proposed by Chierchia (1991, 1993), the possibility of a list answer depends crucially on a semantic property of universal quantification in a way that will be made precise shortly. If *all* were a universal quantifier, therefore, we would expect it to behave similarly to *every* in this respect. The fact that it doesn't suggests that *all* is not a universal quantifier.

Chierchia (1991, and see Chierchia 1993, Dayal 1996) proposed that the availability of list answers to questions with universal quantifiers is based in part on a logical property of universals, discussed by Barwise and Cooper (1981).

In the framework elaborated by Barwise and Cooper, generalized quantifiers denote sets of sets and the common noun argument provides the generator set for that "family" of sets. When the determiner-quantifier combines with the common noun, the result denotes a set of sets in a particular relation to the common noun set; the common noun set itself is not necessarily an element in the generalized quantifier denotation.

This is true for all generalized quantifiers. However, because universal quantification is the subset relation between sets, in a universal generalized quantifier the common noun set is recoverable. The generator set of a universal quantifier will be exactly that set which is the intersection of all the other sets in the generalized quantifier denotation. In other words, universal quantifiers have a unique "witness set".

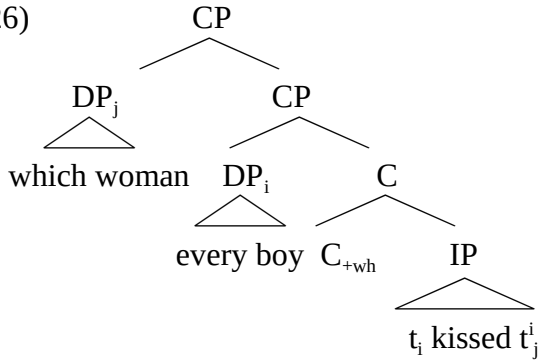
The availability of this unique witness set for a universal generalized quantifier plays a crucial role in licensing list answers for Chierchia, because this set provides the

domain argument for a function whose "spelling out" constitutes a list answer. A discussion of all the details would take us too far afield but we can look at a sketch of how this works. Consider the question (25).

(25) Which woman did every boy kiss?

Abstracting away from details that do not concern us here, Chierchia shows how to interpret this question for a structure like the one given below in (27).

(26)



(27) λP∃A[W(every.boy',A) & P(λp∃f∈[A→woman']∃x∈A[p=kiss(x)(f(x))])]

For Chierchia, the complementizer introduces existential quantification over a function from kissers to kissees. The wh-phrase *which woman* denotes the set of contextually salient women, and so provides the range of the function. To get the domain of the function, we must be able to extract a set from the denotation of DP$_j$. If DP$_j$ is a universal quantifier, as it is here, then we can use its witness set.

We also, of course, expect that this should be possible with all universal quantifiers. But as we have already seen, it is not apparently possible with *all* (see (24)). On any analysis that treats *all* as a kind of universal quantifier, we have to wonder why that universal quantifier does not make a witness set available to serve as the domain of the

function. There would have to be something else about the meaning of *all*, over and above its purported universal quantification, that prevented using its witness set in a functional question.

On the other hand, on the view that *all* is a modifier there is nothing that needs to be explained here. The difference between (23)-(24) is expected.

However, there is one complication in the data which I have so far avoided. In the examples in (23)-(24) the *wh*-terms are singular. However, if we use a plural *wh* term, or one that is unmarked for number (such as *who*), then sentences with definite plurals allow for list answers.

(28) Which women/who did the boys kiss?
  a. Their mothers
  b. Judith
  c. John kissed Mary, Bill kissed Sue...

But if we add *all* the possibility of a list answer, again, goes away.

(29) Which women/who did all the boys kiss?
  a. Their mothers
  b. Judith
  c. #John kissed Mary, Bill kissed Sue...

Dayal (1992) and Krifka (1992) argued that the list reading in (28) is not a functional reading, but a kind of "spell-out" of a cumulative reading, which we find in sentences with two plural arguments (see Scha 1991), such as (30).

(30) The boys kissed the girls

The cumulative reading of (30) is the following. Suppose the boys are Alex, Bill, and Carl, and the girls are Sally, Terry, and Ursula. Then (30) can be true if A kissed S, B kissed T, and C kissed U.

The phenomenon of cumulaltivity is a complicated one, and its relationship to distributivity is unclear. We will not discuss cumulativity in any detail in this dissertation, but I will point out that *all* on a subject DP makes a cumulative reading impossible, or at least much more difficult. Compare (30) to (31).

(31) All the boys kissed the girls

(31) is not true in the cumulative scenario described above.

If *all* interferes with cumulativity, and if plurality-based list answers are based on cumulativity, then the fact that *all* does not license a plurality-based list answer in (29) is expected (on whatever analysis of *all* one may propose). But this sort of explanation will not help save the quantifier analysis with respect to list answers for questions with singular *wh* terms, because cumulativity plays no role there. Hence the argument from questions against treating *all* as a quantifier still stands, although it is a bit more complex than it at first seemed.

1.2.1.4 "Semantic" evidence

The other argument for treating *all* as a quantifier is its apparent synoymy with *every*. This is shown by the data in (32)-(33).

(32) All men are mortal
Every man is mortal

(33) All the girls went to the gym
Every girl went to the gym

One problem for the synonymy argument is that, as I have already discussed, *all* can occur with some collective predicates. This is in contrast to *every*, which is never possible with collective predicates.

(34) a. ??Every student gathered in the hall
b. All the students gathered in the hall

(34) clearly shows that *all* cannot be completely synonymous with *every*.

In addition, as we have already seen, the "mixed-extension" predicates are not disambiguated by *all*. The range of possible interpretations of (35), in terms of distributive and collective interpretations, is exactly the same as the range of (36).

(35) The boys carried the piano upstairs
(36) All the boys carried the piano upstairs

Again, this similarity is expected on the view that *all* is a modifier.

I am not arguing from this kind of data that *all* couldn't be quantificational; an example like (37) shows that *gather*, for example, is compatible with some kind of quantification (unless we want to claim that *no* isn't quantificational). [3]

(37) No students gathered in the hall

What I am arguing is that we should treat characterizations of the semantics of *all* based on intuitionistic comparisons with *every* with suspicion. On closer inspection of the data it appears that *all the N* behaves much more like *the N* than like *every N*. Since it is clear that *all* and *every* are not exactly synonymous, we are not forced to treat *all* as a quantifier. We will have a lot more to say about examples like (34) as this dissertation proceeds.

I conclude that there is ample reason for treating *all* as a modifier rather than as a quantifier. Unlike determiner/quantifiers, *all* does not take a property as its first argument,

[3] Thanks to Angelika Kratzer (p.c.) for pointing this out to me. Also see the work of Winter (1998a,b) in which collective predicates and quantification is discussed quite extensively.

and when it combines with an NP or a DP the type of the resulting constituent appears to be the same as the type of NP or DP that *all* combined with. In this respect its behavior appears to be more like that of a modifier than that of a determiner.

The idea that *all* might be a kind of modifier is not new: this is one aspect of the proposal of Dowty and Brodie (1984) for floated *all*. Here I am proposing to extend this idea to prenominal *all*. A fuller account of just what prenominal *all* 'modifies' and how it combines with the DP will have to wait until I have given an explicit semantics for *all*. But for now, I take it that the preceding discussion gives us evidence for supposing that *all* is closer in spirit to a modifier than a genuine determiner/quantifier.

The fact that floated quantifiers have some kind of relationship to individual-denoting DPs was already noted by Roberts (1987:206), who pointed out that the category of DPs that allow floated quantifiers are those DPs that denote a plural individual (in the sense of Link 1983, and others). In other words, floated quantifiers are licensed by the same kind of DP which I have argued here combines with prenominal *all*.

I conclude that at the very least, it is possible to question the assumption that *all* is a quantifier. It behaves in many ways that are quite unexpected on the view that it is a quantifier. The path is open for us to argue that *all* has a different kind of grammatical function than is commonly assumed.

1.2.2 Both

Before we consider how these arguments apply to *both*, we should briefly review the analyses of *both* that have been proposed in the literature, since it is not widely considered to be the case that *both* is similar to *all*. In this discussion I will limit myself

to prenominal *both*, leaving aside the use of *both* as a kind of "conjunction introduction" which can be used to introduce conjunctions of many categories (ie, *Mary is both rich and intelligent, Grover both sings and dances*); but we will return to this use of *both* in Chapter 5. Also see Stockwell, Schachter and Partee (1973), and Edmondson (1978) for some discussion of this use.

Barwise and Cooper (1981) propose that *both* is a determiner that means the same thing as *the two*. A problem for this idea, which they point out, is that *both* is not licensed in the lower NP in a partitive construction, while *the two* is.

(38) *One of both children sneezed
(39) One of the two children sneezed

Ladusaw (1983) offers an explanation for the difference between (38) and (39) by first pointing out that collective predicates, like partitives, expose a difference in distribution between *both* and *the two*.

(40) *Both students are a happy couple
(41) The two students are a happy couple.

Now, we have already seen that *both* can cooccur with some collectives. However, Ladusaw takes the facts to be somewhat different: he treats *both* as being completely incompatible with any predicate that can be classified as collective.[4] Based on this, he proposes that the facts in (38)-(39) and (40) can be accounted for by hypothesizing that *both* has an essentially distributive component of its meaning which makes it incompatible

[4] Ladusaw (p.c.) has told me that he disagrees with some of the judgments I gave earlier in this chapter for examples like (4)-(6) (ie, *The students both collided*). He finds these sentences to be ill-formed. However, I have checked these judgments with many people and find that most speakers allow them. It appears we might be dealing with a dialectal difference.

with collective predicates and impossible inside partitives.

This apect of Ladusaw's analysis, namely, the idea that *both* is distributive, is shared by many others, including Edmondson (1978), Roberts (1987), and Landman (1989). However, we have already seen evidence that this cannot be correct, because *both* is possible with some collective predicates. If this is true then the fact that *both* cannot occur inside a partitive is left unexplained. We will offer a new explanation for *both*'s inability to occur inside a partitive, one that follows from the modifier hypothesis, in Chapter 5.

1.2.2.1 Arguments that both *is not a determiner-quantifier*

One argument for treating *all* as a quantifier is, as we have seen, its apparent synoymy with universal quantifier *every*. Likewise, *both* has been proposed to be equivalent to *each* or *every* (Roberts 1987, Landman 1989) plus the requirement that we are dealing with a plural individual whose cardinality is two. However, this cannot be correct because as we have already seen, *both* can felicitously combine with some collective predicates.

(42) The boys both collided in the hallway
??Every/each boy collided in the hallway

(43) The girls both shared a cookie (*internally reciprocal possible*)
The girls each shared a cookie (*internally reciprocal not possible)*

Of course, the data introduced by Ladusaw shows us that there are at least some collective predicates that are impossible with *both* (such as *be a happy couple*). So what is needed is a generalization about which collective predicates allow *both* and which do not. This will be the topic of extensive discussion, both later in this chapter and in

Chapters 3 and 4. For now, the point is that the argument that *both* is a quantifier because it acts like a universal quantifier (plus a cardinality requirement) is not decisive, because *both* doesn't act exactly like other universal quantifiers with respect to collective predicates. In this respect it is just like *all*.

In addition, I argued that *all* is not a true quantifier because it can combine with a (plural) individual (type e), unlike *every*. This argument also holds for *both*; it, too, can combine with a plural-individual denoting DP, as in (44).

(44) Both the girls went to the gym.

This shows that *both*, like *all*, can combine with a constituent of type e.

1.2.3 Possible problems for the modifier proposal

1.2.3.1 Both + *CN*

However, there is one argument for not treating *all* as a quantifier that does not apply to *both* in the same way. *Both* can, apparently, combine with a plural common noun to yield a definite description, in contrast with *all* (where, as we saw, the NP is always interpreted generically).

(45) Both girls went to the gym. (cf. (17))

This is not predicted by my characterization of *all/both* as belonging to the same subclass. However, an interesting fact pointed out to me by Bill Ladusaw (p.c.) sheds some light on what might be going on here. The only case where *all* can combine with an indefinite to yield a term that is not kind-denoting is when the indefinite is a numeral indefinite, such as *three girls*.

(46) ??All girls left early
(47) All three girls left early

The phrase *all three girls* behaves like a definite description, in that it must refer to an object previously mentioned in the discourse. For example, consider the difference between (48) and (49).

(48) Three girls came in. They drank some tea, then three girls left.
(49) Three girls came in. They drank some tea, then all three girls left.

In (48) the indefinite *three girls* cannot take as its referent the same three girls who came in and drank tea. In (49), *all three girls* must refer to the same three girls who came in and drank some tea.

Now, we have already seen that *both* can combine with an indefinite to yield a definite description. And of course, *both* has a cardinality requirement 'built in'. I want to suggest that if I'm right about *all/both* comprising a class, then we can explain (49) by saying that whatever contribution in meaning *all/both* make to a sentence, this contribution plus the cardinality requirement (as in *three* in (49)) adds up to definiteness. Then we don't have to state anything special about *both* because the definiteness of the subject NP in (49) follows in part from the (obvious) fact that it always comes with a cardinality requirement. However, I do not at this point have a formal analysis that captures this so for the purposes of this dissertation this is only a hunch.

1.2.3.2 Partitive constructions

One kind of evidence that might be taken to suggest that *all/both* are quantifiers is the fact that they can appear as the "upstairs" constituent in partitive constructions, as shown in (50).

(50) Both of the ducks
All of the correct answers

This property is shared by many quantifiers, as shown in (51).

(51) Most of the pens
Few of the politicians
Every one of the jellybeans
None of the teachers

However, here I will suggest two reasons why this does not force us to conclude that *all/both* are quantifiers.

First, there are many different categories that can appear upstairs in a partitive construction, including the modifier/nominal *half* and nouns that denote some kind of measure or group like *gallon* or *bunch*. (See Jackendoff 1977, Selkirk 1977 for discussion).

(52) Half of my students

(53) A gallon of the orange juice
A bunch of the paper clips

Numerals can also appear upstairs in a partitive construction. (I thank Viviane Déprez for pointing this out to me.) This is perhaps particularly relevant because many authors treat numerals as cardinality predicates (ie, as adjectives) rather than as quantifiers (see, for example, Link 1983, 84).

(54) Four of the cookies

So it is clear that we are not forced to conclude that anything that appears in upstairs position in a partitive must be a quantifier.

Secondly, there is interesting diachronic evidence that *both* and *all* came to occur in this construction from a different route than the quantifiers. The OED reports (compact edition 1971, p. 226) that the use of *all* followed by *of* is "comparatively modern, and is

probably due to form-assoc. with *none of , some of*..."[5] It similarly reports that the use of *both of* is "modern" (ibid, p 1012).

This again suggests that although *both* and *all* can appear in the same position as a quantifier in a partitive, they do not belong to the same category as the quantifiers.

So to sum up, I have argued here that *all/both* are not quantifiers or determiners, and suggested that they are closer in spirit to modifiers. However, I will continue to call them “floating quantifiers” and sometimes even just “quantifiers” simply for convenience, and I don't intend that any theoretical claims should be associated with these terms. A more specific proposal about how they "modify" will be the subject of Chapters 3 and 4.

1.3 The Syntax of **all/both** *and the Phenomenon of 'Floating'*

Perhaps the most well-studied aspect of *all* has been its syntactic behavior, specifically with respect to the phenomenon of "floating". The earliest approaches to this issue in the generative literature, including work by Postal (1974), Kayne (1975), and Maling (1976), proposed that a rule of "quantifier float" applies to this class of quantifiers and moves them rightward, away from the NP. Dowty and Brodie (1984) proposed that floating quantifiers are not really moved quantifiers at all, but adverbial elements that are base-generated in their VP-adjoined positions.

The idea that *all/both* are modifiers is compatible with either a movement or a base-generated analysis of floating quantifiers. In Chapter 5 I will argue that the syntactic and semantic evidence favors the base-generation treatment of floating quantifiers,

---

[5] If *all* came to be possible in partitives by analogy with quantifiers, this might be taken to suggest that *all* really is a quantifier. But speakers can make analogies that are grammatical errors, as the morphological phenomenon of backformation shows.

specifically as adverbs. I will anticipate the conclusions of that chapter by treating floated *all* as an adjoined adverbial phrase. In addition, we need to assume some mechanism for relating floating quantifiers to the subject. Since in English floated *all/both* are always construed with the subject, I will simply assume by stipulation, until Chapter 5, that a floating quantifier is always thusly construed.

Our task in the meantime is to look more closely at the semantics of *all* and *both* in order to make a convincing semantic proposal in which they are treated as modifiers. In the discussion in the rest of this chapter and for much of Chapters 2-4, I restrict my attention to *all* in order to keep the discussion simpler. However, nearly everything I say about *all* applies to *both*, and *both* will be specifically discussed at relevant points throughout the dissertation.

1.4 The semantics of **all**

1.4.1 The maximizing effect

As we have already seen, contrasts like the one in (55) suggest that *all* cannot be characterized simply as a universal quantifier.

(55) (a) ??Every student gathered in the hall
(b) All the students gathered in the hall

The universal quantifier *every* distributes the property of having gathered in the hall down to each individual student: since a single student cannot gather, the sentence (55)a is ill-formed. In contrast, in (55)b the function of *all* seems to be to say that each individual student participated in the gathering.

But if we can't treat *all* simply as a universal quantifier, we do need some way to capture the "strengthening effect" that *all* has on a sentence with a definite plural. This

effect can be illustrated by the examples in (56)-(57) and (58)-(59).

(56) The boys jumped in the lake
(57) The boys all jumped in the lake

(58) The students gathered in the hallway
(59) The students all gathered in the hallway

While (56) can be judged true if one or two of a large group of boys stayed behind on shore, (57) seems to more strictly require that every boy jumped in the lake. Similarly, (58) can be true if even one or two of the students didn't participate in the gathering, but (59) requires that every student was there. Link (1983) called this the "totality" effect and Dowty (1987) called it "the maximizing effect". These authors take it that explaining the meaning of *all* involves explaining its totality effect while still allowing *all* to combine with collective predicates. They propose similar solutions to this problem.

Link's solution is to propose that *all* introduces a 'partakes in' operator, which distributes the property of taking part in the action down to every individual in the subject of the predicate. So, for example, (59) means that every student took part in the gathering. For a distributive predicate like *jump in the lake,* Link introduces a meaning postulate that says 'taking part in' just reduces to ordinary distributivity. So 'every boy took part in jumping in the lake' means the same thing as 'every boy jumped in the lake'.

Dowty (1987) recasts the idea of 'taking part in' by proposing that the lexical representations of certain (collective) predicates provide "distributive subentailments" for *all* to operate on. If a predicate occurs with *all*, then the function of *all* is to distribute the subentailments of that predicate down to each individual in the subject noun phrase.

What are distributive subentailments? Dowty does not provide a definition, but

gives the proposed subentailments for *gather* as an example. The distributive subentailments of *gather* are something like, 'come to be in the same place at the same time as a lot of other people'. Dowty's idea, then, is that *all* is a universal quantifier that distributes this property, the property of 'coming to be in the same place...' down to every individual in the denotation of the subject.

One problem with this proposal is that while it is relatively easy to see what the distributive subentailments of *gather* are, it's harder to see this for other collective predicates. Take the collective reading of *build a raft*: what are the relevant distributive subentailments here? Sawing wood, reading blueprints, hammering nails? We might wonder whether it is reasonable to suppose that all of this information is lexically specified. Another problem is that, as we have already seen in our discussion of *both*, not all collective predicates can cooccur with *all*. For example, the sentences in (60) and (61) are ill-formed.

(60) *The students are all a big group
(61) *The students are all numerous.

According to Dowty, these predicates do not allow *all* because they do not have subentailments, and hence do not give *all* anything to 'operate on'. He calls these predicates "pure cardinality predicates". The problem is that some evidence or explanation must be given about why these predicates lack distributive subentailments. Unfortunately, Dowty does not provide any evidence. So one problem for Dowty's account is that it does not give a satisfactory account of why some collective predicates allow *all*, and others don't.

1.4.2 Taub's criticism of distributive subentailments

Taub (1989) criticizes Dowty's account on the grounds that it is circular: the only evidence for the existence of subentailments is the distribution of *all*, and yet subentailments are called upon to explain the distribution of *all*. I will criticize Dowty's account shortly on different grounds. But first let us take a look at Taub's discussion.

Taub gives an interesting characterization of the collective predicates that do not allow *all*, in terms of the Aktionsart classification of predicates (on Aktionsart classification see also Dowty 1979 and the references cited in Chapter 4, section 4.4). I'll call this characterization Taub's generalization, and it is given in (62).

(62) **Taub's Generalization**: the collective predicates that disallow *all* are the collective predicates denoting states and achievements.

The evidence that Taub gives for this generalization is given in the following examples.

Collective states:
(63) *The boys are all a big group
(64) *The trees are all dense in the middle of the forest (*on collective *dense*)

Collective activities:
(65) All the boys carried the piano around for an hour

Collective accomplishments:
(66) All the students gathered in the hallway
(67) All the girls built a raft

Collective achievements:
(68) *All the senators passed the pay raise
(69) *All the students elected a president

Note that Dowty's cases of "pure cardinality predicates" like *be a big group* and *be a group of four* are states, and so fall under Taub's generalization. But it is not the case that only cardinality predicates disallow *all*. The example in (64) is from Taub's paper

(she attributes the example to Barbara Partee), and I would add (70). On a collective reading both are infelictious with *all.*

(70) *The bottles are all too heavy to carry

Note that Taub's generalization also correctly categorizes the collective predicates that do and do not allow *both*. For example, the collective predicate *be a happy couple* is stative, while *share* is an activity predicate.

(71) *John and Mary are both a happy couple
(72) Sara and Mary both shared a cookie

Taub's generalization about the distribution of *all* (and its extension to *both*) has the advantage that the Aktionsart classes are well motivated on grounds independent from anything having to do with *all*. Thus it does not suffer from the same circularity as Dowty's idea about distributive subentailments. However, the problem is, as she points out, that it's not clear how the generalization about Aktionsart classes relates to the meaning of *all*.

We will give an explanation for Taub's generalization in Chapter 4. For now, I want to turn our attention to a different problem with the Link/Dowty approach to the meaning of *all*.

1.4.3 A problem with the maximizing effect

Link and Dowty agree that the effect of *all* is a strengthening effect. This is shown by the pair of sentences in (56)-(57), repeated as (73).

(73) a. The boys jumped in the lake
b. The boys all jumped in the lake

Recall that the relevant observation is that (73)(b) somehow more stringently requires that

each and every boy jumped into the lake: it shows the maximizing effect. But here is the problem. It is clear that the effect of *all* is to somehow strengthen the truth-conditions of a sentence with a definite plural NP subject, but the meaning that Link assigns to these sentences is already quite strong without *all*. Dowty doesn't give a specific analysis of plurals, but suggests that his proposal for *all* can be implemented in just about any theory, and we will see shortly that the problem arises in several theories, not just Link's. So the problem I will discuss here applies equally well to Dowty's proposal about *all*.

For now I will just briefly sketch a theory of plurals similar to Link's (1983) so that we can get an idea of what the problem is; a more thorough discussion will be the subject of Chapter 2. Consider for now only the case of sentences with predicates that are distributive, like *jump in the lake*. Let us assume that a definite plural NP like *the boys* denotes a (first-order) set (in this characterization I am abstracting away from Link's original formulation a bit). To interpret this predicate as applied to *the boys*, we introduce a covert distributivity operator on the plural predicate, which Link represents as ^{D} (in (75)). The D operator has the effect of introducing universal quantification over the members of the set that the subject denotes, as shown in (76).

(74) The boys jumped in the lake
(75) ^{D}jump.in.lake'(the.boys')
(76) $\forall x[x \in [\![\text{the.boys}']\!] \rightarrow x \in [\![\text{jump.in.lake}']\!]]$

But now we have a problem: it is clear, as we have already said, that (73)b more strongly requires that each boy jump in the lake (for the sentence to be considered true) than (73)a. However, the interpretation we have assigned to (73)a, shown in (76), is already quite strong. How could (73)a be stronger than the universally-quantified

expression in (76)? The problem is that the truth conditions Link assigns to (73)a are too strong.

I have presented this problem in terms of Link's theory of plurals and distributivity, but the same issue arises in many theories of plurality. This includes Lasersohn (1990, 1995), in which a D operator is always introduced when a distributive predicate is applied to a plural individual, and Schwarzschild (1996), which is similar in this respect.

This problem has been noticed before: the discussion in Dowty (1987) makes it clear that he was aware of this problem, but he did not give any explicit semantic account of it. And Lasersohn (1993, ftn. 3) also notices it. However, to my knowledge no proposal for handling it has been presented.[6] We will take up this problem in detail in Chapter 2, where I will propose a 'weakening' of the semantics we assign to sentences with definite plurals.

1.5 An Outline of Things to Come

Before we can give a semantic proposal for *all* we need to deal with the problem of nonmaximality. In Chapter 2 we will look at the problem of nonmaximality with definite plurals in some detail, and develop two desiderata for a theory that captures nonmaximality. An examination of various different proposals in the literature will lead us to conclude that none of these proposals does quite the job we need.

In Chapter 3 we will propose a theory that captures nonmaximality. It is a theory of distributivity with a pragmatic component based on the proposals in Schwarzschild

---

[6] Actually, in work in progress Lasersohn has taken up this challenge. We will look at his proposal at the end of Chapter 4.

(1992, 1994, 1996). Thus it satisfies the two main desiderata for a theory of nonmaximality that emerge from Chapter 2: it captures the quantificational nature of nonmaximality and its pragmatic character. This theory of nonmaximality paves the way for a very simple account of the meaning of *all*/*both*, in which they are proposed to be modifiers of the distributivity operator, and which captures their "maximizing" effect.

In Chapter 4 we extend the empirical coverage of the proposal for *all*/*both* in several ways; and in every case we will see that the scope of the distributivity operator that *all/both* modify plays a crucial role in our explanation. We introduce an event semantics as the framework for capturing *all*'s behavior with collectivizing adverbials, and for a proposal about how lexical aktionsart interacts with distributivity to account for Taub's generalization.

In Chapter 5 we turn our attention to the syntax of floating and prenominal *all/both*. I argue that the distribution of floated *all* is better accounted for under the hypothesis that it is an adverb, and show how to account for the syntactic link between adverbial *all* and the subject. I also propose that prenominal *all* is in the specifier of D, which explains many peculiarities of its distribution.

Chapter 2

Plurality and Nonmaximality

2 Introduction

In Chapter 1 we saw that a theory of the meaning and distribution of floating quantifiers raises, in particular, two issues related to theories of plural noun phrases. First is the collective/distributive distinction, which is important because floating quantifiers are allowed with some, but not all, collective predicates. In this chapter we will look in more detail at the collective/distributive distinction, and at how it should be represented in the grammar.

The second issue was the issue I call "nonmaximality," after Dowty's (1987) term, 'the maximizing effect'. The question there is whether sentences with definite plurals should apply "maximally", that is, apply some property to every individual making up the plural DP, and if so, how that should work for both distributive and collective predicates.

In this chapter we will look at some proposals for handling distributivity and especially nonmaximality. The goal of this chapter will be to develop an idea of what kinds of things a theory of nomaximality should take into account. We will develop two desiderata for a theory of nonmaximality. Then we will examine the mainstream view on nonmaximality, and see that it does not meet our desiderata.

Our discussion will proceed with a review of some approaches to definite plurals that have been proposed in the literature, paying special attention to these two issues, the

collectivity issue and the nonmaximality issue. Since we are restricting our review in this way, there are many works on plurality that I will necessarily be forced to leave out of this brief survey, but I hope that what I lose in breadth I gain in clarity of focus on the issues of concern to us here.

2.1 Distributivity and collectivity

While I discussed the behavior of some distributive and collective predicates in Chapter 1, up to this point we have not had an explicit discussion of the difference between distributivity and collectivity. Before we can go any further we need to get more specific. Consider the pairs of sentences in (1)-(2) and (3)-(4) .

(1) George and Sandra left.
(2) George left and Sandra left

(3) George and Sandra are a happy couple
(4) *George is a happy couple and Sandra is a happy couple

The oft-reported difference between (1)-(2) on the one hand, and (3)-(4) on the other, is that (1) entails (2) (although we will soon have reason to question this assumption) and (3) doesn't entail (4) (since (4) is neither true nor false). We call a sentence that exhibits the entailment pattern in (1)-(2) distributive — the property of having left distributes to each member of the NP *George and Sandra*. A sentence that does not license this type of inference, as in (3), is called collective.

Note that the distributive/collective distinction does not arise only for conjoined NPs. For the examples in (5)-(8), suppose that *the girls* are Lauren, Sue, and Jill. Then (5) entails (again, apparently entails) (6), but (7) doesn't entail (8) .

(5) The girls left
(6) Lauren left, Sue left, and Jill left.

(7) The girls gathered in the clubhouse
(8) *Lauren gathered..., Sue gathered..., and Jill gathered...

Once we recognize the distributive/collective distinction the question arises as to how it should be represented in the grammar. For example, is distributivity/collectivity a property of a noun phrase, a verb phrase, or both? (or neither?). Proposals have been made in the literature for all of the above possibilities: some claim the distinction is in the NP (Bennett 1975, Gillon 1992, van der Does 1993, Heim, Lasnik, and May 1991); in the VP (Lønning 1987, Schwarzschild 1991, 1996, Lasersohn 1990, 1995); in both (possibly Link 1983, 1984, Landman 1989, 1996); or in neither (Roberts 1989, Schwarzschild 1992b, where distributivity applies when a predicate is combined with its plural argument). And some of the above authors make use of more than one strategy.

Lasersohn (1995, ch. 7) gives a very thorough discussion of this debate. Here I will simply review his best argument in favor of treating collectivity/distributivity as a property of the VP. The crucial sort of example comes from VP conjunction (Lasersohn attributes this argument to Dowty[1]), as shown in the sentence in (9).

(9) Rick and Ilsa met in a bar and had a drink.

The point is that *Rick and Ilsa met in a bar* has to be interpreted collectively. It does not follow from *Rick and Ilsa met in a bar* that Rick met in a bar. But *Rick and Ilsa had a drink* is at least allowed to be interpreted distributively; (9) does not require that Rick and Ilsa share a drink but allows for the possibility that each had their own drink. If

[1] Roger Schwarzschild (p.c.) points out that Roberts (1987:121) attributes this argument to Karina Wilkinson, and that in his own work (Schwarzschild 1996:15) he attributes it to G. Massey (1976:103).

collectivity were a property only of the NP, it would be impossible to represent the truth conditions of this sentence because the same NP would have to be both a collective NP and a distributive NP[2]. However, if collectivity is a property of the VP, then we may simply assume that one conjunct in the VP is a collective predicate, and the other is distributive. From this I take it that the collective/distributive ambiguity is located in the VP.

For now, then, let us assume this simple theory of the distributive/collective distinction. A plural noun phrase denotes a set. A verb phrase denotes a one-place predicate. Distributivity is a property of a verb phrase (ie, a predicate), and it is represented in the grammar by an optional D operator (as we saw in the previous chapter). Lack of a D operator will yield a collective reading. Our D operator is defined as in (10), which means that a D operator applied to a predicate P will yield the expression in (11).

(10) $^{D}=_{df} \lambda P \lambda x \forall y[y \in x \rightarrow P(y)]$

(11) $^{D}P = \lambda x \forall y[y \in x \rightarrow P(y)]$

The function of the D operator is to introduce universal quantification over the plurality introduced by the subject. This move is justified for two reasons. First, it is widely observed that for some predicates, at least, sentences with definite plurals apparently have universal quantification as part of their meaning. The examples in (12)-(14) are taken from Higginbotham (1981), Link (1983), and Yoon (1996), respectively.

(12) The men are left-handed

---

[2] Actually, van der Does and Gillon both say that there are ways around this argument against the NP strategy; van der Does argues that a strategy like quantifying-in gives us a way to have one noun phrase in two places, and Gillon proposes that we can use small *pro* to have two noun phrases in an anaphoric relationship. But then we would have to say that the two NPs, though coindexed, are not coreferential, since one denotes a quantified expression (the distributive NP) and one is a definite description (the collective NP).

(13) The pigs died
(14) The girls are 8 years old

The authors cited report the intuition that for these sentences to be true, all the men have to be left-handed, all the pigs have to have died, and all the girls have to be 8 years old. In upcoming sections I will question whether these sentences are really strictly universal but for now let us accept this judgment.

Furthermore, it can be argued that there is evidence we're dealing with universal quantification even for sentences where this intuition is not so strong, like (15) from Kroch (1975) (cited in Lasersohn 1993).

(15) The townspeople are asleep

It seems that we might be willing to say (15) is true even if a few townspeople are not asleep. However, Kroch points to (16) as evidence against this.

(16) *Although the townspeople are asleep, some of them are awake.

Kroch argues that the infelicity of (16) suggests that we really are dealing with universal quantification here.

Another argument for introducing a universal quantifier comes from sentences like (17).

(17) The boys ate a sandwich.

It is at least possible to interpret this sentence as meaning that each boy ate a <u>different</u> sandwich. If the VP is interpreted with a D operator, which introduces a universal quantifier, then we get the right interpretation because the universal quantifier

introduced by the D operator can take scope over the existential *a sandwich*[3]. So (18) is interpreted as shown in (20).

(18) The boys ate a sandwich
(19) ^{D}ate.a.sandwich'(the.boys')
(20) $\forall x[x \in [\![\text{the.boys}']\!] \rightarrow \text{ate.a.sandwich}'(x)]$

On the other hand (18) can also be true in a situation where the boys shared a single sandwich (maybe it was a 3-foot-long hero). We can represent this, collective, reading by predicating *the boys* of the VP *ate a sandwich* without any intervening D operator, as in (22).

(21) The boys ate a sandwich
(22) ate.a.sandwich'(the.boys')

This is the same way we would represent a predicate that is always interpreted collectively, such as *gather*, shown below.

(23) The boys gathered in the hallway
(24) gathered.in.the.hallway'(the.boys')

So for now, the distributivity/collectivity distinction is represented as follows: the distributive interpretation is represented by means of a D operator on the VP. A sentence that is interpreted collectively involves direct predication of the plural subject and the VP, without an intervening D operator and hence without any quantification involved.

At this point, then, we will take it to be true that universal quantification is part of the meaning of at least some sentences with definite plurals, in the form of the D operator. But of course we saw in Chapter 1 an argument from *all* that the D operator has truth

[3] It doesn't matter to us whether *a sandwich* is treated as an existential quantifier or whether it introduces a variable that gets its quantificational force from existential closure.

conditions that are too strong. In the discussion to follow we will see more evidence that the D-operator semantics we are tentatively employing is too strong. However, we will also see that an alternative that doesn't make use of the D operator is too weak. These facts will play an important role in our development of desiderata for a theory of nonmaximality.

In the next section we will review some proposals from the literature for handling nonmaximality.

2.2 Nonmaximality

In Chapter 1 we observed that (25) can be interpreted to mean that each boy is building a separate raft.

(25) The boys are building a raft

As we have just seen, we are able to obtain the "one-raft-per-boy" reading by using a D operator, which introduces universal quantification over *the boys*. However, I argued in Chapter 1 that even on this distributive reading, the sentence doesn't entail that every single boy is building a raft; if we are dealing with a large enough group of boys, and one or two of the boys is cleaning up from lunch, or napping in his cabin, the sentence is still true. This is a problem because we need the universal quantifier to obtain the "one-raft-per-boy" reading, but it looks like using the universal gives us truth conditions that are too strong. This point can be made most clearly by comparing (25) and (26).

(26) Every boy is building a raft.

In (26) we have the universal quantifier *every*, so we expect that as long as *the boys* in (25) and the contextually-restricted set of boys in (26) are the same, the two sentences should be synonymous. But the sentences are not synonymous. So the question we address

in the non-maximality issue is whether or not sentences with definite plurals are really synonymous with sentences with universally-quantified subjects, as the D-operator analysis predicts.

In the works cited in the previous section, some authors mention nonmaximality, but most of those who do mention it do not propose a way for the semantics to capture it (notable exceptions are Lasersohn 1990, 1995 and Landman 1989, 1996, which we will discuss shortly). So even though most authors are aware of the phenomenon of nonmaximality, it does not occupy a central place in theorizing about definite plurals.

On the other hand, a closely related issue has occupied a central place of importance in the study of reciprocals, especially reciprocal *each other*. Since reciprocals must have plural antecedents, the two topics are very closely related. Therefore a look at what has been said about nonmaximality in the literature on reciprocals will prove illuminating.

2.2.1 *Nonmaximality and reciprocals*

In his early paper on conjoined structures, Dougherty (1970) proposed that a sentence like (27) is derived via transformation from a sentence like (28).

(27) The men are hitting each other
(28) Each of the men is hitting the others

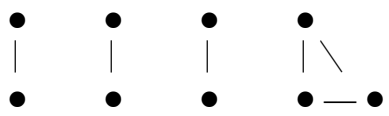
Fiengo and Lasnik (1973) object to this proposal, partly on the grounds that whenever the plural *the men* is a group greater than two in number, the truth conditions of (27) and (28) can be distinguished. They call the two different truth conditions the *'each-the-other'* reading (28) and the 'reciprocal' reading (27). Their discussion of the differences between the *each-the-other* relation and the 'reciprocal' relation is important here because

it's an early account that highlights some of the differences we find between sentences with overt quantifiers (like *each*) and sentences with plurals.

A sentence like (28), with *each-the-other,* is true in a situation only if every individual in the plural antecedent (here, *the men*) is in the relation R (here, *hitting*) with every other member. The reciprocal relation, however, is weaker: Fiengo and Lasnik claim that it requires simply that there be some way of carving up the plural antecedent into subgroups such that *each-the-other* holds within each subgroup. Note that the case where there are just two individuals in the plural antecedent is a special case of the reciprocal relation: the only way to divide the antecedent into subgroups in this case is trivially to have a single subgroup.

Fiengo and Lasnik illustrate with the diagram shown in (29). Suppose that each of the circles represents one of the men and the arrows represent the hitting relation. Their point is that (27) can be true in the situation illustrated by (29), but (28) cannot.

(29) a b c d



The *each-the-other* relationship holds for each subgroup. This set of subgroups is called a *partition*; the definition of partition is given below.

(30) For some set X, a set Y is a partition of X iff:
a. Y is a set of subsets of X
b. $\forall x \in X\ \exists Y \in Y\ [x \in Y]$
c. $\forall Z,Z' \in Y[Z \neq Z' \rightarrow Z \cap Z' = \varnothing]$
d. $\varnothing$ is not in Y

The requirement in (c) means that the sets of the partition must be nonoverlapping.

(Note that a partition where the sets are allowed to overlap is called a *cover*; so a partition is actually a kind of cover. We will see covers again later.) Fiengo and Lasnik claim that this requirement is crucial to prevent (27) from being true in a scenario like the following: four men are in a circle, and three of them are hitting a single man who is hitting the other three back. In this case our intuition is that (27) is not true, and the partitional requirement captures this, since the man who is hitting the other three makes the subgroups of *each-the-other* overlapping.

However, they also discuss a scenario in which it appears that even the partition requirement is too strong. This kind of situtation is one in which someone walks into a room to find a "general brawl" going on. In this situation, (28) is clearly false, but it seems that (27) is true, although it is perhaps not necessarily so clear what the partition should be, and not so clear whether the partitions are really nonoverlapping. What this shows, they say, is that (27) admits a kind of vagueness that (28) does not. They attribute this to the fact that *each other* sentences can "characterize the entire set", while *each the other* sentences "characterize each member of the set" (p. 452).

In other words, they point out that the plural anaphor *each other* allows a kind of vagueness that the quantifier *each* does not. The point here is that this difference is very much like the difference we saw earlier between the D operator and the overt quantifier *every*. I think it's also worth pointing out that the most plausible 'general brawl' scenario involves many more than four men, and so their examples also show that the number of individuals in the plurality we are talking about makes a difference.

Finally, Fiengo & Lasnik point out one other difference between *each-the-other*

and reciprocal sentences, and attribute it directly to the close relationship between reciprocals and plurals. The difference is between (31) and (32); (31) can describe the situation explicated in (33) , but (32) cannot.

(31) Each of the cars bumped into the other
(32) The cars bumped into each other

(33) The Pontiac bumped into the Plymouth on Monday, and the Plymouth bumped into the Pontiac on Tuesday.

It appears that (32) must refer somehow to a single event, or at least to two events in "the same general time span,"(p. 451) but (31) doesn't require this. This requirement, they say, isn't restricted to reciprocal sentences, but follows "from the general semantic properties of singulars and plurals". (p. 451) While this characterization is vague and suffers from some apparent counterexamples (*the women left* doesn't require that the women all left at the same time) we would like an account that does justice to their intiutions about the "general semantic properties of singulars and plurals".

There are two main lessons to be drawn from Fiengo and Lasnik's work that are important here. First, their discussion brings out some of the differences between plurals and universally-quantified expressions like *each*. These differences are not expected on the D-operator account. A similar point is made by Langendoen (1978) and Dalrymple, Kanazawa, Mchombo, and Peters (1994). Langendoen proposes several types of formulas for capturing the truth conditions of sentences with reciprocals, and points out several different kinds of cases where the quantification involved is different from the kind of quantification introduced by the universal quantifier *each*. Dalrymple et al take a similar approach to reciprocals.

Secondly, Fiengo & Lasnik introduce the idea of a partition of a plurality. This idea was used later by Higginbotham, who proposed that partitions can be used not just in modeling reciprocals, but also simple plural sentences. This was picked up by Gillon (1987) and later by Schwarzschild (1994, 1996), and it will prove important to us as we proceed.

Fiengo and Lasnik's point about the differences between *each other* and *each-the-other* is echoed and amplified in a later discussion by Williams (1992), which we will look at here. But since Williams' paper is a reply to Heim, Lasnik, and May (1991; hereafter HLM), first we need to briefly summarize the ideas in the latter paper.

2.2.1.1 Williams, and Heim, Lasnik, and May

HLM propose in their work to use reciprocal *each other* as a probe into the syntactic and semantic properties of anaphora, plurals, and quantification. The basic idea is that *each other* is built from *each* and *other* and its meaning and distribution should be predictable from its parts: in particular, the quantificational force of *each other* should derive from the fact that *each* moves and introduces distributivity, and the anaphoric properties of *each other* should derive from the fact that *each* leaves a trace in the position next to the anaphoric and relational noun *other*.

HLM go on to propose how these elements of their analysis can predict the various ambiguities that are (and aren't) found with sentences with reciprocals. For example, they discuss the following example from Higginbotham (1981).

(34) John and Mary think that they like each other.

There are at least two readings available for this sentence. One is that John and

Mary think the same thing: that John and Mary like each other. Another is that John thinks that he likes Mary, and Mary thinks that she likes John. The availability of these possibilities is explained in part by the possibility of *each* raising to attach to *they* or to attach to *John and Mary*. In what is for the purposes of this discussion an important decision, all of the examples HLM consider are cases where the antecedent of *each other* denotes a plurality of just two, which they justify as a move to simplify the data.

Perhaps the biggest problem for this account is that it predicts that a sentence with *each other* should always yield an *each-the-other* reading, because *each* is a distributive quantifier[4]. This is true for all of the data they examine but it is not true when a wider range of cases is considered. As Fiengo and Lasnik had already pointed out, the reciprocal relation will boil down to *each-the-other* whenever the antecedent of the reciprocal is a group of two (or three), but when the antecedent is a larger group the truth conditions get weaker.

Williams (1991), in his reply, takes this point even further. Williams calls the case of *each-the-other* "strong distributivity" and *each other* "weak distributivity," because he argues that the 'weak' and 'strong' distinctions are found not only in reciprocals, but with plural noun phrases in general. He cites Fiengo and Lasnik's examples (repeated here), and adds to them the sentence in (37) .

(35) They were each hitting the others
(36) They were hitting each other

[4] Strictly speaking, this is not quite correct. HLM begin the article by supposing that the *each* of *each other* is just the same as floated *each*, but find reasons later in the article, and in their reply to Williams, to back away from this position. It can be argued, however, that doing so makes their theory somewhat less compelling, since then they are not so far from claiming that *each other* is simply a distinct lexical item, rather than a complex constituent whose meaning is derivable from the meaning of its parts.

(37) They were hitting Bill

He recalls that (36) can describe a "general melee" which allows there to be non-hitters among those referred to by *they*. The same is true for (37), Williams says: "it is compatible with a situation in which there were some nonhitters." (ie, among the group picked out by *they*, p.162). Williams argues that the property of allowing weak distributivity is shared by both plural noun phrases and reciprocal *each other*, and so the meaning of *each other* is best explained, not by compositionally building it up from the meaning of *each* and *other*, but by treating *each other* as an anaphor which must be represented as a "plural variable."

Williams' discussion is important here because he explicitly connects the "weakness" of the reciprocal *each other* to the "weak distributivity" of plural noun phrases. Weak distributivity might be one way of capturing the phenomenon that I have here been calling nonmaximality: it would capture the idea that the property named by the predicate distributes down to some, but not necessarily all, of the members of the plural. But Williams unfortunately doesn't give us an explicit theory of how weak distributivity could work. A more explicit proposal about how weak distributivity might work is given in a recent paper by Yoon (1996) which is concerned with the truth conditions of donkey sentences with plurals. We will ultimately decide that Yoon's theory is still not quite what we want but she makes important arguments for the existence of weak distributivity that we should consider here.

2.2.2 Yoon's theory of the lexicon and weak distributivity

Yoon proposes that weak distributivity is not a property of plurals in general, but

rather a property of certain classes of predicates. For example, compare the two sentences in (38)-(39).

(38) The glasses are dirty
(39) The glasses are clean

In a situation where someone is setting a table for a formal dinner, she argues, (38) would count as true even if only a few of the glasses are dirty; but for (39) to be true it seems we would say that all the glasses have to be clean. The difference between (38) and (39), according to Yoon, is that *dirty* is a 'partial predicate' and *clean* is a 'total' predicate. Yoon argues that the distinction between total and partial predicates can be part of an explanation for the proportion problem in donkey sentences like (40)-(41).

(40) Most boys who had a baseball card in their pockets soiled it while playing in the mud.
(41) Most boys who had a baseball card in their pockets kept it clean while playing in the mud.

Suppose there are five boys, each of whom has three baseball cards in his pocket. The intuition about (40) is that it is true if three out of those five boys soiled at least one of their baseball cards. This is the 'weak' or 'existential' reading of a donkey sentence. On the other hand, (41) is true, it appears, only if at least three boys kept all three of their baseball cards clean; this is the 'strong' or 'universal' reading of a donkey sentence. Yoon points out that the weak and strong readings of the donkey sentences in (40)-(41) correspond to the weak and strong truth conditions we assigned to the ordinary sentences in (38)-(39).

She proposes that the weakness or strength of the sentence is due to a lexical difference between the predicate *dirty/soiled* and the predicate *clean*, and that these

predicates belong in two different lexical classes: total predicates and partial predicates. They are defined as follows (p.224):

(42) If P and Q are a pair of lexicalized antonyms, and it holds that
a. if $P(x) \& y \sqsubseteq x \rightarrow P(y)$
b. if $Q(x) \& x \sqsubseteq y \rightarrow Q(y)$
then P is a total predicate and Q is a partial predicate, where $\sqsubseteq$ is the semantically relevant part relation.

Since we are here considering only pluralities, and we take plural DPs to denote sets, the "semantically relevant part relation" can be considered just the $\in$ relation. Yoon is very specific that (42) requires that "total and partial predicates come as pairs of lexical entries." (p. 224) A predicate can be classified as total or partial only if it has a lexicalized antonym. Periphrastic negation does not count in determining total and partial predicates.

So Yoon proposes that the sentence *the glasses are dirty* only requires that some of the glasses be dirty because this 'weakness' is a lexical property of the partial predicate *dirty*. At this point, then, we have one proposal for implementing Williams' idea of 'weak distributivity'. However, as Yoon points out, this classification doesn't completely explain the possibility of nonmaximality (as we are calling it) because in fact most predicates do not come out as either total or partial predicates.

Yoon then goes on to discuss other factors that may influence whether a predicate receives a strong or a weak interpretation. She uses (43) and (44) to argue that episodic predicates are more likely to allow a weak interpretation than stative predicates.

(43) The children (who ate pizza here last night) got food-poisoned
(44) The children (who are playing in the garden) are eight years old.

The intuition is that there is a difference between (43) and (44) in that (43) less

strongly requires each individual child to have gotten food poisoning than (44) requires each child to be eight years old.

But while it is generally true that stative predicates yield a strong interpretation, there are episodic predicates that also seem to prefer a strong interpretation, as in (45).

(45) The children walked to school

Yoon attributes the strength of (44) and (45) to the fact that the VP is a change of state predicate. She finally concludes that there is a variety of factors that influence whether an episodic predicate is interpreted strongly or weakly, including lexical factors as in the total/partial distinction, and contextual factors.

I think Yoon is quite correct in arguing for the importance of context, but for Yoon these contextual factors are supposed to be influential only for non-total or non-partial predicates. (That is, only for the predicates that haven't already been lexically classified as total or partial.) That this might not be correct is already pointed out in a footnote in her article, in an example based on a suggestion from Manfred Krifka. The relevant case is the following.

Yoon classifies *open* as a partial predicate, in contrast to the total predicate *closed*.[5] But the idea that the partial/total distinction is just a lexical distinction is questioned by Krifka (1996), with an example like the following. Imagine there is a safe in a high-security building, behind three locked doors. If the outermost door is locked, but the inner

---

[5] The kind of example that motivates this classification of *open* is a sentence like (i) or (ii).

(i) If a man has a garage with a window, he keeps it open

(ii) Oh no, I left the windows open!

(i) can be true if most men have garages with two or three windows, but they leave only one open. (ii) is felicitous in a situation where someone is driving away from their home on vacation and they suddenly realize they left just two windows open in their house, even if their house has many more than two windows.

two doors are open, then we would not normally say *the doors are open*, even though *open* is classified as a partial predicate so we expect that this sentence could be true in this situation. In this case it seems that the context, namely a safe behind three big doors, plays a more important role in determining the strength or weakness with which the predicate is interpreted. But for Yoon if a predicate is classified as total or partial, it's not clear what role context can play (if any) in determining the predicate's "strength" or "weakness." This is perhaps in part because Yoon does not give an explicit proposal for how "contextual factors" play a role in determining the interpretation of a sentence.

Another reason why we won't be able to use Yoon's proposal to capture nonmaximality is that for her, strength or weakness is a lexical property. This doesn't help us for a number of cases we have been considering here. First,Yoon's own examples like *got food poisoned* and *built a raft*, and Williams' example *were hitting each other*, show that the strong/weak distinction needs to be statable on complex predicates, not just lexical predicates (see also Roberts 1987 for discussion on this point). But the proposal she gives doesn't give an explicit way to do that. This problem is even worse for cases like the distributive interpretation of *ate a sandwich*, because she doesn't make reference to any explicit quantifiers so something more would need to be said in order to get the one-sandwich-per-person reading.

Finally, I think that even the predicates that Yoon classifies as 'strong' or 'total' do not necessarily require true universal quantification over all of the parts of the plurality. For example, she cites the stative predicates *like jazz* and *be from Texas* as strong predicates. But I think both of the following sentences allow for exceptions, especially if

the group of children is large.

(46) The children like jazz
(47) The children are from Texas.

What this shows is that the distinction between the strong or universal reading on the one hand, and the weak or existential reading on the other, is not fine-grained enough to capture the somewhat vaguer effect that I have been calling nonmaximality.

Nevertheless, Yoon gives good arguments for a theory of plurality in which a notion of 'weak distributivity' is part of the semantics. It's just that it appears from the point of view of the data we are concerned with here that she doesn't go quite far enough. We need two things that her account doesn't provide: a quantifier (for non-total predicates), and an account of the role of context that is as explicit as her account of the role of lexical meaning in determining the truth conditions of a sentence with a definite plural.

In the next section we will take stock of what we have learned about maximality from the literature, and add some new observations of our own. Then we will look at a completely different kind of approach to nonmaximality, one which treats nonmaximality as a variety of collectivity. This will require that we learn a bit more about theories of plurals and collectivity.

2.2.3 Taking stock of nonmaximality

2.2.3.1 Nonmaximality is affected by lexical meaning

The first thing we have learned about nonmaximality, which we take from Yoon's work, is that the effect of nonmaximality is influenced by the predicate. To take just two of Yoon's examples, something about the lexical meanings of predicates like *got food*

poisoned and *went to school* makes the former allow nonmaximality more easily (and perhaps to a greater degree) than the latter.

2.2.3.2 Nonmaximality is affected by the size of the plurality

Another factor influencing nonmaximality, which we saw from our discussion of Fiengo and Lasnik's paper, is the size of the plurality denoted by the subject DP. If the set of people we are talking about is a large set (say, the set of men in a crowded bar), then it is easier to allow weakened truth conditions than it is for a small set of people (say, Bill, Sam, and Tom). The size of the plurality obviously depends on the context.

2.2.3.3 Nonmaximality affects collectives and distributives

A third point, which perhaps has not been stressed in the discussion so far, is that nonmaximality can occur equally well with (at least some) collective predicates as well as distributive ones. For example, we have been using the following example to show nonmaximality with distributive predicates.

(48) The boys ate a sandwich

But of course this sentence can also be interpreted collectively. And on its collective reading it also allows nonmaximality. If Adam, Bill, Chris, and Dan are the boys, then (48) can be true even if Chris didn't eat any part of the sandwich at all.

2.2.3.4 Nonmaximality is quantificational

A fourth, and very important point, is that the sentence in (48) shows that we need a quantifier in our representation, to account for the one-sandwich-per-boy reading. But we also need that quantifier to be able to be weakened appropriately, to account for the fact that ate-a-sandwich doesn't have to distribute down to every single boy. So it appears that

nonmaximality may be a quantificational effect; but it is worth pointing out that it is not obvious how to reconcile this idea with the fact that nonmaximality is also possible with collectives, which at this point, by hypothesis, involve no quantification. *2.2.3.5 The form of the DP interacts with nonmaximality*

Here I want to add some new observations about plurality and nonmaximality that I will add to those we have just seen. First, it appears that the linguistic form of the DP can affect how easy it is to get nonmaximality. For example, compare the sentences in (49)-(50).

(49) The girls ate a sandwich
(50) Alice, Betty, Carmen, and Diane ate a sandwich

I think that it is much harder to allow nonmaximality for the second sentence than for the first sentence, even if we take the two subject DPs to be coreferent. The reason seems to be that if we went to all the trouble of mentioning Alice, Betty, Carmen, and Diane by name, it is hard to exclude them.

However, it is not completely impossible to exclude people named in a conjoined DP. The following example from Lasersohn (1990:47) shows that with a rich enough context nonmaximality can apply even in these types of cases. I quote Lasersohn at length:

> Imagine a comptetition in which teams are required to attempt various stunts, including lifting a piano. John and Mary form one team, Bill and Susan form another. During the competition, John lifts the piano; meanwhile Mary performs one of the other stunts, say shooting herself out of a cannon. When Bill and Sue's turn arrives, they succeed in doing almost all the stunts that John and Mary did, but fail at lifting the piano, and therefore lose the competition. In this sort of situation, it seems fair to say that John and Mary won the competition because THEY lifted the piano, while Bill and Sue didn't. This is despite the fact that Mary played no role in the actual lifting.
>
> In other words, given Lasersohn's scenario, the sentence below can be judged true,

precisely because the fact that John and Mary form a team is so salient.

(51) John and Mary lifted the piano

So even though DPs that are conjoined proper names tend not to allow nonmaximality, in the right context this tendency can be overridden. This is somewhat similar to the case of Yoon's partial and total predicates, where we found that there, too, a rich enough context could lead to the overriding of the semantics we expect for a partial or a total predicate.

2.2.3.6 All *is incompatible with nonmaximality*

In this chapter we have not specifically discussed *all* very much but in our discussion of nonmaximality we should recall that *all* is the reason we are concerned with nonmaximality in the first place. We saw in Chapter 1 that *all* has a maximizing or totality effect. Another way to think about this, from the point of view of the discussion in this chapter, is to say that *all* is incompatible with nonmaximality; it's a kind of nonmaximality-canceller.

2.2.3.7 "Maximality" is an implicature

Finally, we can apply standard tests for implicature versus entailment to sentences with definite plurals, and when we do, we find that the purported maximality of definite plurals comes out as an implicature, rather than an entailment. For example, the "universal" force of plural definites is cancelable without contradiction. (Grice; see also Levinson 1983, Horn 1989) Compare (52) and (53) (both allow either a collective or a distributive reading) with the infelicitous (54).

(52) The boys ate a sandwich, but Adam didn't/but not all of them.
(53) The girls gathered in the auditorium, but Mary was in the bathroom /but not all of

them.

(54) #John managed to solve the problem but he didn't solve it.

The alleged universal force is also reinforceable without redundancy (Sadock 1978).

(55) The boys -- in fact, all the boys -- ate a sandwich
(56) The girls -- in fact, all the girls -- gathered in the auditorium
(57) #John managed to -- and in fact did -- solve the problem

If the apparent maximality of these sentences is a conversational implicature, this suggests that maximality/nonmaximality has a pragmatic character.

Recall that we saw above that Kroch (1975) argued from the infelicity of (58) that definite plurals really are universally quantified.

(58) #Although the townspeople are asleep, some of them are awake.

However, if in fact the universality of quantification over *the townspeople* is an implicature, we wouldn't expect that the construction in (58) should be able to cancel it. Only certain locutions can be used to cancel implicatures (I thank Veneeta Dayal for pointing this out to me). (59) is equally awkward in its attempt to cancel the "not-all" implicature of *some.*

(59) #Although some people are asleep, all of them are asleep.

2.2.4 Desiderata for a theory of plurality and nonmaximality

So from our discussion so far, we can extract two important desiderata for a theory of plurality and nonmaximality, which are given in (60)

(60) **Desiderata for a theory of plurality and nonmaximality**:
a. The theory should make use of a quantificational operator
b. The theory should make room for both lexical and pragmatic factors

In crafting a theory of nonmaximality, we need to keep these desiderata in mind.

However, the mainstream view of nonmaximality in the literature on plurality does not take these ideas as central; it takes a very different approach to nonmaximality.

2.3 Collectivity in theories of plurals

In the many different works on plurality cited at the beginning of this chapter, a rather different approach to nonmaximality from the one suggested by Yoon's and Williams' work prevails. Instead of proposing a kind of weak distributivity, these theories treat nonmaximality as essentially a variety of collectivity. I'll call this type of approach to nonmaximality the 'groups approach'.[6] Since on this approach nonmaximality is a subcase of collectivity, rather than a weakened form of distributivity, to see how this type of approach works we need to return to the issue of the distributivity/collectivity distinction.

Recall that we have proposed that the difference between (61) and (62) is the presence (61) or absence (62) of a D operator on the VP.

(61) George and Sandra left
(62) George and Sandra are a happy couple.

But this has been something of a simplification. There are more issues surrounding D operators and how they are used on the one hand, and collective readings (ie, absense of a D operator) and when they are licensed on the other, that we have not yet discussed. To look more closely at these issues, we need to look more closely at the literature on

---

[6] Landman used the term 'group' in a very technical sense, but not all proponents of the groups approach to nonmaximality believe in Landman's theory of Groups. So to distinguish the two I will write Groups whenever I mean Landman's technical term and groups when I am referring to a type of approach to nonmaximality. It is understandable if the reader finds this to be less-than-perspicuous. But in the large body of work on plurals, most of the terms that could be used to refer to plurals, including *plurals, sums, pluralities, bunches,* and *groups* have already been used in a very specific way by one author or another.

plurals, and in particular at the papers in which the issues of distributivity and collectivity are discussed.

2.3.1 *Link*

I'll begin with an influential paper by Link (1983). In it he proposed that modeling plurality using lattice structures would capture several important properties of sentences with plural terms. The phenomena he is concerned with are first, the distributive/collective distinction, which we have already seen. Secondly, there is the property of cumulativity; this is the fact that (63) and (64) together entail (65).

(63) George and Ringo are pop stars
(64) John and Paul are pop stars
(65) John, Paul, George and Ringo are pop stars

Finally, Link proposes that a desireable semantics of plurals should capture the similarity between plural terms and mass terms because mass terms also show cumulativity, as shown in the entailment from (66) to (67).

(66) The stuff in my cup is tea, and the stuff in your cup is tea.
(67) The stuff in my cup and your cup is tea.

To capture all of the above, Link proposes that a lattice structure be used to model plural and mass terms. There are two main reasons for doing this.

The first is that in traditional treatments of plurals (Link cites Bennett 1975 in this regard) there is a distinction between singular terms, which denote objects in a model, and plural terms, which denote sets of objects. However, in natural language the vast majority of predicates freely take both singular and plural arguments, so it would appear that if we treat plurals as sets, we would have to complicate the grammar with two entries for every predicate (one for singular or one for plural) or with some sort of operation that will

transform singular predicates into plural predicates. (But note that Scha (1981) deals with this problem in a set-theoretic model by proposing that individual terms denote a singleton set, so both singular and plural terms denote sets in his model.) In a lattice structure, singular terms denote individuals, and plural terms denote sum individuals: both are of the same logical type. So, Link argues, we are not forced to make a distinction that appears to be unwarranted.

Secondly, and less importantly for our purposes, the use of lattices allows us to model the semantics of plurality and the semantics of mass terms with the same system, thus capturing the similarities between them.

Link incorporates lattice structures by requiring that the domain of individuals be a complete join semilattice. That means it consists of a set of individuals E which is closed under a summing operation +.[7] The summing operation combines two individuals, say a and b, and yields a third individual, a+b, which can be read, "the sum of a and b". The important point for Link is that a, b, and a+b are all the same type of object in the model. Link also proposes a partial ordering relation, $\leq$, [8] and a plurality operator on one-place predicates, *. To see what all this does for us, consider just a small part of a model as is shown in (68). (68) E= {a, b, c, a+b, a+c, b+c, a+b+c}

⟦girl'⟧ = {a, b, c}

The star operator * can apply to the denotation of *girl* to yield a plural denotation,

---

[7] Really $\oplus$ for Link; he defines several different sorts of sum operations, but the differences between them don't concern us here.

[8] Really $\leq_i$ for Link, because the relation we want here is the individual-part relation. Link distinguishes between i-parts, m-parts (mass parts) and "the material part-whole relation". Since we are only concerned with i parts here I will leave off the subscript.

ie, the denotation of *girls*. When it applies to a one-place predicate it yields a set that contains the individuals in the singular denotation plus all the sums of individuals that can be formed therefrom. This will yield (69) as the denotation of *girls*.

(69) $[\![*girl]\!] = \{a, b, c, a+b, a+c, b+c, a+b+c\}$

Finally, the $\leq$ operator is in some sense the inverse of the + operator, because the following holds: if $x \leq y$ then $x+y = y$. To use the examples from our model, $a+ (a+b) = a+b$, and $a \leq a+b$. In other words, the sum of the individual a and the individual a+b is again a+b.

The $\leq$ operator is also useful in establishing a formal distinction between individuals that are sums, and individuals that are 'singular' in the common-sense notion of the term. Although we are treating plural and singular expression as denoting the same sort of entitity, we will still sometimes want to distinguish the two. So we will call a 'singular' (in the everyday sense of the term in which I am a singular entity and a pencil is a singular entity, etc) individual an 'atomic' individual. An atomic individual is defined as follows.

(70) An individual x is *atomic* iff $\forall y[y \leq x \rightarrow y=x]$

Link uses the lattice structure of the plural and singular domain to define a supremum operator, σ, which captures the meaning of the definite description. The σ operator picks out the maximal individual in a predicate's denotation; this is defined as the individual x such that for some predicate P, every individual y in P is such that $y \leq x$. Where the extension of the predicate is a singleton, σxPx will pick out the unique individual in the extension of P. Where the extension of the predicate is plural, σx*Px will

pick out the maximal individual in the extension of *P: in the case of *the girls* in our model above, this would give us the individual a+b+c. (The idea of treating the definite determiner as a maximality operator was already proposed by Sharvy 1980.) Note that in the model above, *the girl* fails to denote because there is no object in the extension of *girl*' of which every other object in the extension is a part.

In his 1983 paper, Link proposes that not all plural predicates should be translated with the * operator: only lexically distributive predicates should be. So, for example, a lexically distributive predicate like *die*, when applied to an argument that is a plural individual (henceforth: a sum), is starred, as in (72).

(71) The pigs died
(72) *die'(σx*pig'(x))

However, a predicate like *gather* takes a sum argument but is not distributive, so (73) is translated as in (74).

(73) The children gathered in the hall
(74) gather'(σx*child'(x))

This distinction between * and no * corresponds to the semantics we have been using so far, in which distributive predicates have a D operator and collectives have no D. But Link's original proposal about D operators was a bit different from the use I have put it to here.

For Link, the D operator was originally proposed to account for distributivity on a predicate that is not lexically distributive like *die*, but ambiguous between a distributive and a collective reading; we've seen examples of these like *build a raft* and *eat a sandwich*. Another example is *carry the piano upstairs.* This predicate can be applied to

an atomic individual or to a sum; (75) can mean either that John and Bill each carried the piano upstairs in separate events, or that they carried it together.

(75) John and Bill carried the piano upstairs

Since *carry* is ambiguous in this way, we cannot represent the predicate *carry* with a * operator, because it is not lexically distributive. Link instead introduces the D operator to mark distributivity in these cases. It can be applied to a predicate like *carry the piano* to yield a distributive reading. So on its collective reading, (75) is translated as (76), and on its distributive reading, as (77).

(76) carry'(j+b)
(77) ^{D}carry'(j+b)

At this point we have a model which includes two types of individuals: plural individuals (sums) and atomic individuals. In later work, Link (1984) proposes that we need a new kind of individual in the model: something that's made out of a plural individual, but that behaves like an atomic individual. He proposes to capture this by introducing the idea of a 'group': a group is a sum that has been 'closed off''. It is derived from a sum, but it is treated as an entity that has no parts; that is, as an atomic individual. The type of sentence that motivates this analysis is given in (78).

(78) George and Martha and Nick and Honey hate and humiliate each other.

This sentence can mean that George and Martha on the one hand, and Nick and Honey on the other, are couples who hate and humiliate each other. The point is that the 'hating and humiliating each other' is distributed down to the separate couples. This is not captured by either of the following translations that we so far have available to us, in (79) and (80).

(79) hate.e.o.'(g+m+n+h)
(80) ^{D}hate.e.o'(g+m+n+h)

The expression in (79) means that all of the four individuals are in the 'hate.each.other' relation, but this is not so in the scenario Link is concerned with (although it is of course possible); only the couples are in that relation. (Note that a partition analysis of *each other* would solve this problem for us, but Link doesn't consider this possibility; but this is essentially what Schwarzschild 1994, 1996 will propose later.) On the other hand, the expression in (80) means that George hates each other, Martha hates each other, etc., which is nonsense.

To solve this problem, Link introduces the idea of a "group". The operator that takes a sum into a group is represented as↑ (this is actually Landman's 1989 symbol but I adopt it here for convenience). A group is "made out of" a plural individual (ie a sum) but its part structure is 'erased' by the ↑ operator. Thus a group behaves like an atomic individual: it has no parts for a distributivity operator to operate on. This means that we can translate (78) as (81).

(81) ^{D}hate.e.o'(↑(g+m)+↑(n+h))

The translation in (81), Link argues, represents the meaning of (78) that we are after. The property of hating each other distributes down to the group of George and Martha, and the group of Nick and Honey.

So now we have a model in which a plural term can denote either a sum or a group. Recall that we translated (82) as (83).

(82) The children gathered in the hall
(83) gather'(σx*child'(x))

Now that we have groups in our model, we expect that there should be another 'reading' of (82) available, namely (84).

(84) gather'(↑(σx*child'(x))

In other words, there should be an ambiguity between a collective reading (which concerns a sum) and a group reading (which concerns a group). But it's not clear that this ambiguity really exists, and this forms part of the basis for Landman's (1989) criticism and revision of Link's theory.

2.3.2 Landman's theory: Groups

Landman (1989a) argues that the distinction between collective and group readings is not well motivated, and proposes that all collectivity be represented in terms of Groups. (Here I will begin using Groups with a capital *G* to refer both to Landman's theory and to the formal object in his theory.) He furthermore argues that the distinction between the * and the ^{D} operator is not well-motivated for the following reason. Link introduces the * operator to capture cumulative inference, but for any sentence that is distributively true, cumulative inference also holds; distributivity and cumulativity are essentially "two sides of the same coin". Link's system, with both a pluralization operator and a distributivity operator, fails to capture this generalization.

To solve these problems, Landman introduces a system of plurality where predication of a sum is always distributive predication, and predication of a Group is always collective predication. This means that a sentence with a distributive reading is always translated as predication of a sum to a starred predicate, and a sentence with a collective reading is always translated as predication of a Group to an unstarred predicate.

So the two readings of (85) are translated as (86) and (87) in Landman's system.

(85) John and Bill carried the piano upstairs
(86) *carry'(j+b)
(87) carry'(↑(j+b))

To make the proposal clear, it is perhaps worth pointing out that for Landman, the expression in (88), which is consistent with our earlier proposal (section 2.1) for representing a collective reading, is ill-formed. Verbal predicates only have individuals or groups in their extensions, and j+b is a sum.

(88) carry'(j+b)

So Landman's system is one in which the distinction between collectivity and distributivity is rigidly codified: only a certain kind of NP can be interpreted distributively (namely, a sum), and it must be predicated of a certain kind of VP (a starred predicate). Likewise, only a certain kind of NP can be interpreted collectively (a Group), and it, too, must be predicated of a certain kind of VP (a non-starred predicate).

Landman also argues that the use of lattices is not crucial in modeling the meaning of plurals. It is possible to give the set-theoretic equivalents of the crucial lattice-theoretical operations that Link uses (for example, + is set union, and ≤ means subset-of or element-of). He further argues that the class of lattices that can always be mapped onto a set-theoretic equivalent is a subset of the possible lattice structures, and it is just this class of lattices that can be used to model the semantics of plurals. This fact, he argues, is more naturally explained by using sets to model the plural domain, rather than using lattices and placing a stipulative restriction on the kinds of lattices that can be used.

The only thing we lose by using sets to model plurals is the similarities to the mass

domain, because while sets are made up of atomic elements, mass terms do not have elements of the same sort. But Landman says that he is 'reluctant' to accept this as a knock-down argument that we must use lattices. (And see Chierchia 1995 for recent work in which the semantics of mass nouns is represented set-theoretically.) Since he is concerned in his paper only with plurals, he essentially sets aside the mass domain as a separate issue. We will follow Landman in this respect, and of course we have already been treating the plural domain as set-theoretic rather than lattice-theoretic.

Now we have Landman's proposal for representing the collective/distributive distinction, so we can return to the other problem we began this discussion with, the nonmaximality problem. In a model of plurals that includes Groups, can we use Groups to handle nonmaximality?

2.3.2.1 Nonmaximality on the groups approach

In a later paper Landman (1996) makes very explicit arguments that we can and should use Groups to handle nonmaximality. His system of Groups and distributivity already treats collectivity and singular predication on a par: both are cases where an atomic individual is the subject of a predicate.

Landman argues that this similarity accounts for nonmaximality. For example, consider the sentences in (89)-(90).

(89) John touched the ceiling
(90) The boys touched the ceiling.

(89) will be considered true even if John just reached up and touched the ceiling with his hand (in fact this is perhaps the most likely scenario in which it is true). It is not necessary that John's entire body be in contact with the ceiling; just one part of him,

namely his hand, will do.

Now consider (90) in a scenario in which the boys form a human pyramid so that the topmost boy can reach the ceiling. In this scenario it seems the sentence can be judged true. Landman argues that the way we verify the sentence as true in this situation does not have to do with distributing the predicate *touch* down to one or more individuals who make up the collection of boys. Instead, he argues, just like part of John's hand touching part of the ceiling is sufficient to make (89) true, part of *the boys* touching part of the ceiling is sufficient to make (90) true. In other words, nothing special needs to be said about plurality here; in particular we don't need to invoke distributivity. Collective predication can be thought of as "collective body" formation, and a collective body behaves semantically just like a physical body.

Another example that Landman discussess is due in part to examples discussed by Lasersohn (1990). Lasersohn discusses cases where we are willing to give "team credit" to a group of individuals for some action that is actually performed by just one, or just some, of the individuals in that group. We saw one case earlier with John and Mary lifting the piano. Another example of this is given below.

(91) The gangsters killed the stool pigeon.

It is possible that the stool pigeon was shot by just one of the gangsters, but nevertheless we are willing to say that this sentence is true by virtue of that action, because the gangsters act as a team.

This type of approach to nonmaximality is not unique to Landman's proposal. Other authors, who do not use Landman's type of Groups, propose a similar approach. For

example, the proposals of Lasersohn (1990, 1995) and Schwarzschild (1996) have in common that plural noun phrases may denote only one type of entity and the distributive/collective distinction is represented by the presence or absence of a D operator on the VP. (Note that for Lasersohn plurals always denote groups, although they're not the same as Landman's Groups, and for Schwarzschild plurals always denote sums.) In fact this is just the basic theory of distributivity/collectivity that I sketched earlier in section 2.1. Both Lasersohn and Schwarzschild propose that there are some kinds of cases where nonmaximality should be treated as a kind of collectivity (that is, plural predication without distributivity), so in this respect they are in agreement with Landman.

The groups approach to nonmaximality can be summarized as follows. A plurality (whether it be a Group, a group (Lasersohn) or a sum (Schwarzschild)) is a type of individual, and it can be the subject of a predication. When a plurality is the subject of a predicate, we expect that whatever inferences follow from that should be attributable just to the nature of the predicate and the fact that the group is plural. No appeal to distributivity is needed for these cases. Nonmaximality is essentially just a subcase of collectivity.

2.3.2.2 Problems with the groups approach to nonmaximality

The groups approach to nonmaximality clearly does not meet at least one of our desiderata for a theory of nonmaximality: it does not involve quantification.

In this section I will criticize the groups approach to nonmaximality on three different grounds, including points that are related to our desiderata of sections 2.2.3-2.2.4. I criticize Landman's proposals in the most detail, because he has offered the most detailed

arguments in favor of the groups approach. But the arguments given here apply generally to the groups approach, not just to Landman's theory.

The first problem is that we find, if we look at an array of examples that are "collective" (ie nonmaximal) that we have to make contradictory assumptions about whether or not collective predication is just like singular predication.

Since on the groups approach nonmaximality is supposed to be the result of collective predication, and since collective predication is a type of singular predication, we expect that nonmaximality effects will have the same distribution as "part of" effects in singular predication. Landman argues that the 'touch the ceiling' case shows this to be so. But there are examples where this prediction is not met. Consider the sentences in (92).

(92) a. The soldiers of the 4th platoon were captured by the enemy
b. Bill was captured by the enemy

The (a) sentence in (92) can be true if one or two of the soldiers managed to get away without being captured; in other words, it exhibits nonmaximality. But (b) shows no such part-of effect. This is unexpected if a "collective body" is, semantically speaking, the same thing as a "physical body", as was claimed earlier.

Other examples that seem to show a difference between singular and collective predication are the following.

(93) a. Jane knows the answers to these questions
b. Jane knows the answer

Imagine a context for the (a) sentence in which we are looking at a list of a dozen questions. Then the sentence can be considered true even if Jane doesn't know the answers to one or a few of them. But on the other hand, suppose that the context for (93)b is a long

complicated mathematical problem that has several subparts. (93)b can only be true if Jane knows the answer to the entire problem including all of its subparts.

Similar problems arise for the examples in (94)-(95). The (a) examples contain a singular argument that can plausibly contain salient "parts" (panes in a window, parts of a long exam) but the sentence seems not to allow the possibility of applying the predicate to just some of these parts. These contrast with the (b) examples, where nonmaximality is possible.

(94) a. The explosion shattered the window
b. The explosion shattered the windows

(95) a. Polly graded the exam
b. Polly graded the exams

Now, Landman considered this type of counterexample to the groups approach, and argues that they are not really counterexamples. Landman argues that the differences between the (a) and the (b) sentences do not force us to abandon the idea that collective predication is the same as singular predication. The differences between the (a) and (b) sentences, he says, are attributable "to the differences in the relation between [singularities] and [their] parts and collections and their parts." (p.430)

But this contradicts the earlier claim that "collective bodies" are just like "physical bodies". Even if he did not intend that particular claim to be fully general (for example, is a math problem with many parts a "physical body"?), the contradiction in the system is still there. We begin by claiming that collective predication is exactly the same as individual predication; in particular, we claim that the part-of relation between individuals and collections is analogous to the part-of relation between atomic individuals and their

"parts" (in the sense that my arm is part of me). We find examples where this appears to be difficult to maintain. We then claim that collective predication is not exactly like individual predication, after all; the parts of a plural individual are a little bit different from the parts of an atomic individual. The "part of" notion at work here must simultaneously treat the parts of pluralities like the parts of individuals; and treat the parts of pluralities differently than the parts of individuals.

Notice that on a weak distributivity theory of nonmaximality, the difference between plural individuals and atomic individuals is distributivity. So a theory of weak distributivity would do some work for us in distinguishing the (a) cases from the (b) cases.

The second problem is that the groups approach doesn't help us capture nonmaximality in the cases where we have evidence that a universal quantifier is taking scope over an existential, as in the familiar (96).

(96) The boys ate a sandwich

It is of course possible to interpret (96) distributively (one sandwich per boy) and at the same time allow the sentence to be true even if, out of a large group of boys, Billy and Jimmy didn't eat a sandwich because they weren't feeling well. The groups approach to nonmaximality would tell us that (96) is translated as (97).

(97) ate.a.sandwich'(↑(the.boys'))

This sentence can only be true in a scenario where the boys shared a single sandwich. As it stands, the groups approach cannot predict the fact that (96) allows for there to be many sandwiches, but not necessarily a sandwich for every single boy. This is precisely the kind of case that led us to conclude that a theory that captures

nonmaximality must include a quantificational component.

Finally, the groups approach to nonmaximality cannot help us with our original problem with *all*, which was brought to light in Chapter 1. Recall that the problem with *all* is explaining its maximizing effect with both distributive and collective predicates. Since on the groups approach nonmaximality is just a species of collectivity, there is no way in this approach to distinguish between a nonmaximal collective reading and a maximal collective reading. Both are simply represented as collective predication. But *all* makes exactly this distinction: it is compatible with a maximal collective reading, and incompatible with a nonmaximal collective reading. Since the groups theory gives us no way to distinguish between the two, it won't help us discover the meaning of *all*.

So I conclude that the groups approach to nonmaximality is not a good way to handle nonmaximality. It does not seem to capture the things we need a theory of nonmaximality to capture.

In Chapter 3 we will develop a theory of nonmaximality that does meet our desiderata. That theory will be an adaptation of Schwarzschild's "generalized distributivity", which crucially for us involves distributivity with a pragmatic component.

Chapter 3

Capturing Nonmaximality and the Meaning of *all*

3 Introduction

In Chapter 2 we discussed the phenomenon of nonmaximality in detail and argued that the following two desiderata should be met by a theory of the interpretation of definite plurals in order to capture nonmaximality: the theory should have a quantificational component, and a pragmatic component. Then we looked at the groups approach to nonmaximality and found that it didn't adequately account for the full range of phenomena associated with nonmaximality, mostly because it treated nonmaximality as the absense of quantification (ie, a kind of collectivity).

In this chapter we will develop a theory of distributivity that captures the features of nonmaximality discussed in Chapter 2. We will then go on to see how this theory of distributivity makes possible a novel idea about the meaning *all* and *both*, and show that this proposal about *all* and *both* captures a wide range of facts about the meaning and distribution of these floating quantifiers.

3.1 A criticism of Groups (not groups)

In section 3.2 we will adopt a variation of the theory of plurality proposed by Schwarzschild (1996). But in order to appreciate the use that we can put that theory to, we need to first reexamine a debate over the use of Groups (Landman 1989) to represent the semantic "ambiguity" of DPs between collective and distributive readings.

In Landman's account, the dividing line between distributive and collective

predication is expressed both on the NP and on the VP; distributive predication is predication of a sum to a distributive VP, and collective predication is predication of a Group to a non-distributive VP. The motivation for introducing this structure is to handle examples like the following (the examples are from Schwarzschild 1992; but note their similarity to the examples we saw from Link's work in Chapter 2, section 2.3.1, ie. *George and Martha and Nick and Honey hate and humiliate each other*.)

(1) The cows and the pigs are separated
(2) The young animals and the old animals are separated
(3) The animals are separated

Suppose that all of the animals in our discourse are just the cows and the pigs, and among them are old and young cows, and old and young pigs. Then if (1) is true, it seems we don't want to say that it follows that (2) is true (although we probably do want to say that (3) follows). This is someting we would like our semantics to capture.

If we treat conjunction as set union (which is the set-theoretic equivalent of the sum operation familiar from Link's proposal), then without groups the two noun phrases *the cows and the pigs* and *the young animals and the old animals* are in this context coreferent: they each are coreferent with the noun phrase *the animals*. If *the cows* denotes a set, and *the pigs* denotes a set, then the union of these sets is the set containing all the animals. But of course this is also the denotation of *the young animals and the old animals*, by the same reasoning. This means that in fact the unwanted inference would be predicted to go through: (1) would entail (2).

Groups are the firewall that prevents this inference from going through, because the union of the Group of cows and the Group of pigs is just the set containing the Group

of cows and the Group of pigs. This is, of course, not the same set as the set containing all the animals, and it is not the same set as the set containing the Group of old animals and the Group of young animals. So if we have Groups in our system, the truth of (1) does not license the inference that (2) is true, which seems to be the right result.

However, Schwarzschild, in his criticism of Landman, points out that there are sentences where this sort of inference does go through, though this is not predicted if, as the Groups approach proposes, *the cows and the pigs*, *the young animals and the old animals*, and *the animals* are not coreferent. For example, (4) does seem to entail (5) and (6).

(4) The cows and the pigs filled the barn to capacity
(5) The young animals and the old animals filled the barn to capacty
(6) The animals filled the barn to capacity

To capture the inferences here, we need to introduce type lifting operations that will allow the (here) collective predicate *fill the barn*, which applies to first-order Groups as in (4), to apply to higher-order Groups as in (5) and (6).[1] Furthermore, we need a condition that tells us that if the predicate can truthfully be applied to the Group *the animals*, then the higher-order versions of the predicate should truthfully apply to higher-order Groups that are formed from the same individuals as *the animals* is.

It should be clear that the point of introducing these sorts of rules is to eliminate the distinction between the Groups in question in (4)-(6). But if we need to eliminate the distinctions in some cases, we might wonder whether the distinctions were well-motivated

[1] The Groups denoted by the DPs *the cows and the pigs*, and *the young animals and the old animals* are higher-order because they are themselves composed of Groups.

to begin with. Schwarzschild argues that they are not. He argues that even the cases that were originally supposed to motivate the existence of Groups turn out to make an argument against them.

For example, if we return to the examples we saw at the beginning of this section, Schwarzschild argues that while (7) does not apparently entail (8), it does entail (9).

(7) The cows and the pigs were separated
(8) The young animals and the old animals were separated
(9) The animals were separated

Although (7) is a more informative statement than (9), since it provides information about how the animals were separated, it is nevertheless true that (7) entails (9). This means that we also need to introduce a type-shifting operation that would lower the type of *separate* so that it can apply to the Group *the animals*, and furthermore we need a condition that says that if the original form of *separate* can be truthfully applied to *the cows and the pigs*, then the lowered form of *separate* will truthfully apply to any lower type of Group that is made out of the same individuals as the higher-order group *the cows and the pigs* is ultimately composed of. Again, we need to obliterate the distinctions that Groups impose, but in this case we are going 'down' in types, rather than up.

At this point the motivation for Groups looks rather weak, and Schwarzshild argues that we should abandon the use of Groups in our semantics. But still we need to account for the examples that originally motivated Landman's use of groups. An explanation for these will be given in the next section.

3.2 Schwarzschild: generalized distributivity

Schwarzschild (1996, 1994, 1992) proposes that the distributivity/collectivity

distinction doesn't work in quite the way we have been assuming. Up until now we have been assuming that a predicate is either interpreted distributively, in which case the property it denotes is distributed down to each atomic individual in its plural subject, or collectively, in which case there is no distributivity. But Schwarzschild, following Higginbotham (1980), Gillon (1987) and others, points out that there are predicates that can be true on a reading that we can call 'intermediate' distributivity — that is, distributivity down to subpluralities of the plural subject but not "all the way down" to atoms. Furthermore, and crucially, the relevant subpluralities are pragmatically determined.

Here the idea of a cover, introduced in our discussion of reciprocals in Chapter 2 (section 2.2.1), will prove useful. Covers give us a formal way to make the D operator range not only over atoms of a plurality, but subpluralities of a plurality. Schwarzschild argues quite persuasively that the relevant cover that the D operator will range over is determined in large measure by the context of utterance: whatever subpluralities are salient in a discourse, those will make up the "cells" of the cover. Let's first look at how Schwarzschild's proposed semantics works, and then come back to the arguments in favor of using covers as part of the meaning of the D operator.

Schwarzschild calls his D operator *Part*, for partition, and he proposes that the Part operator is always accompanied by a context-dependent domain selection variable he calls *Cov*, because the value assigned to the variable always takes the form of a cover of the universe of discourse. Because a sentence can contain more than one Part operator, each with its own *Cov* variable, Schwarzschild assumes that each *Cov* variable may carry a

different index (and see also the discussions of Westerstahl 1985 and von Fintel 1994 on quantifiers and their domain restrictions).

For now I will adopt Schwarzschild's use of the term Part but it is important to keep in mind that the Part operator is just a kind of D operator.

Let us see how the Part operator and *Cov* work together in the interpretation of a sentence with the following example.

(10) The boys are hungry

In translation, Part applies to the plural predicate *be hungry*, just the way a D operator does, yielding (11).

(11) $(\text{Part}(\text{Cov}_i)(\text{hungry}'))(\text{the.boys}')$

The meaning of the Part operator is quite similar to the meaning of the D operator. (Because we always interpret expressions relative to a model *M* and an assignment function *g*, I will leave off these superscripts.)

(12) $\text{Part} =_{df} \lambda P \lambda x \forall y[y \in \text{Cov}_i \,\&\, y \subseteq x \,\&\, \rightarrow P(x)]$

Using this definition of the Part operator, (11) would be interpreted as in (13).

(13) $\forall x[x \in [\![\text{Cov}_i]\!] \,\&\, x \subseteq [\![\text{the.boys}']\!] \rightarrow x \in [\![\text{hungry}']\!]]$

Note that the universal quantifier contributed to (13) by the Part operator has two conditions in its restriction: *x* must be a subset of *the boys* (with the D operators we have previously seen, *x* is required to be an element of *the boys*) and *x* must be an element of the cover assigned to the variable *Cov*.

In order to evaluate the truth conditions of (13), we must assign a value to *Cov*, and to do that we have to have a universe of discourse to refer to. A universe U and some

possible covers of the set of singularities of U is given in (14).

(14) U = {a, b, c, s, t, {a,b}, {a,c}, {a,s}, {a,t}, {a,s,t,}...}

⟦the.boys'⟧ = {a,b,c}

I = {{a}, {b}, {c}, {s,t}}
J = {{a}, {c}, {b,s,t}}
K = {{a,b,c}, {s,t}}
L = {{a,b}, {c,s,t}}

Suppose the context assigns the value I to Cov_i in (13). Then the sentence will be interpreted exactly as it would be on a Link or Landman-style semantics with a D operator, because since each boy occupies a singleton set of the cover assigned to Cov_i, each boy is asserted to be in the extension of *be hungry*. Here I follow Schwarzschild in assuming that an individual and the singleton set containing that individual are indistinguishable from the point of view of natural language semantics. (He calls this idea Quine's Innovation after Quine 1980; see Schwarzschild 1996 for details.)

So assignment of a cover like I to Cov_i in (13) leads to what we have been calling a distributive interpretation of the sentence. Note that we must assume that lexical semantics plays some role in constraining the felicity of certain types of covers. For example, if we take *be hungry* to be a predicate that only applies to atomic individuals, then we must assign a Cov that has singleton cells, or else we would have a sentence that would be doomed to be false (because there are no sets of boys in the extension of *be hungry*, only individual boys). Likewise, as we will see shortly, the fact that mixed-extension predicates allow singleton or non-singleton cells is information that comes from the predicate itself: a predicate like *eat a sandwich* can apply to a plurality or to a singleton.

3.2.1 Generalized distributivity and context-sensitivity

At this point our theory of plurality with a Part operator and a Cov variable doesn't do anything for us that our old D operator didn't do. But the cases that motivate the presence of *Cov* are cases where subpluralities, not just atomic parts, of a plurality are relevant. We can illustrate this with an example from Gillon (1987:212) (cited in Schwarzschild 1996).

(15) The men wrote musicals

Suppose that *the men* denotes Rodgers, Hammerstein, and Hart. Gillon points out that the sentence above is true by virtue of the fact that Rodgers and Hammerstein wrote musicals (together) and Rodgers and Hart wrote musicals (together). This can be captured if we suppose that the universe of discourse looks something like (17), and we choose a value for Cov_i like the one given in Q. (Suppose that r, h, t are Rodgers, Hammerstein, and Hart, respectively.)

(16) (Part(Cov_i)(wrote.musicals'))(the.men')
(17) U = {a, b, c, r, h, t...}
Q = {{a,b,c},{r,h},{r,t}..}

If the value assigned to Cov_i in (16) is Q, then the sentence is correctly predicted to be true. Note also that this example shows that we want to use a cover of the plurality, rather than a partition; recall that the difference is that a partition requires the sets that make it up to be nonoverlapping.[2]

Gillon's example gives us some evidence that we want to use covers to constrain

[2] In nearly all the examples that I discuss I use nonoverlapping covers to keep things simpler, but overlapping covers are in principle permitted.

the domain of the Part operator, and Schwarzschild goes on to give even more. Moreover, Schwarzschild's examples show that the pragmatic context of utterance plays an important role in establishing what the value of Cov_i should be. It is important to point out here that "context" includes non-linguistic context as well as linguistic context, and we will see examples where both kinds of information is relevant. First, here is an example where non-linguistic context plays the crucial role, which I will quote at length (from Schwarzschild 1996:67):

> Imagine a situation in which two merchants are attempting to price some vegetables. The vegetables are sitting before the merchants, piled up in several baskets. To determine their price, the vegetables need to be weighed. Unfortunately, our merchants do not have an appropriate scale. Their grey retail scale is very fine and is meant to weigh only a few vegetables at a time. Their black wholesale scale is coarse, meant to weigh small truckloads. Realizing this, one of the merchants truthfully says:

(18) a. The vegetables are too heavy for the grey scale *and*
b. The vegetables are too light for the black scale

In the situation described above, (18)a is false on its "pure" distributive reading, where we apply *too heavy for the grey scale* to each vegetable. It is true on its "pure" collective meaning, but that is not what the merchant means. On the other hand, (18)b is false on a pure collective meaning but true on a distributive reading, but again, this is not what the merchant means.

It is clear that both sentences are true because in this context we are distributing the predicates *too heavy...* and *too light...* down to subpluralities of vegetables, according to "the physical arrangement of the vegetables in baskets". (p. 67) The theory proposed by Schwarzschild provides a way to capture that: we use a distributivity operator (Part) and choose a value for *Cov* in which the cells of the cover correspond to the baskets full of

vegetables.

To see an example where linguistic information tells us what the relevant subpluralities are, we can return to our examples involving reciprocal predicates, repeated below.

(19) The cows and the pigs are separated
(20) The young animals and the old animals are separated
(21) The animals are separated

Suppose for ease of exposition that *separate* has the following truth conditions: *separate(P)* is true of a plurality P just in case for all relevant subparts x, y that are part of P, x is separated from y and y is separated from x. Then it should be clear that the three subject DPs in (19)-(21), although coreferent, suggest different sets of relevant subpluralities. (19) suggests that the relevant subpluralities are the cows and the pigs, so these are the sets that we take to be separated from each other when we hear that sentence. (20) suggests instead that the relevant subpluralities are the young animals and the old animals. And finally, (21) is just not linguistically explicit about the intended subpluralities. Thus we can explain the data that was used to motivate the existence of Groups by appealing to context rather than inventing new semantic types.

Let us sum up what we have seen in this section. The theory of plurals proposed by Schwarzschild retains the idea of a D operator (in the form of the Part operator) but makes two innovations. First, in this theory, distributivity doesn't have to be all the way down to atomic individuals. Borrowing from the literature on reciprocals (which we looked at in some detail in Chapter 2), Schwarzschild introduces the idea that the D operator has a resource domain variable whose value is a cover of the domain of

discourse.[3] Since distributivity no longer is required to look just at atomic individuals, this theory is called the generalized distributivity theory. Secondly, Schwarzschild argues that pragmatics plays an important role in determining just what entitites the predicate is supposed to be distributed down to.

3.2.2 Generalized distributivity and collectivity

The generalized distributivity theory also offers a new way of looking at collective predicates. Since our Part operator has *Cov* in its restriction, it is possible to interpret even collective predicates with a Part operator. If the value assigned to *Cov* includes a cell that is equivalent to the set denoted by the subject DP, then the sentence will have a collective reading.

For example, consider the predicate *ate a sandwich*, which as we have seen can be interpreted collectively or distributively. In the generalized distributivity theory, the sentence in (22) can be translated as (23), and interpreted as in (24). Our covers from (14) are repeated below.

(22) The boys ate a sandwich
(23) (Part(Cov_i)(ate.a.sandwich'))(the.boys')
(24) $\forall x[x \in [\![Cov_i]\!] \& x \subseteq [\![the.boys']\!] \rightarrow x \in [\![ate.a.sandwich']\!]]$

$[\![the.boys']\!] = \{a,b,c\}$

I = {{a}, {b}, {c}, {s,t}}
J = {{a}, {c}, {b,s,t}}
K = {{a,b,c}, {s,t}}
L = {{a,b}, {c,s, t}}

If the context assigns I to Cov_i in (23), the sentence will be interpreted distributively, and

[3] For more discussion of resource domain variables, see section 3.2.4.1.

for exactly the same reason as (13), above, was.

On the other hand, if the context assigns K to Cov_i, then the sentence will assert that the set containing the boys is in the extension of *ate a sandwich*; in other words, it will assert that there was a single collective eating of a sandwich.

We might wonder whether this is really desireable. (23) with K assigned to Cov_i of course turns out to be equivalent to *ate.a.sandwich'(the.boys')*, because the D operator quantifies over a singleton domain (the domain contains just one object: the set {a,b,c}). It is debatable whether natural language allows quantification over a singleton domain. Consider a party where there is only one man in attendance. Is it felicitous to say *every man at this party failed chemistry*? Schwarzschild himself does not take a strong stand on this matter; from the point of view of the data he is concerned with, if we took away quantification over a singleton domain the meanings of those sentences could still be represented without the D operator.

For now I will assume that quantification over a singleton domain is possible, for reasons that will become clear shortly. But we will have reason to revisit this issue in Chapter 4.

Another aspect of the generalized distributivity theory that bears discussing is the proposal that the value of *Cov* has to be a cover of the whole domain of discourse. We might wonder whether *Cov* could be a cover of the set denoted by the subject DP. But Schwarzschild shows that the value assigned to *Cov* does indeed have to be a cover of the domain of discourse with the following example. Again I quote (p.76):

> Apparently, in the last five years, an unsavory Mr. Slime has made several purchases from a computer store: 4 computers and 1 cartonful of diskettes. These

purchases were made over the course of a few years and each time, Mr. Slime paid an initial amount in counterfeit currency and the remainder he paid for with a valid credit card. The following remark is entered in the police report:

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

In this example the plural *the computers* is being distributed over (each computer was paid for in two installments) and the plural *the diskettes* is being interpreted collectively (the entire carton was paid for in two installments). Since the second VP is elided, we assume that it must be identical to the first VP; but the first VP has a *Cov* with singleton cells corresponding to each computer, and we need the second VP to have a *Cov* with a nonsingleton (ie collective) cell corresponding to the set of diskettes. This is not a problem if we assume that *Cov* is a cover of the whole domain of discourse: a single value for *Cov* can contain both the singleton cells (corresponding to the computers) and the nonsingleton cell (corresponding to the diskettes).

So in the generalized distributivity theory the domain of quantification of the D operator is crucially influenced by the context of utterance. Consider now our earlier discussions of the problem of nonmaximality. We concluded in Chapter 2 that nonmaximality involves quantification, and it is influenced by pragmatics, and these are the two main ingredients of the generalized distributivity theory. This suggests that we can use the generalized distributivity theory to handle nonmaximality, and in the next section I will propose that it gives us a very simple way to instantiate the notion of "weak distributivity."

2.3.3 A potential problem

The generalized theory of distributivity makes it possible to give an explicit and

general account of nonmaximality that meets the desiderata proposed in Chapter 2. The basic idea will be to let the pragmatics influence the meaning of the D operator in such a way that the domain of the D operator is smaller, hence 'weakening' its quantificational force. Since this theory involves a weakening of the quantificational force that is caused by pragmatics, I will rename the phenomenon of nonmaximality, and now call it "pragmatic weakening."

The idea was actually inspired by an objection to the generalized distributivity theory raised by Lasersohn (1995). So to show how this will work, we will first discuss his objection.

Recall that the sentence in (26) will be interpreted as (27), and furthermore that the values assigned to *Cov* are covers of the universe of discourse, not just the DP denotation.

(26) The boys are hungry
(27) $\forall x[x \in [\![Cov_i]\!] \ \& \ x \subseteq [\![the.boys']\!] \rightarrow x \in [\![hungry']\!]]$

U = {a, b, c, s, t, {a,b}, {a,c}, {a,s}, {a,t}, {a,s,t,}...}
$[\![the.boys']\!]$ = {a,b,c}

I = {{a}, {b}, {c}, {s,t}}
J = {{a}, {c}, {b,s,t}}
K = {{a,b,c}, {s,t}}
L = {{a,b}, {c,s,t}}

It is possible that we might assign the cover J to Cov_i in (27). The difference between J and I (the value we assigned earlier) is that in J Bill does not occupy a singleton cell: he is in a cell with the two non-boys, Sue and Tina. Call this an ill-fitting cover, because it is ill-fitting with respect to the set of boys — there is no set of cells whose union is equivalent to the set of boys.

A consequence of assigning this type of cover to Cov_i in (27) is that the semantics

in some sense 'doesn't care' whether Bill is hungry or not. Since the set {b,s,t} is not a subset of the set {a,b,c}, there is no cell containing Bill that satisfies the restriction of the quantifier. The sentence can come out true whether Bill is hungry or not.

Lasersohn objects that this is highly undesirable, on the grounds that it means that *John and Mary went to school* could come out true even if Mary stayed home, a somewhat counterintuitive result. In response, Schwarzschild argues that to allow such a choice of cover would be so uncooperative of a speaker as to be "pathological," and suggests that while it might be possible to formulate a rule that prevents the choice of such a cover, it is more plausible to simply assume that pathological covers are ruled out on general principle.[4]

I agree with Schwarzschild that we should assume that speakers are cooperative in their intended choice of covers, but wish to suggest that ill-fitting covers are not necessarily so 'pathological' as either he or Lasersohn believe them to be. We quite commonly find ourselves in circumstances where it is not necessary to be precise down to each and every individual, as we saw in great detail in Chapter 2. So perhaps in precisely these circumstances ill-fitting covers do some work for us — they allow us to be a little bit vague. In the next section I will show how this works.

3.2.4 Capturing nonmaximality

I propose that we adopt the generalized theory of distributivity, and allow for possibility of ill-fitting covers. This is essentially a way to extend the arguments

[4] Recall that to the extent that he discusses nonmaximality, or pragmatic weakening as we are now calling it, Schwarzschild adopts the general approach that I discussed above as the groups approach.

Schwarzschild gives for the interaction of pragmatics with distributivity so that we capture pragmatic weakening as just another type of domain selection effect. I'll argue here that this gives us just the right results with respect to nonmaximality (pragmatic weakening) effects.

Consider the case that was most difficult for the groups approach to nonmaximality, repeated here as (28).

(28) The girls ate a sandwich

On the groups approach, if the sentence is interpreted distributively then we predict it should be equivalent to *every girl ate a sandwich*, and if it is interpreted collectively then we allow for pragmatic weakening but we expect that it should only have a reading where the girls are sharing a single sandwich. This approach does not predict that (28) can be both distributive and allow for pragmatic weakening. But in the generalized distributivity approach we expect that some pragmatic weakening should be possible.

Consider the sentence in (28) in a situation where *the girls* are Alice, Betty, and Carmen. Then (28) is interpreted as in (29) , and a possible assignment to Cov_j is the cover shown in (30).

(29) $\forall x[x \in [\![Cov_j]\!] \,\&\, x \subseteq [\![\text{the.girls}']\!] \rightarrow x \in [\![\text{ate.a.sandwich}']\!]]$

(30) {{a}, {c}, {b,m,n}}

Then the sentence will be true if Alice ate a sandwich, and Carmen ate a sandwich, and it 'doesn't care' in some sense whether Betty ate a sandwich or not. This is exactly what we want; a universal quantifier ensures that we have a one-sandwich-per-girl reading, but its scope over the set denoted by *the girls* is "weakened" by the ill-fitting

cover, so it doesn't have to be the case that every single girl ate a sandwich.

Note that we can capture pragmatic weakening with collective predicates with exactly the same mechanism: (28) can be interpreted collectively, and nonmaximally, by assigning the cover in (31) to *Cov*.

(31) {{a,b},{c,s,t}}

If (31) is assigned to Cov_j in (29) then the sentence (28) can be judged true, even if Carmen didn't participate in eating the sandwich that Alice and Betty ate. Here we see why I am assuming that quantification over a singleton domain is possible: it allows us to capture pragmatic weakening effects with collective predicates using the same mechanism that we used with distributive predicates.

So we have a theory of distributivity that interacts crucially with pragmatics in such a way that the universal force introduced by distributivity is weakened just a bit by the pragmatics. Note also that since collective readings in this theory can involve a D operator, the weakening of both collective and distributive predicates is treated by the same mechanism. Both kinds of weakening turn out to be domain of quantification effects.

3.2.4.1 Domain-of-quantification effects

It is perhaps worth pausing here to look at the general phenomenon of contextual domain of quantification effects. Von Fintel (1994) argues in favor of the idea (see also Westerstahl 1985, and references cited in both works) that the use of (covert) resource domain variables to restrict quantification is pervasive in natural language. To take one very simple example, consider the sentence in (32).

(32) Every chicken laid three eggs last week

The domain of quantification of *every* in this sentence is ostensibly the common noun *chicken*. But *chicken* denotes every chicken in the world (and every chicken in every possible world, if we consider its intension), and it is highly unlikely that (32) is really about every chicken in the world.

It is clear that an utterer of sentence (32) has in mind some contextually salient set of chickens, perhaps the chickens that he owns. In order to capture this, we can introduce a resource domain variable, C, which is just the set of contextually relevant things. By intersecting C with *chickens*, we get a set of contextually relevant chickens, which is what we understand to be the relevant domain of quantification in (32).

(33) $\forall x[x \in [\![\text{chicken}']\!] \ \& \ x \in C \rightarrow x \in [\![\text{l.3.e.l.w.}']\!]]$

Von Fintel argues that we find contextually-licensed resource domain variables wherever we find quantifiers, including determiner quantifiers and adverbial quantifiers (which he argues quantify over situations).

If this is true, then it leads us to expect that we should find a resource domain variable with the D operator as well, since it introduces a quantifier. In fact on this view of quantification, it would appear to be quite odd if the D operator didn't have a resource domain variable of some sort, as is the case in every account of distributivity that I know of except the generalized distributivity proposed by Schwarzschild.

But then the question naturally arises: does the D operator have a variable like C, an unstructured set of contextually relevant things, or a variable like *Cov*, which is a set with some internal structure? Or does it even have both?

I think there are good reasons to say that the D operator has a single resource domain variable, and that variable is *Cov*. The starting point for this argument is to say that we have good evidence already, from Schwarzschild's work, that we need a variable like *Cov* to account for distribution down to contextually-relevant subgroups of a plurality. While von Fintel's work leads us to expect that a quantifier should have some resource domain variable, it doesn't force us to conclude that that variable must be C. (Indeed, the value of the variable appears to depend on the type of quantifier, because for adverbial quantification the resource domain variable has as its value a set of situations.) So if we were to suppose that a D operator has some variable in addition to *Cov*, we should show that it does some work for us.

C certainly couldn't do the same work in distributive quantification that it does in determiner quantification as in (32). As we have seen, the determiner *every* takes a common noun like *chicken* as its first argument. The function of *C* in (33) is to reduce the set that *chicken* denotes into a more specific, contextually relevant set. But since distributive quantification in contrast is always over a definite plural, we already have a specific, contextually relevant set: the set that the definite plural refers to.[5] So we don't need *C* for this purpose; it wouldn't appear to do any work for us in this regard.

Another way to look at the function of *C* is that it eliminates things from the

---

[5] Note that a sentence like (i) shows that pragmatic weakness is independent of the process of identifying the referent of a definite DP.

(i) The girls went swimming, and then they played basketball

In (i) we understand *the girls* and *they* to refer to the same set of girls. But pragmatic weakening can apply to each VP independently so that the "actual participants" in swimming do not have to be the same as the "actual participants" in playing basketball.

domain of quantification. Could it be useful there? If we could use *C* to eliminate objects from the domain of quantification, this could constitute another possible mechanism leading to pragmatic weakening (as pointed out to me by Angelika Kratzer, p.c.). However, given that ill-fitting values of *Cov* will do this same work for us, and that we have better motivation for the existence of *Cov*, we can assume that the D operator does not make use of a variable like C and we would lose nothing from our system.

In short, the differences between determiner quantification and distributive quantification lead us to expect that the resource domain variable of the latter would be of a different type than the resource domain variable of the former. A resource domain variable like *C* wouldn't be very useful in distributive quantification, but we have plenty of evidence that a variable like *Cov* is. In addition, the generalized distributivity theory, which makes use of *Cov*, fits nicely into the picture of natural language quantification in which quantifiers always come with resource domain variables. I conclude that *Cov* is the context-dependent variable that is associated with the D operator.

3.2.5 More on the notion of an ill-fitting cover

3.2.5.1 The "junkpile"

From the discussion so far we have concluded that in order to get a "pragmatically weakened" reading of a sentence like (34), we have to have a salient cover that lumps one of the girls in with one of the non-girls.

(34) The girls ate a sandwich

It is not difficult to imagine this kind of context. Suppose that a large shipment of unpainted jeeps, cars, and trucks was delivered to a factory where they can be painted. At

the end of the day, the vice-president who shipped the vehicles comes to the factory to check on the status of his job. He talks to the manager, who brings him out to the parking lot where they can see several rows of freshly-painted vehicles, lined up neatly. Off to the side and very obvious (ie, salient) are a car and a truck, still wearing their drab-brown coats of primer. In this context, the factory manager can truthfully say (35).

(35) The jeeps were painted blue, the cars were painted red, and the trucks were painted green. But we didn't get to those two over there.

In this context it would appear quite likely that the unpainted car and truck are in a cell together. This would have the effect that the value assigned to *Cov* would be ill-fitting both with respect to ⟦the.cars'⟧ and with respect to ⟦the.trucks'⟧, which correctly predicts that (35) counts as true in this context, because pragmatic weakening is permitted.

However, I don't believe that it is crucial to have this kind of cell in the cover. If we take this same context, except instead of a truck and a car off to the side, we see just a car over there, it is still possible to say (35). But now what is the cell that the car has been cordoned off to, to ensure that it doesn't enter into the scope of the D operator? Note that it is not enough for the car to be in a cell by itself — if it were it would still fall into the scope of the D operator. The car that we want to exclude must be in a cell with non-cars.

For this case, and in fact for most cases of ill-fitting covers, I think the cell that non-participants can be lumped into is a kind of pragmatic "junkpile".[6] The junkpile can be thought of as a cell in the cover that contains all those objects which are present in the

[6] Thanks to Roger Schwarzschild for the catchy name!

universe of discourse, but which we are not concerned with at a particular moment. It seems reasonable to suppose that pragmatically salient covers do not necessarily structure the entire universe, that speakers may employ a junkpile as a kind of shortcut.

So far so good. We might want to allow for the existence of a pragmatic junkpile. But on the other hand, in the case I described above, that one primer-colored car is not being simply ignored — it is being ignored in virtue of its salience. Is this a contradiction? I don't think so.

We already know from our extensive discussion of nonmaximality, both here and in Chapter 2, that objects in discourse can be "saliently ignorable". Perhaps the best example of this is Lasersohn's scenario in which we can say that John and Mary carried the piano up the stairs in virtue of the fact that John actually did the feat by himself. In this case Mary becomes saliently ignorable. If the junkpile is the set of things that we are ignoring, then there isn't any contradiction in saying that an object can become salient enough to be ignored. We can imagine that in virtue of Mary's salience as part of a team with John, a speaker can relegate her to the junkpile.

So what we have already seen is that there is plenty of evidence that it is possible to ignore someone (or something) in virtue of their salience, given that other kinds of contextual conditions are met (such as the notion of a team). Here I am proposing that the junkpile is the mechanism that we use to do it, in many cases.

3.2.5.2 "Mindreading" and context-sensitive variables

Although I use very specific values for *Cov* in my illustrations, I do not suppose that speakers and hearers are always aware of the values to such a high level of precision.

I think it is more accurate to say that speakers and hearers are in some sense guessing at the sort of cover that the other has in mind.

We see examples of this in everyday life all the time. Suppose we are in a retail clothing store and a manager is giving instructions to one of the stock clerks. They are looking at a shipment of four boxes of sweaters that just arrived this morning. One of those four boxes is set off to the side, in a space usually reserved for boxes that the manager plans to deal with himself (perhaps they are special items). The manager says to the clerk the sentence in (36).

(36) The sweaters that came in this morning go in aisle three.

At this point the clerk might be slightly puzzled. *The sweaters that came in this morning* denotes the set of sweaters contained in the four boxes. But he knows that the manager might have in mind that the fourth box of sweaters, the one off to the side, should be treated differently than the others. The clerk is forced to guess at what sort of cover the manager has in mind. Note that if the clerk asks for clarification, he would most likely do it by asking a question like the one in (37).

(37) Do you mean <u>all</u> the sweaters that came in this morning?

As long as we allow that speakers will sometimes have in mind a value for *Cov* that is ill-fitting, then a sentence will allow for pragmatic weakening. The value of *Cov* is determined contextually, but speakers are constantly engaged in guessing about what context the other speaker takes him or herself to be in. This is true not only for the *Cov* variable of a distributivity operator; I take it that it is true for pragmatically-determined domain-of-quantification variables in general (see von Fintel 1994).

What is important is that we allow for the possibility of ill-fitting covers. Then the process of figuring out what value one's conversational partner has in mind will mean that one has to leave open the possibility of an ill-fitting cover, and therefore of pragmatic weakening. Thus I am claiming that part of the effect of pragmatic weakening comes not only from the fact that a speaker may choose an ill-fitting cover, but also from the fact that her conversational partner must guess at what the speaker has in mind.

2.3.5.3 Context and covers (again)

Two more points are worth emphasizing before we go on to *all*. First is that I do assume, like Schwarzschild, that there is a general principle prohibiting "pathological" use of ill-fitting covers. The difference is that I think it is wrong to say that every possible ill-fitting cover is necessarily pathological; as we have seen in great detail here, sometimes ill-fitting covers can be useful.

Secondly, the same kinds of factors that we have already seen can make one or another cover salient can influence whether or not pragmatic weakening is allowed; this suggests that we are correct in treating pragmatic weakening with ill-fitting covers.

For example, we saw previously that it is more difficult, although not impossible, to get pragmatic weakening with a plural DP that consists of conjoined DPs. This is an example like Lasersohn's *John and Mary went to school*. But notice that it was another example of Lasersohn's, *John and Mary lifted the piano*, that showed that it is not impossible to get pragmatic weakening in these cases.

These types of examples are very similar to the ones discussed by Schwarzschild, e.g. *The cows and the pigs are separated, The young animals and the old animals are*

separated. There Schwarzschild argued that mention of *the cows* and *the pigs* makes them so salient that the most natural interpretation of *the cows and the pigs are separated* is that the separation is according to species.

We see the very same phenomenon here. The form of the DP influences the type of cover one takes as salient. This is most likely due to some kind of Gricean reasoning, along the following lines. If a speaker goes to the trouble of mentioning each of the individuals by name, it is unlikely that they would have in mind a cover that excludes one of those individuals. A more cooperative thing to do, under most circumstances, is to simply not name that individual. Still, as Lasersohn's example shows, this is true only in most circumstances, not in all circumstances. A rich enough context can change the Gricean calculus.

3.3 A Proposal for the Semantics of all

At this point we can finally return to the problem that we closed Chapter 1 with: how to capture *all*'s 'maximizing' or 'totality' effect, even with collective predicates. In light of our current account of nonmaximality or pragmatic weakening, in this section I will propose that the 'maximizing' effect of *all* is essentially an 'anti-weakening' effect. We will see that one consequence of this approach is that *all* does not have any quantificational force of its own. We will also see that this analysis of *all* offers a fairly straightforward way to capture the idea proposed in Chapter 1, that *all* is a modifier.

Let us return to the contrast that we concluded Chapter 1 with.

(38) The boys are hungry
(39) The boys are all hungry

The difference between these two sentences is that (39) is a stronger statement than

(38). A distributivity theory of plurals without pragmatic weakening does not explain this. However, given the proposal I made for capturing pragmatic weakening in the previous section, we would say that the source of the relative weakness of (38) is the fact that it allows at least for the possibility of pragmatic weakening.

In contrast, let us say that *all*'s contribution to (39) is to rule out the possibility of pragmatic weakening. Since the cause of pragmatic weakening is (the possibility of) an ill-fitting cover, we propose that the function of *all* is to disallow the choice of an ill-fitting cover. Or, another way to say this is that *all* requires a good-fitting cover, where good fit is to be understood as the opposite of ill fit.

Let us define a relation **good fit** between a cover and a definite DP denotation (i.e. a set). A cover is a good fit for this set if there isn't any element or member of the set that's stuck in a cell with some non-members. Another way to say this is that the cover is a good fit if every element of the set is in a cell of the cover that is a subset of that set. We can define this formally as in (40).

(40) **Good fit**: For some cover of the universe of discourse *Cov* and some DP denotation X, Cov is a good fit with respect to X iff $\forall y[y \in X \rightarrow \exists Z[Z \in Cov \ \& \ y \in Z \ \& \ Z \subseteq X]]$

The function of *all* in a sentence will be to ensure that the value assigned to *Cov* is a good fit with respect to the subject DP. Recall from Chapter 1 that we are stipulating, for the present, that floated *all* is always construed with the subject. Our task here is to show how, given the connection between *all* and that DP, *all* will ensure that the value assigned to *Cov* is a good fit with respect to that NP.

All does this by eliminating from the set of contextually salient covers any cover

that isn't a good fit. So *all* is essentially an operator on the set of contextually salient covers. This means that the assignment function that assigns a value to *Cov* will have only good-fitting covers (with respect to some particular DP) to choose from, thus guaranteeing that the value of *Cov* will be a good fit.

The denotation of a DP with *all*, then, will be the same as a DP without *all*. The same will be true for floated *all*; the denotation of a VP with floated *all* is the same as the denotation of that same VP without *all* (we will make some minor revisions to this in Chapter 5). But to keep things a bit more perspicuous in our examples, I will use the superscript ^{gf} on a DP whenever that DP is antecedent to *all*, so that we can tell which interpretations are interpretations of sentences containing *all*.

Let's return to our sample universe and sample values of *Cov* to see how this works. Consider the sentence in (41), which is interpreted as in (42).

(41) All the boys ate.

(42) $\forall x[x \in [\![Cov_i]\!] \ \& \ x \subseteq [\![the.boys^{gf}]\!] \rightarrow x \in [\![ate']\!]]$

U = {a, b, c, s, t, {a,b}, {a,c}, {a,s}, {a,t}, {a,s,t,}...}
[⟦the.boys'⟧] = {a,b,c}

I = {{a}, {b}, {c}, {s,t}}
J = {{a}, {c}, {b,s,t}}
K = {{a,b,c}, {s,t}}
L = {{a,b}, {c,s,t}}

If *all* interacts with the context to limit the set of possible values for Cov_i in the way I have described, then J and L are not possible values for Cov_i in (42). They have been eliminated as possibilities by the presence of *all*. This leaves I and K as possibilities. (There are in principle others but I am restricting our attention to keep our discussion

simple.) If we assume, as we discussed earlier, that lexical semantics can play a role in determining which cover will be appropriate, then I will be assigned as the value of Cov_i in (42). Hence the D operator will have true universal quantificational force over the set denoted by *the boys*.

This explains the "maximizing" effect of *all*, and hence the difference between (38) and (39). (38) can be true, in some contexts, if a speaker is willing to allow assignment of an ill-fitting value to *Cov*. But (39) will not be true unless every boy ate because the presence of *all* ensures that an ill-fitting cover cannot be assigned to Cov. Now let us see how it will work for collective predicates. Because I have adopted from the generalized distributivity theory a mechanism for collectives that is the same as for distributives, except for the value assigned to the variable *Cov*, this hypothesis about the meaning of *all* can be applied to collective predicates straightforwardly.

As we have already seen, the interpretation of (43) is (44).

(43) All the boys built a raft
(44) $\forall x[x \in [\![Cov_i]\!] \& x \subseteq [\![the.boys^{gf}]\!] \rightarrow x \in [\![built.a.raft']\!]]$

Because *all* eliminates all ill-fitting covers from consideration, only K and I are possible choices for Cov_i in (43).

(45)
K = {{a,b,c},{s,t}}
L = {{a,b},{c,s,t}}
I = {{a}, {b}, {c}, {s,t}}
J = {{a}, {c}, {b,s,t}}

Thus we expect that (43) should be interpretable either collectively or distributively, but in either case, "maximally." (43) cannot be true in a situation where Chris, for instance, did not participate in the raft building. The same mechanism that

accounts for the meaning of *all* with distributive predicates accounts for the meaning of *all* with collective predicates.

3.3.1 Linguistic forms and context

The proposal I have made about *all* here makes some assumptions about context, and about the interaction of language and context, that are probably nontrivial. In this section I will just briefly explore some of the most obvious implications of the proposal, noting with humility that it deserves a fuller treatment than I can give it here.

For instance, I must assume that the 'context' of a conversation includes not just the set of all the individuals in the context (ie, the set U from above) but also a set of covers of that set. This assumption was already implicit in Schwarzschild's formulation, and I am exploiting it a bit more here. To make it more convenient to talk about this set of covers, we will give it a name: let us call this set U_c, because it is a set of covers of U.

I have claimed that *all* performs an operaton on U_c, which is that it shrinks U_c so that it includes only covers that are good-fitting with respect to a particular DP. One very obvious question that arises is, how long in the discourse does this effect last? If eliminating covers from U_c is something like eliminating worlds from a context set, then we would expect that those covers should be more-or-less permanently eliminated from U_c (or at least as long as the conversation lasts).

But it appears that eliminating covers from U_c is not like eliminating worlds from a context set. Consider the sentences in (46), as part of a discourse.

(46) a. The girls all went outside.
b. They went swimming, and played basketball.

If the ill-fitting covers had been "permanently" eliminated from U_c, then we should

understand from (46) that every girl went swimming, and every girl played basketball. But I don't think that this is the case. It appears that the effect of *all* 'wears off' at the end of a sentence.

I can imagine two ways to capture this fact. One is to suppose that U_c is in some sense "constructed" anew for each sentence. Another is to suppose that the effect of *all* is not so much to eliminate covers from U_c, but to make some covers more or less salient.

I think the second option looks more promising, because the notion of saliency is already central to the use that we're putting covers to in this theory. A cover is just a way of organizing U according to the salience of things in U. We already know from VP-conjunction that we have to suppose that there is more than one cover available in a discourse; so here we can suppose that *all* raises a certain class of covers to discourse prominence so that the assignment function has to choose one of those covers.

It is perhaps worth pointing out that, based on the evidence from ambiguous predicates like *build a raft*, we cannot suppose that *all* fixes the value of *Cov* — it only constrains it. Consider again a sentence like (47).

(47) The girls all built a raft

We know that this sentence doesn't allow pragmatic weakening, but *all* doesn't dictate whether the sentence should be interpreted collectively or distributively. There is still a role for context to play here.

A similar proposal is made for conditional clauses by von Fintel (1994), in which he argues that conditional clauses constrain, but do not determine, the domain of quantification of an adverbial quantifier. His proposal is not fully worked out (as he

himself points out) but because of its similarity to the present proposal we will look at a sketch of it here. Consider a sentence like (48).

(48) (If the sun shines, we play soccer on Sundays, but) If it rains we always stay home.

The standard analysis of (48), following Lewis (1975) and others, is that the conditional clause serves as the restrictor to the quantificational adverb *always*. One way of instantiating this is to say that *always* is a quantifier over situations, where the tripartite structure of (48) is something like (49) (s is a variable over situations).

(49) $\forall$ s (it rains in s)(we stay home in s)

Von Fintel argues that this semantics appears to be too strong, for several reasons.[7] It is also inconsistent with claims he makes elsewhere in the dissertation that the restriction of adverbial quantifiers comes from the context, specifically from the discourse topic. But an alternative, which is to treat the conditional clause as though it is the discourse topic (we might suppose that it is an overt reminder of what the topic is; this was proposed by Haiman 1978), appears to be incorrect also.[8]

---

[7] One reason is that it leads us to expect that the relationship between an adverbial quantifier and its domain restriction is as strong as the relationship between a determiner quantifier and its domain restriction. But adverbial quantification and determiner quantification differ in several ways. The hallmark way in which they differ is that the relationship between an adverbial quantifier and its restriction is relatively "loose": the restriction doesn't even have to be overt, as shown in (i), and when present there are several syntactic ways to realize it, and positions where it may appear, as shown in (ii) and (iii).

(i) Bill always takes the bus
(ii) When he goes to work, Bill always takes the bus
(iii) Bill always takes the bus to work

[8] Von Fintel argues that this idea, that conditional = topic, cannot be correct because conditional clauses can be focused, as shown in (i) (the example is from Givon 1992).

(i) A: Under what conditions will you buy this house?
B: I'll buy this house if you give me the money

The conditional clause in (B) represents new information, therefore it is not the topic but the focus. Since conditional clauses cannot be simultaneously topic and focus, we have to abandon the idea that conditionals are just a "spelling out" of the topic.

Von Fintel proposes a middle ground, in which the restriction of the adverb of quantification is constrained, but not determined, by the proposition provided by the *if*-clause (which denotes a set of situations). He implements this by intersecting the set of situations denoted by the *if*-clause with the contextually-determined domain variable of the adverbial quantifier (which presumably gets its value from the topic, which is also a set of situations[9]). This way, the semantics of conditionals with adverbial quantifiers is similar to adverbial quantification in general, but also the *if*-clause does some work for us in determining the domain of the quantifier.

Von Fintel's idea and the proposal I have made here are similar in that both propose that overt linguistic material partly, but not completely, constrains the set of possibilities that are made available by the context. But von Fintel's implementation is a bit different from mine. I have proposed that *all* interacts directly with the context to constrain the set of possibilities, and then the contextually-determined domain variable Cov can be assigned a value only from this new, restricted set of possibilities. Von Fintel proposes that the quantifier takes the information provided by the contextually-determined domain variable and the information provided by the *if*-clause and intersects them. For von Fintel, the *if*-clause interacts with the quantifier, not with the context. For me, *all* interacts with the context, not with the D operator.

So von Fintel's proposal raises the question of whether *all* interacts directly with the D operator, or with the context (that is, with U_∂). It seems that there should be

[9] Actually, I'm simplifiying a bit here. A topic is a set of propositions, which is a set of sets of situations. We need to assume, as does von Fintel (p. 41), that we can take the union of the sets in the topic, which is then a set of situations.

consequences for choosing one possibility over the other, but I want to defer exploring those consequences for a later time.

3.3.1 All *and* any

The idea that *all* signifies a 'strength' of quantification has some interesting points of commonality with two recent proposals about *any*, one by Kadmon and Landman (1993) and another by Dayal (1995, and to appear).

Kadmon and Landman give an analysis of *any* in which several of its properties, including its distribution and its use as a negative polarity item and a free-choice item, follow from its interpretation. Their analysis has two points that are relevant here.

First, Kadmon and Landman propose that *any* doesn't have quantificational force of its own. Instead they argue that an NP with *any* is an indefinite NP. The apparent quantificational force of *any* comes from two sources that have been independently shown to be associated with indefinites: the existential force of negative polarity *any* comes from existential closure, and the universal force of free-choice *any* comes from generic quantification.

Similarly, here I have proposed that *all* does not contribute any quantificational force of its own: the quantification comes from VP distributivity that has been independently proposed by many authors working on the semantics of definite plurals.

Secondly, Kadmon and Landman argue that the contribution of *any* to a sentence is to indicate "reduced tolerance of exceptions" (p. 356). It does this by performing the operation of 'widening' on the domain of quantification. The idea is that *any* indicates we should consider a wider domain of quantification, and in so doing will make the statement

with *any* stronger than the corresponding statement without *any*.

The proposal I have made here for *all* is quite similar. The anti-weakening effect of *all* is quite similar to the notion of reduced tolerance of exceptions, and the mechanism employed to account for these notions are similar as well. Anti-weakening is accomplished by the good fit requirement imposed by *all*. The effect of requiring a good fit is that the domain of quantification of the D operator is as "wide" as it can be with respect to the subject DP.

Dayal (1995, to appear) argues against Kadmon and Landman's proposal. Specifically, she argues that *any* is a universal quantifier. She proposes that in addition to introducing universal quantification, the semantic contribution of *any* is to signal the speaker's lack of commitment to the existence of individuals in the domain of quantification. Although she dispenses with the idea that *any* doesn't have quantificational force of its own, the idea that *any* supplies information about the speaker's intentions for the domain of a quantifier is quite similar to the present proposal about *all*.

Of course, many years ago, looking at a different set of data (*all* and *any* in generic contexts) Vendler (1967) argued that *all* and *any* are more similar to each other than either is to *each* and *every*. I take it to be an argument in favor of current approaches to *any* and the proposal made here for *all* that these proposals bear out what Vendler had observed in a pre-theoretic way.

3.4 Extending the proposal to both

As I discussed in Chapter 1, the distribution of *all* and *both* are very similar, giving us reason to suppose that they share a common component of their meanings. So I will

propose that *both* does the same thing that *all* does: it eliminates from U_c any covers that are not a good fit with respect to the DP it is construed with. Of course, *both* also carries the requirement that we are talking about exactly two individuals; this is traditionally treated as a presupposition and I will adopt that position here.

This works without any further conditions for floated *both* because the subject is a definite DP, as in (50), and for some uses of prenominal *both*, as in (51). But it doesn't work for (52), because *girls* is not a DP.

(50) The girls both went to the movies
(51) Both the girls went to the movies
(52) Both girls went to the movies.

As we saw in Chapter 1, the phrase *both girls* in (52) acts like a definite description, even though there is no definite determiner present (if I am right about *both/all* being modifiers). For now I do not have an explanation for this fact, but I will assume that *both* works in this case the same way it works for the other two cases; that is, it shrinks U_c so that it includes only good-fitting covers.

If *both* imposes a good fit requirement, then we can capture the difference between the sentences in (53)-(54). (53) will be true only if each of the students drove to school, whereas (54) can be true even if only one student drove and the other one was a passenger.

(53) The students both drove to school
(54) The students drove to school

In general, though, it is harder to find contexts with a plurality of two in which pragmatic weakening is felicitious. For example, Anna Szabolsci (p.c.) has pointed out to me that it appears that (55) cannot be true if only one of my parents is tall. This would seem to suggest that the function of *both* isn't really to rule out pragmatic weakening, since

(55) doesn't seem to allow pragmatic weakening anyway.

(55) My parents are tall
(56) My parents are both tall.

I think the difficulty in getting pragmatic weakening with (55) is due to the fact that we are dealing with a plurality that is a small group of just two individuals, and an individual level predicate. (For a different kind of explanation for the difference between (55)-(56) see Schwarzschild 1994.) We saw in Chapter 2 that both factors (the size of the plurality and the type of predicate) independently make it harder to get pragmatic weakening, so putting two hard-to-weaken factors together in one sentence makes it very difficult indeed to get pragmatic weakening. But I would argue that it is in principle possible to get pragmatic weakening in (55), and so the presence of *both* in (56) is not rendered redundant.

In addition, other people have pointed out to me (including Jason Merchant and Geoff Nunberg, p.c.) that *ever* seems to bring out the weakening effect, as in (57)-(58), which seem to demand an affirmative answer even if only one parent has been to Europe and one boy (out of a contextually salient group of two) has hit a girl.

(57) Have your parents ever been to Europe?
(58) Have the boys ever hit the girls?

If indeed definite plurals in general allow pragmatic weakening, no matter how small the plurality they denote, then the behavior of (57)-(58) is not unexpected. So I conclude that pragmatic weakening with a group of just two is in principle possible, and the function of *both* is to rule it out.

Recall also that, as we noticed in Chapter 1, the distribution of *both* follows Taub's

generalization. It is hard to find collective achievements that are felicitous with *both* (we would have to imagine a context in which two people elected a president, and still I'm not quite sure we can use *elect* in that kind of context), However, it is certainly true that collective states are not felicitous with *both.*

(59) *John and Mary are both a happy couple
*Connors and Borg are both a good doubles team
*Jim and Rita are both outnumbered by Hank, Bill, and Kim (**on coll. reading*)

This is of course expected if *both* and *all* have the same meaning, as I've proposed here.

3.5.1 Both *and reciprocals*

The behavior of *both* with reciprocal predicates deserves some special attention. As we have already seen, *all* is quite freely licensed with reciprocal predicates, such as *share a cookie, meet, like each other,* etc. But *both*'s distribution with reciprocal predicates, in contrast, shows some interesting restrictions. It can occur with some reciprocal predicates, but not others, as observed in Brisson (1996). Consider the contrasts in the data below.

(60) Les and Pete both collided
(61) Jane and Sarah both left together
(62) Mike and Danny both looked at each other

(63) *The oil and the vinegar both separated
(64) *Jack and Emily both met

In addition to the data in (60)-(64), there is the curious fact that for most of these reciprocal predicates, the position of the floating quantifier makes a difference. The sentences that are good in (60)-(62) with floated *both* degrade significantly with prenominal *both.* (I include *VP+together*, whose status as a "reciprocal" is unclear,

because it shows this difference between prenominal/floated *both*.)

(65) *Both Les and Pete collided
(66) *Both Jane and Sarah left together
(67) ?Both Mike and Danny looked at each other

I mention this fact here for completeness, but to keep the issues clearer I will defer discussion of it to Chapter 5 (section 5.2.3.3). In this section we will be concerned only with the facts in (60)-(64).

Brisson (1996) proposed that the crucial property distinguishing among the predicates in (60)-(64) is *symmetry*. That analysis was couched in terms of a different analysis of *both*, one in which *both* was proposed to have a particular kind of presupposition (in addition to its cardinality presupposition). Nevertheless, the central idea can be translated into the proposal I have made here for *both* in terms of the notion of good fit.

3.4.1.1 Collide-*type predicates*

To do that, we have to think about what it would mean for pragmatic weakening to occur with a reciprocal predicate like *collide*. Let us consider the sentence (68) in more detail.

(68) Pete and Les collided.

Suppose that we interpret (68) as (69), and we assign to *Cov* an ill-fitting cover which excludes Les, like the one in (70).

(69) $\forall x[x \in Cov_j \ \& \ x \subseteq [\![Pete.and.Les']\!] \rightarrow collided(x)]$

(70) $Cov_j = \{\{p\}, \{l,a,b,c\}\}$

This might be expected to be problematic, because on its reciprocal reading, *collide*

cannot apply to the atomic individual Pete. However, it appears that reciprocal predicates with a transitive alternant allow a kind of accommodation process here, which licenses a kind of "covert" use of the transitive variant.[10] In Brisson (1996) I proposed that this process obligatorily supplies the other member of the conjoined subject as the object of the verb. If this is correct, then (69) (where, recall, Cov_j is ill-fitting) might be interpreted as something like (71), where boldface **y** represents the other member of the conjoined subject (ie, Les).

(71) $\forall x[x \in Cov \ \& \ x \subseteq [\![Pete.and.Les']\!] \rightarrow collided.into(\mathbf{y})(x)]$

Thus the sentence would be interpreted as roughly synonymous with "Pete collided into Les."

This implies that in general, given that there is the possibility of pragmatic weakening, the sentence (72) is vague between three possible meanings: Pete collided into Les, Les collided into Pete, or Pete and Les (more or less equally) collided.

If we apply a familiar test for pragmatic weakening, continuation with a *but*-clause, we find some evidence that this is indeed the case. If (68) were incompatible with an asymmetrical reading we would expect (72) to be ill-formed, but if we suppose that some kind of pragmatic weakening is licensed with reciprocal *collide,* then (72) is expected.

(72) Les and Pete collided, but it was really Pete's fault.

3.4.1.2 Separate-*type predicates*

On the other hand, let us apply this same kind of "covert transitivity" analysis to

[10] In Brisson (1996) I called this process "context reevaluation", but in light of the fact that I am not using the presupposition part of that analysis this term is no longer felicitous.

a predicate like *separate,* using the sentence in (73).

(73) The oil and the vinegar separated

By analogy with *collide*, we would expect that if pragmatic weakening can apply to this predicate via covert transitivity, then (73) should in principle be vague between three interpretations: The oil separated from the vinegar, the vinegar separated from the oil, and the oil and the vinegar (more or less equally) separated.

But the meaning of *separate* intervenes here. It is arguable that these three interpretations are not really truth-conditionally distinct. It is impossible for the oil to be separated from the vinegar unless the vinegar is also separated from the oil.

If this is true, then even if we allow the same process of covert transitivity to apply to *separate* just as it does to *collide*, we will still not get any effect of pragmatic weakening. The symmetricality inherent in the lexical meaning of the predicate ensures that there is no truth-conditional difference between the three meanings that are made available by covert transitivity.

If there is no pragmatic weakening with symmetrical predicates, then we would expect that *both* should also not be licensed, since the function of *both* is to rule out pragmatic weakening. Hence the difference between the predicates in (60)-(64) is explained.

The class of symmetrical reciprocal predicates, which are predicted to be infelicitious with *both*, includes those in (74) and (75).

(74) John and Mary both *met
*married
*divorced

(75) John and Mary are both *similar
*different

The class of asymmetrical reciprocal predicates, which are predicted to be felicitous with *both*, include those in (76) (and more generally, almost any predicate containing the reciprocating operator *each other*, for reasons that should be clear), and those in (77).

(76) Brian and Carla both love each other
looked at each other

(77) Brian and Carla both flirted
embraced
agreed

In fairness, it must be pointed out that are some predicates that appear to be quite symmetrical, that nevertheless allow *both*. The most striking example is *look like each other*; clearly, if Carla looks like Brian then Brian looks like Carla, but nevertheless *both* is possible with this predicate. The only explanation I can offer for this anomaly is that "perspective-taking" seems to be such a prominent aspect of the meaning of the predicate *look like x* that it really does have an asymmetric aspect to its meaning.[11]

In addition, despite the difficulty of a predicate like *look like each other*, there is more evidence that symmetricality is the right notion for explaining the distribution of *both*. First there is the fact that adding *both* takes away the asymmetrical reading, which is what we would expect if the asymmetrical reading is really a kind of pragmatic weakening. Note that (78), with *both*, suddenly disallows the *but*-phrase that was a

[11] By "perspective-taking" I mean the fact that when we compare a grandmother and a granddaughter, for example, it is much more felicitous to say that the granddaughter looks like the grandmother than vice versa. Similarly, if we are comparing something familiar and something new, it is much more felicitous to say that the new thing looks like the old thing than the reverse.

permissible continuation to (72).

(78) *Pete and Les both collided, but it was really Pete's fault

Secondly, it is possible to "asymmetricalize" a symmetrical predicate. This is shown by the contrast between (79) and (80). (I thank Veneeta Dayal for this example.)

(79) *Dixie and Jerald both married
(80) Dixie and Jerald both willingly married

The difference between *marry* and *willingly marry* should be clear: if Dixie married Jerald then it entails that Jerald married Dixie. But of course if Dixie willingly married Jerald then Jerald could have nevertheless got to the altar very reluctantly. Since we have taken the very same predicate, *marry*, and simply added an adverb to it, this suggests that it is symmetricality that is crucial to the failure of (unmodified) *marry* to occur with *both*.

So pragmatic weakening, with some help from the process of covert transitivity, correctly predicts the more limited distribution of *both* with reciprocal collective predicates.

3.5 A problem with collective predicates

In my discussion I have been claiming that it is an advantage of my proposal that it predicts we should expect *all* to occur with collective predicates. This is true, and it means that I do not face the problem that Dowty and Link faced, which was how *all* can combine with collective predicates.

But instead, I face an inverse problem, because although it is true that *all* occurs with some collective predicates, there are other collectives that do not allow *all*. Instead of trying to explain why *all* can occur with any collective predicates, I have to explain why

there should be any collective predicates that disallow *all*.

Recall the predicates of Taub's generalization, which we saw in Chapter 1. These were collective state and achievement predicates, which don't (usually) allow modification with *all*, as in (81)-(83)

(81) *The boys are all a big group
(82) *The senators all passed the pay raise
(83) *The students all elected a president

There are also collective predicates that disallow *both*. (84) is the example we saw from Ladusaw (1981).

(84) *John and Mary are both a happy couple

It is worth pointing out that *be a happy couple* is a collective stative predicate, and so it too falls under Taub's generalization. This is expected under the hypothesis put forward here, that *both* and *all* have a meaning component in common.

The problem for the analysis I have given here is that it predicts that this subclass of collective predicates should allow *all* because nothing would prevent us from interpreting (81) as (85).

(85) $(\text{Part}(\text{Cov}_j)(\text{big.group}))(\text{the.boys}^{gf})$
$\forall x[x \in [\![\text{Cov}_j]\!] \,\&\, x \subseteq [\![\text{the.boys}^{gf}]\!] \rightarrow x \in [\![\text{big.group}']\!]]$

We can choose a cover for Cov_j in (85) that would yield a perfectly sensible interpretation of (81), namely, a cover in which the set denoted by *the boys* occupies one cell of the cover.

The issue of the ungrammaticality of (81) and sentences like it will be the topic of extensive discussion in Chapter 4. But for now, I wish to point out two phenomena that have the same distribution as *all* with respect to Taub's generalization, as evidence that the

generalization is tapping into something real.

First, the predicates of Taub's generalization do not allow pragmatic weakening. We can see this by applying the same tests for entailment/implicature that we have been applying since Chapter 2 (section 2.2.3.7) For example, we have seen that for a sentence like *the boys are hungry*, the 'maximality' of *the boys* can be denied without contradiction. This is not so for the sentences in (86)-(87), which are quite odd.

(86) *The boys are a big group, but not John and Bill /but not all of them

(87) *The students elected a president, but Kerry and Jan didn't do anything/but not all of them.

We also saw that the 'maximality' of the plural NP could be reinforced without redundancy, but again, with the subclass of predicates that disallow *all*, the sentences come out distinctly odd.

(88) #The boys -- in fact, the entire lot of them -- are a big group
(89) #The students -- in fact, the entire lot of them -- elected a president

So it appears that these sentences do not allow pragmatic weakening. This suggests that there is indeed a link between pragmatic weakening and *all*, which provides support for the hypothesis I've proposed here.

Another phenomenon that picks out this class of predicates is the distribution of *except*-clauses. *Except* clauses are acceptable with distributive predicates, as (90) shows, and acceptable with some collectives, as (91) shows. But as (92)-(93) show, they are distinctly odd with collective states and achievements.

(90) The girls went to the gym except for Donna
(91) The girls gathered in the hallway except for Joanne

(92) #The boys are a big group except for Jason/except for Jason and Andrew

(93) #The students elected a president except for Carol/ except for Carol and Bill

The continuation after the / in (92)-(93) is there to show that it's not the singularity of the DP in the *except*-clause that makes the sentence bad.

According to von Fintel (1994 and see references cited there), *except*-clauses are 'domain subtractors'; that is, they subtract things from the domain of quantification of a quantifier that is present elsewhere in the sentence. Note that this is similar to the idea that I have proposed here for *all*: *all* does something that affects the domain of quantification of a D operator.

If collective predication of *gather*, for example, can involve a D operator, then we expect that an *except*-clause should be possible as in (91), which is a welcome result. (And note that this can be taken as additional evidence for the claim made here that even collective predication involves a quantifier.) But again, if this is so, it is not clear why it shouldn't be possible to use a D operator with the collective states and achievements, and hence to license an *except* clause with those predicates.

So it appears that *all*, pragmatic weakening, and a new phenomenon, *except*-clauses, are sensitive to the presence of something in the structure of a sentence with a definite plural, and the most plausible candidate for that something is a D operator. However, we have not explained why these phenomena should not be possible with the predicates of Taub's generalization, because we don't have any reason to think that a D operator isn't licensed with those predicates. In the next chapter I will propose a theory of the 'scope' of distributivity which will explain why we find this restriction on the predicates of Taub's generalization, and will explain other phenomena involving *all* as well.

Chapter 4

The Scope of Distributivity

4 Introduction

Up until now I have been concerned almost exclusively with plural definites and *both/all* in subject position (or floated position, which in English is always construed with the subject). In this chapter we will see that the properties that led us to postulate the theory of pragmatic weakening and the meaning of *all/both* for plural DP's in subject position are also exhibited by plural DP's in object position. We will adopt a recent proposal on distributivity by Lasersohn (1998a) to provide an account of distributivity over object DPs. Once we can distribute over objects, pragmatic weakening and the meaning of *all* in object position will follow straightforwardly from the account of *all* and pragmatic weakening that I have already given.

We will then look at three other phenomena related to *all*: scopal differences between *all* and *every*; differences between *all* and *every* with respect to certain 'collectivizing' adverbials; and Taub's generalization. We will find that the latter two phenomena are explained by the same hypothesis that explains the first; that is, all three phenomena turn out to hinge crucially on the 'scopal' properties of *all* and distributivity, which will be discussed in detail in this chapter.

4.1 Pragmatic weakening with objects

Plural definite NPs in object position show the same sort of pragmatic weakening that we saw in subject position. So for example, take the sentence in (1).

(1) Jane tasted the cakes

In this sentence, the implication that Jane tasted all of the (contextually relevant) cakes has the status of an implicature, as shown by the fact that it can be canceled, and reinforced without redundancy.

(2) Jane tasted the cakes, but she overlooked the lemon creme cake.
(3) Jane tasted the cakes — in fact, she tasted all of them.

As we might expect, modification by *all* induces the same sort of maximizing effect on objects that it does on subjects. *All* eliminates the possibility of pragmatic weakening.

(4) Jane tasted all the cakes.

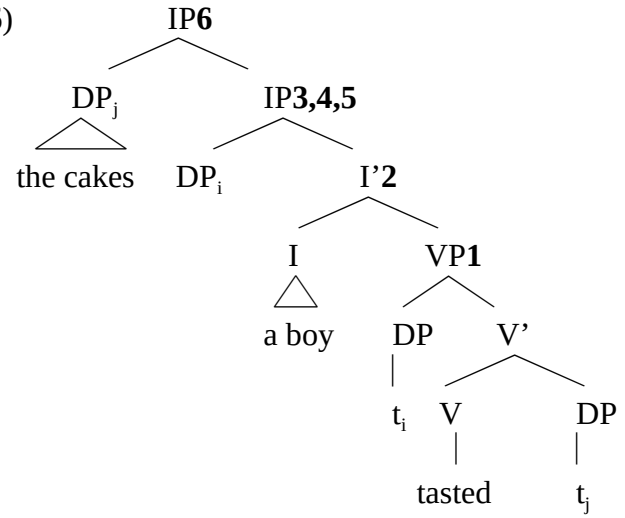
So we would like to extend our theory of pragmatic weakening and the meaning of *all/both*, which up to now has been concerned exclusively with subjects, to objects.

4.1.1 A problem with the QR approach

My analysis of pragmatic weakening requires a distributivity operator. So to apply it to object position, I will have to apply distributivity just to an object. One way of applying distributivity to an object is to QR the object (assuming we may optionally QR definite plurals) and then assume that the remaining structure forms a derived predicate that can be distributed over the object DP (see, for ex., the discussion in Roberts 1987, and Lasersohn 1998a). So, for example, the sentence in (5) could be interpreted as in (6)

(5) A boy tasted the cakes

(6)



1. tasted'(x_j)(x_i)
2. λx_i[tasted'(x_j)(x_i)]
3. ∃x[boy'(x) & tasted'(x_j)(x)]
4. λx_j∃x[boy'(x) & tasted'(x_j)(x)]
5. ^{D}λx_j∃x[boy'(x) & tasted'(x_j)(x)](the.cakes')
6. ∀y[y⊆⟦the.cakes'⟧ & y∈⟦Cov$_j$⟧ →∃x[boy'(x) & tasted'(y)(x)]]

There is a problem with this, however: the expression in line 6 of (6) can be true if for each cake there is a different boy who tasted it, but the sentence in (5) doesn't allow this. The problem is that we have given the universal quantifier introduced by distributivity scope over the existential from *a boy*.

One way to avoid this problem is to adopt the groups approach to nonmaximality. Then we would say that in (5), the object *the cakes* isn't interpreted distributively, but as a group. Then we don't need to QR the plural DP (or we could QR it, but it wouldn't affect the interpretation). (5) would be interpreted simply as (7).

(7) ∃x[boy'(x) & tasted'(the.cakes')(x)]

But I argued in the last chapter that the groups approach is not the best way to handle pragmatic weakening. Here we will adopt a theory in which we do not need groups, and we

do not need QR.

4.1.2 "Generalized Distributivity"

Lasersohn (1998a) provides a general semantic rule for distributivity that can apply to any constituent that can take a plural DP as its argument. The rule is modeled after generalized conjunction (Partee and Rooth, 1983, Keenan and Faltz 1985); for this reason Lasersohn calls his D operator a "generalized D operator". So from this point forward we will use this term to refer to Lasersohn's D operator (where in Chapter 3 we had used it to refer to a D operator that can range over pluralities as well as atomic individuals; now I will simply assume that the D operator can always range over pluralities and individuals).

The rule crucially makes use of a generalized conjunction operator, so Lasersohn gives the following (essentially standard) definition of a generalized conjunction operator $\sqcap$.

(8) a. t is a conjoinable type
b. If <**a,b**> is a type and **b** is a conjoinable type, then <**a,b**> is a conjoinable type.

(9) a. If $X \subseteq \mathbf{D}_t$, then $\sqcap X = 1$ if $X = \{1\}$; $\sqcap X = 0$ otherwise.
b. If $X \subseteq \mathbf{D}_{<a,b>}$ (where <a,b> is a conjoinable type), then $\sqcap X$ is that function $f \in \mathbf{D}_{<a,b>}$ such that for all *a*, $f(a) = \sqcap \{f'(a) \mid f' \in X\}$

Lasersohn defines a set of "distributable types"; these are the types with which the D operator can combine. Distributable types are defined with respect to conjoinable types. Then the definition of the D operator is given below (I have altered Lasersohn's notation slightly so that it conforms with the conventions I am using here).

(10) *Distributable types*:
If **a** is a conjoinable type, then <e,**a**> is a distributable type.

(11) Where α is an expression of some distributable type <e,**a**> and *x* is any individual (ie, $x \in \mathbf{D}_e$): $[\![^{D}\alpha]\!]^{M,g}(x) = \sqcap\{[\![\alpha]\!]^{M,g}(y) \mid y \leq x\}$

The generalized D operator that Lasersohn gives does not make use of the variable *Cov*; but this can be incorporated quite easily (as Lasersohn himself points out). We simply change the interpretation rule for the D operator to the one below.

(12) Where α is an expression of some distributable type <e,**a**> and *x* is any individual (ie, $x \in \mathbf{D}_e$): $[\![^{D}\alpha]\!]^{M,g}(x) = \sqcap\{[\![\alpha]\!]^{M,g}(y) \mid y \subseteq x \ \& \ y \in [\![Cov_i]\!]^{M,g}\}$

Let us suppose that the D operator is a syntactic object, whose presence can be (but doesn't have to be) triggered by the presence of a plural DP. I will assume here that syntax doesn't play any role in constraining the insertion of the D operator. The only constraint on the distribution of the D operator is its semantic requirement that it apply to a thing that takes a plural DP as its (first) argument.[1]

The definition in (12) takes a "mode of composition" approach to distributivity. That is, it applies to a predicate in combination with its (plural) argument, rather than applying either to the predicate or to the argument directly. (D operators that are similar in this respect can be found in Roberts 1987 and Schwarzschild 1992b.) For our purposes it will be more convenient if we adapt Lasersohn's definition of a D operator into something that can apply to objects of type <e,**a**>. So I give a revised definition of a generalized D operator below.

(13) Where *Z* is a variable of type <e,**a**> (ie of a distributable type),
$^{D}=_{df} \lambda Z \lambda x \sqcap\{Z(z) \mid z \subseteq x \ \& \ z \in [\![Cov_i]\!]\}$

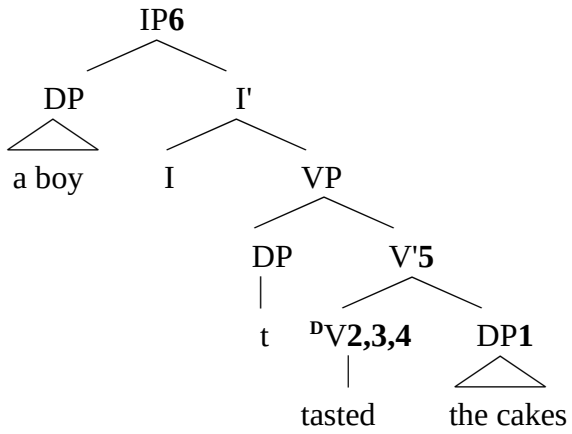
With this D operator, we can account for pragmatic weakening on objects. I give a derivation below.

---

[1] I will later argue that general principles of economy can also play a role in restricting the distribution of the D operator but this sort of restriction is not unique to the syntax or semantics of the D operator.

(14) A boy tasted the cakes

(15)



1. the.cakes'
2. λxλy[tasted'(x)(y)]
3. ^{D} = λRλx⊓{R(z) | z⊆x & z∈⟦Cov$_i$⟧}
4. λx⊓{λxλy[tasted'(x)(y)](z) | z⊆x & z∈⟦Cov$_i$⟧}
5. ⊓{λxλy[tasted'(x)(y)](z) | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}
   =⊓{ λy[tasted'(z)(y)] | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}

5'. (=5) λy'[y'∈ ⊓{ λy[tasted'(z)(y)] | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}]
6. ⟦a.boy'⟧ ∈⊓{ λy[tasted'(z)(y)] | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}

Let us consider this derivation in some detail. The generalized conjunction operator always takes a set of functions (except in the case of sentence conjunction, where it takes a set of truth values) as its argument. We will need to refer to that set here and throughout this chapter, so let us name it: it is a set of functions *F*. What the D operator does is define *F* as follows: *F* is the set of functions that you get by plugging in the parts of a given plural object (*x*) as the argument of a given function of type <e,**a**> (in this case, *R*, type <e,<e,t>>). This restriction on *F* is where we get universal quantificational force over the parts of that object.

In the example in (15) the elements of *F* are functions of type <e,t>. Then the meaning of the operator ⊓ comes in: ⊓ turns the set of functions *F* into that function *f* , also of type <e,t>, such that for all *x*, $f(x) = \sqcap\{f'(x) \mid f' \in F\}$.

Therefore, an expression like $\sqcap\{\lambda y[\text{tasted}'(z)(y)] \mid z \subseteq \text{the.cakes} \;\&\; z \in [\![\text{Cov}_i]\!]\}$, from line 5, is itself a function of type <e,t>. This is why 5 and 5' are equivalent; the equivalence between 5 and 5' is the same as the equivalence between *girl'* and $\lambda x[x \in [\![\text{girl}']\!]]$. We will sometimes use this alternate notation where it makes our derivations easier to read.

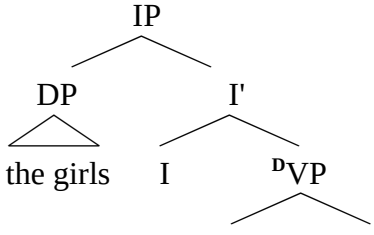
So the function yielded by the generalized conjunction operator is a function that takes individuals as its arguments and yields truth just in case that individual tasted (each of) the cakes (modulo pragmatic weakening). If there is *a boy* who has that property, then the sentence in (13) will come out true. Hence we get distributive quantification with "low scope".

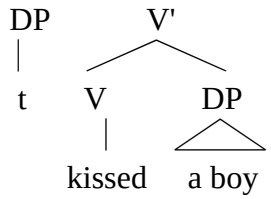
4.1.3 A subject-object asymmetry

Since we are treating the D operator as a syntactic object, this leads us to expect that there should be an asymmetry between distributivity over object DPs and distributivity over subject DPs. Since object distributivity is due to a D operator on V, and subject distributivity is due to a D operator on VP, and VP of course contains V, we expect an asymmetry of the type illustrated below.

(16) The girls kissed a boy

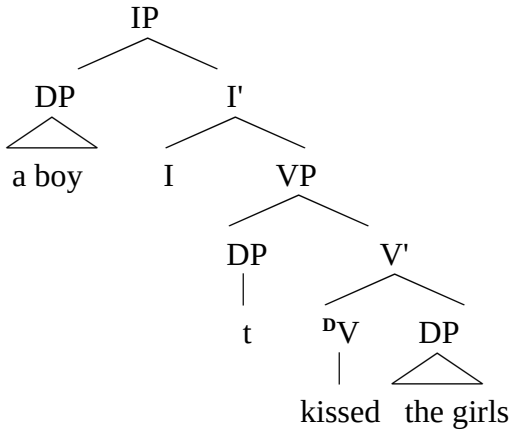
(17)





(18) A boy kissed the girls

(19)



(16) can be true if each of the girls kissed a different boy, but the sentence in (18) requires that there be a single boy who kissed every girl. This is predicted because in (16) the D operator is applied to *kissed a boy*, so it has scope over the indefinite in object position. (18) the D operator is applied only to the predicate *kissed*, so it does not have scope over the indefinite in subject position.

Of course, this argument would not go through if it were possible for definite descriptions to undergo QR, as has been argued by, for example, May (1985) and Pritchett (1990). I will take this argument as evidence, then, that definite descriptions do <u>not</u> undergo QR (see also Williams 1992, Dayal 1992, Krifka 1992), and I will assume for the rest of this

dissertation that this is the case.[2,3]

4.1.4 A scope puzzle solved

This treatment of object distributivity, combined with the proposal I made in Chapter 3 for *all*, immediately provides a solution to the following scope puzzle.

(20) A doctor examined every athlete
(21) A doctor examined all the athletes

The puzzle that these sentences present is the following. (20) (apparently) allows for either *a doctor* to take wide scope, or for *every athlete* to take wide scope. In the former case, it means every athlete was examined by the same doctor, and in the latter, there could have been several doctors who did the examining, as long as each athlete got examined. (21), in contrast, only allows for the first possibility. So it appears that *every athlete* can

---

[2] This is of course only a cursory examination of many issues that are related to QR. For example, Veneeta Dayal (p.c.) points out that QR to VP might make it possible to derive the desired "low scope" for objects; but note that if we allowed for this we would still need to find a way to rule out QR of a definite plural object to IP. Roger Schwarzschild (p.c.) points out that if D operators were not licensed on derived predicates we could QR the object but this would not lead to a distributive, wide scope reading (only a "group" reading would be allowed, making the QR semantically vacuous in the same way that it is vacuous when applied to proper names). The point I wish to make in the text is that I have made one argument for prohibiting QR to definites, even if there are other issues involved.

[3] The derivation of certain kinds of "cumulative" readings might prove problematic under this idea. (I thank Roger Schwarzschild for pointing this out to me, and for the example in (i).) For example, the sentence in (i) is false on a cumulative reading (in which it is asserted that the total number of New Jerseyans who voted for one or the other of the mainstream candidates is few).
(i) Few New Jerseyans voted for the mainstream candidates
However, if we cannot QR *the mainstream candidates*, then it appears we incorrectly predict that this sentence is (vacuously) true, because it is impossible to vote for the group of Dole and Clinton, and because it is impossible to vote for Dole and vote for Clinton (ie, vote for them distributively). Since, as I said in Chapter 2, cumulative readings are not part of the focus of this dissertation, I will set this problem aside as one that should be treated as part of an analysis of cumulativity.

take wide scope, but *all the athletes* cannot.[4]

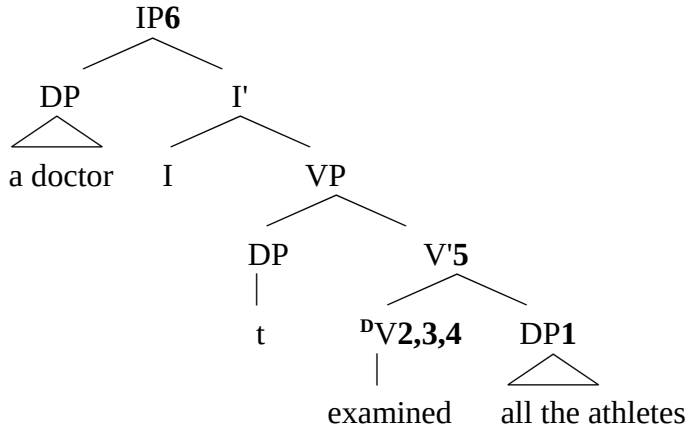
Note that this is exactly the same as the subject-object asymmetry between indefinites and definite plurals that we saw in the previous section.

The proposal I have made for *all* gives us an account in which the parallelism is accounted for: it is due to the syntactic scope of the D operator. Since *all the athletes* is, by hypothesis, a definite description and not a quantified expression, then we expect that it should not undergo QR. Thus its "scope," like the "scope" of any other definite description, will be the scope of the D operator on which it is dependent.

The interpretation of (21) proceeds exactly like the interpretation of (15), above, except that now the presence of *all* ensures there will be no pragmatic weakening.

(22) A doctor examined all the athletes

(23)



1. the.athletes[gf]

---

[4]Some speakers allow a wide scope reading for *all*, but find it is strongly dispreferred. Even the widely-cited 'inverse scope' examples from the literature (e.g. Hirschbüler 1982) sound much worse with *all* than with *every*. (I thank Roger Schwarzschild for bringing this data to my attention.)
(i) A flag hung in every doorway/??all the doorways
(ii) A policeman stood on every corner/*all the corners
Given that inverse scope for *all* is at best strongly dispreferred and at worst impossible, I'll put this aside as being due to interference from the perceived synonymy of *every* and *all*.

2. $\lambda x \lambda y$[examined'(x)(y)]
3. $\lambda R \lambda x \sqcap \{R(z) \mid z \subseteq x \ \& \ z \in [\![Cov_i]\!]\}$
4. $\lambda x \sqcap \{\lambda x \lambda y$[examined'(x)(y)] (z) $\mid z \subseteq x \ \& \ z \in [\![Cov_i]\!]\}$
5. $\sqcap \{\lambda x \lambda y$[examined'(x)(y)] (z) $\mid z \subseteq [\![$the.athletes$^{gf}]\!] \ \& \ z \in [\![Cov_i]\!]\}$
   $= \sqcap \{ \lambda y$[examined'(z)(y)] $\mid z \subseteq [\![$the.athletes$^{gf}]\!] \ \& \ z \in [\![Cov_i]\!]\}$

5'. (=5) $\lambda y'$[$y' \in \sqcap \{ \lambda y$[examined'(z)(y)] $\mid z \subseteq [\![$the.athletes$^{gf}]\!] \ \& \ z \in [\![Cov_i]\!]\}$]
6. [a.doctor'] $\in \sqcap \{ \lambda y$[examined'(z)(y)] $\mid z \subseteq [\![$the.athletes$^{gf}]\!] \ \& \ z \in [\![Cov_i]\!]\}$

This parallelism between *all the CN* and *the CN* in object position is unexplained on the view that *all* is a universal quantifier.

4.1.5 *Preview of things to come*

One difference between *all* and *every*, then, is the difference between the scope of distributivity and the scope of *every*: *every* undergoes QR and hence can take wide scope, while distributivity is interpreted in situ and hence has its scope fixed wherever it is inserted. In the rest of this chapter we will be concerned with other phenomena related to *all*, and the scope of the distributivity operator that *all* is dependent on will play an important role in our discussion throughout.

Here is a preview of some of the phenomena that we will be concerned with for the rest of this chapter. One issue is the contrast between *all* and *every* in their ability to license "collectivizing adverbials", as illustrated in (24)-(25).

(24) All the planes landed together/as a group/in formation/at once
(25) *Every plane landed together/as a group/in formation/at once

I will argue that the possibility of combining *all* with collectivizing adverbials is due to possibility of low scope of distributivity.

The other major issue we will address will be an explanation for Taub's generalization. In Chapter 3 we saw that our proposal for *all* predicts that *all* should be able to combine with collective predicates, but fails to predict that *all* cannot combine with a

subclass of the collective predicates. In this chapter we will revise our analysis of collective predicates and offer an explanation for Taub's generalization. Here again, the low scope of the D operator will play a crucial role. In addition, we will see that our explanation for the distribution of *all* also explains the distribution of pragmatic weakening, and of exception phrases.

In our analysis of these phenomena, we will adopt a theory of collectivizing adverbials inspired in part by Lasersohn (1995), and a theory of "event identification" from Kratzer (1994), that both crucially make use of event semantics (although they do not share all of their assumptions about event semantics). So before we can get to that analysis, we have to adopt a particular framework of event semantics.

4.2 Events

In a neo-Davidsonian event semantics, a verb is a predicate of events and it is linked to its arguments through thematic roles. So a sentence like (26) is interpreted as (28), as compared to the representation in (27) that I have been using so far.

(26) Alvin kissed Roxy.
(27) kiss'(Roxy)(Alvin)
(28) λe[kiss'(e) & Ag(e,Alvin) & Th(e, Roxy)]
(29) ∃e[kiss'(e) & Ag(e,Alvin) & Th(e,Roxy)]

The verb *kiss* is treated as a predicate of events, and individuals participating in the event as agent or theme are linked to the event by means of theta roles. A rule called 'existential closure' takes the predicate of events in (28) and changes it to a proposition as in (29) by binding the variable over events *e* with an existential quantifier. Once we have introduced an event argument on the verb, we will have to adapt our D operator to take events into account. Distributivity in event semantics involves not just universal

quantification over a DP denotation, but also existential quantification over the parts of an event. It is perhaps easier to see how this will work if we consider just VP distributivity first. Let's take our old example, repeated below as (30).

(30) The boys ate a sandwich

On its distributive reading in our old semantics without events, (30) is interpreted as in (31).

(31) $\forall x[x \subseteq [\![\text{the.boys}']\!] \ \& \ x \in [\![\text{Cov}_i]\!] \rightarrow x \in [\![\text{ate.a.sandwich}']\!]]$

Our old distributivity operator (the one that only applied to VP, from Chapter 3) was defined as $\lambda P \lambda y \forall x[x \subseteq y \ \& \ x \in [\![\text{Cov}_i]\!] \rightarrow x \in P]$. Its counterpart in event semantics could be defined as in (32). Note that I use the subset relation as the part-of relation on events.

(32) $\lambda P \lambda x \lambda e \forall y \exists e'[y \subseteq x \ \& \ y \in [\![\text{Cov}_i]\!] \rightarrow P(e')(y) \ \& \ e' \subseteq e]$

This will apply to a VP like (33) to yield (34). (I treat P here as a variable that ranges over functions from individuals to sets of events (type <i,<v,t>>, see below), because this is what a VP denotes in event semantics.)

(33) $^{D}\lambda x \lambda e[\text{ate}'(e) \ \& \ \text{Ag}(e,x) \ \& \ \text{Th}(e,\text{a.sandwich})] =$

(34) $\lambda x \lambda e \forall y \exists e'[y \subseteq x \ \& \ y \in [\![\text{Cov}_i]\!] \rightarrow \text{ate}'(e') \ \& \ \text{Ag}(e',y) \ \& \ \text{Th}(e',\text{a.sandwich}) \ \& \ e' \subseteq e]$

When we have fed in *the boys* as the subject argument and existentially bound the variable *e*, we get (35).

(35) $\exists e \forall y \exists e'[y \subseteq [\![\text{the.boys}']\!] \ \& \ y \in [\![\text{Cov}_i]\!] \rightarrow \text{ate}'(e') \ \& \ \text{Ag}(e',y) \ \& \ \text{Th}(e',\text{a.sandwich}) \ \& \ e' \subseteq e]$

Thus we take the distributive reading of *the boys ate a sandwich* to assert the existence of an event which contains several subevents: these subevents are individual ate-a-sandwich events for each boy. We will carry this idea over to our definition of a generalized D operator in event semantics. But first we need to discuss the assumptions

about event structure that allow us to talk about the parts of an event.

4.2.1 Plural events and parts of events

I will follow Landman (1996, 97) and others (eg Krifka 1992) in assuming that we can define plural events. Landman and Krifka use a simple summing operation. So, for example, if there are three different singing events in our model, e, e', and e", then the extension of *sing* is as given in (36). Then we can apply the star operator to the predicate of events in just the same way that we apply it to nominal predicates, and the result is a plural event predicate **sing*, whose denotation is given in (37).

(36) ⟦sing'⟧ = {e, e', e"}

(37) ⟦*sing'⟧ = {e, e', e", e+e', e+e", e'+e", e+e'+e"}

Landman assumes that there is a difference between *sing'* and **sing'*. I want to suppose something a little bit different about the structure of our model, and the denotation of *sing*'. Let us suppose that there is a set of atomic singing events in a model, but this is not the denotation of *sing'*. Let us call this set *sing♯*. The extension of *sing'* is the power set of the set *sing♯*. In other words, the extension of an ordinary verbal predicate is similar to the extension of a plural common noun phrase: it is a set of sets. One consequence of this is that now the set of atomic events, *sing♯*, doesn't correspond to any lexical predicate. In a model where *sing♯* is the set in (38), the denotation of *sing*' is the set shown in (39).

(38) sing# = {e, e', e"}

(39) ⟦sing'⟧ = {{e},{e'}, {e"}, {e,e'}, {e,e"}, {e',e"}, {e,e',e"}}

One might wonder what counts as an "atomic" singing event. For example, suppose Mary sings a six-minute aria. How many "atomic" singing events is this? It seems to me that

there are basically two plausible answers to this question in the literature. One answer is: it is one atomic singing event. This is because a singing of an aria is simply perceived to be some sort of coherent whole. (See also Moltmann 1997, and the discussion of an event "for" a proposition in section 4.2.1.1 below.) Another answer is: many many singing events. This is the kind of answer we find in the work of McClure (1994), who proposes that activities are actually many many transitional events put together.

I think it's probably the case that both answers are correct, but different phenomena in language are sensitive to different levels of detail in event structure. Whether one chooses the first answer or the second answer will depend largely on whether or not the phenomena one is looking at is sensitive to the difference. In our case we will choose the first answer, because it seems to be the right level for the things we are looking at. The definition of an event "for" a proposition, discussed below, will help give us a formal way to do this.

Once we have a notion of atomic events and plural events, then we can talk about the part-of relation on events. In this chapter we will crucially want to distinguish two kinds of part-of relation.

The first part-of relation is the $\subseteq$ relation, and it is useful in talking about the parts of plural events. Suppose we assert (40).

(40) John and Mary ate dinner

Let us assume that there is a D operator on the VP. Then (40) is interpreted as (41).

(41) $\exists e \forall y \exists e'[y \subseteq [\![J\&M]\!] \ \& \ y \in [\![Cov_i]\!] \rightarrow ate'(e') \ \& \ Ag(e',y) \ \& \ Th(e,dinner') \ \& \ e' \subseteq e]$

This sentence asserts the existence of a (plural) event, which has subevents of the following sort: one subevent is an event of Mary eating dinner, and one is an event of John

eating dinner. (Note that the variable *e* ranges over atomic and plural events.)

There is another kind of part-of relation on events that will prove relevant in later sections. There is a clear intuition that tells us that the event of John picking up his fork, for example, or Mary cutting into her pork chop, are also "subevents" of an eating dinner event. However, although they are subevents of eating-dinner events, they are not themselves eating-dinner-events. To represent this part-of relation on events, I will use the symbol ◂. So, for example, (42) is an expression that makes use of the ◂ relation.

(42) ∃e∃e'[eat(e) & Ag(e,M) & Th(e,dinner') & cut'(e') & Ag(e',M) & Th(e'.p.c.) & e'◂e]

The expression in (42) says that there exists an event of Mary eating dinner, and an event of Mary cutting her pork chop, such that the latter is ◂ the former.

One way to think of the difference between the ⊆ and the ◂ relation on events is that they correspond to Link's "i-part" and "material part-whole" relations, respectively (I thank Angelika Kratzer for this suggestion). For example, in the material domain, the handle of my coffee cup is in the material part-whole relation to my coffee cup. What this means is that "the portion of matter constituting the handle of my coffee cup is m-part of the portion of matter constituting my coffee cup," according to the paraphrase suggested by Link.

In order to carry this over to events we have to assume there is some kind of eventive equivalent to "matter"; we'll assume that this notion is plausible and call it "event stuff". Then the ◂ relation does indeed look like the material part whole relation. This leads us to expect that (42) can be paraphrased as, "the portion of event stuff constituting the event of Mary cutting her pork chop is part of the portion of event stuff constituting the event of Mary eating dinner", which, I submit, is true.

The i-part, or ⊆ relation, is coarser than the material part whole, or ◂ relation, as Link points out, because the following conditional holds, but it is not a biconditional.

(43) e'⊆e → e'◂e

This is a welcome fact, because in our upcoming discussion of *together* the difference between subevents that are ◂ , but not ⊆, a larger event will be crucial.

4.2.1.1 An event for a propositon

There is another potential problem in event semantics that we should deal with here. Because we existentially quantify over events, there are many events that could potentially make the sentence in (44), translated as (45), true.

(44) John ate beans
(45) ∃e[ate'(e) & Ag(e,J) & Th(e,beans')]

For example, obviously an event in which John ate beans makes (45) true. But so does a plural event, one that contains as its subevents the event of John eating beans and the event of the Yankees winning the World Series in 1996. Another event that makes this sentence true is the plural event containing the event of John eating beans, the event of the Yankees winning the world series in 1996, and the event of Nixon's visit to China. And so on. This extraneous stuff would prove troublesome to our analysis in the coming sections if we let it in.

One approach to this problem is taken by Lasersohn (1995). For every case where it makes a difference (particularly in the analysis of *together*) Lasersohn introduces the part-of relation as part of the semantics of the relevant operator, so that the extraneous event-stuff doesn't interfere in the evaluation of the truth of the sentence. But in addition to being cumbersome, it fails to capture the fact that it seems to be the case generally that when we

talk about events we want to talk about something like the "minimal" event that makes a sentence true.

But using minimal events is not a very good solution either. Kratzer (p.c.) points out that if we always evaluate propositions with respect to minimal events, this seems to give us the wrong results. For example, take a sentence like *Mary sang*. Consider the case we saw earlier, where there is an event of Mary singing an aria. If we use minimal events to evaluate the sentence *Mary sang*, then it would appear to be the case that a minimal event that makes this sentence true is any of the instants during which Mary was singing. But our perception is that the event that makes the sentence true is the big event that includes the whole aria.

Angelika Kratzer suggests that her idea of "an event for a proposition" as the default mechanism for selecting an event that makes a sentence true will do the job we need here. The basic idea is that we don't have to evaluate a sentence with respect to a minimal event in order to keep out extraneous events. Instead, the way the notion of an event for a proposition is defined, we are allowed to choose sets of (or sums of, in Kratzer's formulation) minimal events. So for example, even if we count every instant of singing an Aria as a singing event, this would allow us to choose the event of Mary singing the whole aria as the event for the proposition *Mary sang*, as long as it doesn't include any parts that are not events of Mary singing.

The definition of an event for a proposition is given in (46). (Due to Angelika Kratzer, p.c.; but her formulation uses $\leq$. Because of the way it is defined, it actually doesn't matter whether we use $\blacktriangleleft$ or $\subseteq$ as the relevant part-of relation in my version.) (Recall

that propositions are taken to denote sets of events before existential closure.)

(46) An event *e* is an event for a proposition *p* iff:
∀e'[e'◂e & e'∈p → ∃e"[e'◂e"◂e & min(e")[e"∈p]]]

The expression *min(e")[e"∈p]* is meant to be read as, "e" is a minimal event such that it is in p".

When we evaluate the truth of a sentence, we always choose an event for that proposition, and hence that there is no extraneous stuff in the event.

4.2.2 A model with events, and some syntactic considerations

By treating verbs as predicates of events, we have a different kind of model, one in which the universe of entities is sorted so that it includes events and individuals as separate types. Following the notation used in Lasersohn (1998), we will say that individuals are of type i and events are of type v. (Type t, the type of a truth value, remains the same.) Now a predicate like *sing* is not type <e,t>, but <i,<v,t>>. It can be represented equivalently as either (47) or (48).

(47) sing'
(48) λxλe[sing'(e) & Ag(e,x)]

To keep things simple, even though we have some new types in our system I will try to stick to the conventions of variable use as closely as possible. So, for example, as we saw above, I will assume that *P* ranges over (one-place) verbal predicates — only now a verbal predicate is a function from individuals to a function from events to truth values. I will follow suit with other variables (such as *R*, a variable over relations, which are now type <i,<i,<v,t>>).

We will also adopt a stipulation that Landman (1996) proposes on thematic roles,

which he calls the "unique role requirement" (URR). He argues that thematic roles are partial functions, not relations, from events to individuals. This means that every event can have at most one individual satisfying its agent role, its patient role, etc. Landman argues that adopting this stipulation leads to some welcome results, including a solution to the problem of linking the argument of a *by*-phrase to a passive sentence. I will adopt the unique role requirement here, because allowing multiple agents, say, of a single event would complicate our semantics in unwelcome ways. It should be noted, though, that it is possible to have an event semantics without it (see Schein 1993 for one such system, where events can have multiple agents and multiple themes, etc).

Finally, I will also adopt a variation on a principle proposed by Landman called the "Scope Domain Principle" (SDP). This principle says that quantifiers must take scope outside of the existential quantifier introduced by existential closure (this is the scope domain). Here we will take Landman's SDP as the instantiation in event semantics of the idea that quantified expressions must undergo QR.[5] For concreteness, we will suppose that existential closure is introduced at the IP level, and that QR is adjuction to IP or a higher verbal projection. (If we allow QR to VP, then existential closure would have to be permitted to apply at this level; I won't explore that possibility here.)

The SDP is motivated in part by the fact that negative quantifiers must take scope over existential closure if we are to generate the right interpretation for a sentence like (49).

(49) No girl left.

---

[5] Landman suggests that QR might be a viable syntactic instantiation of this idea, but he does not explicitly adopt the idea, preferring instead not to take an explicit stand on the syntax-semantics mapping.

If *no girl* is allowed to take scope inside the existential quantifier, then the sentence would be interpreted as in (50).

(50) $\exists e[\neg\exists x[girl(x) \& left'(e) \& Ag(e,x)]]$

The expression in (50) means something like, 'there is an event in which no girl left,' but this is clearly not what the sentence in (49) means, so (50) must be wrong. (50) is almost guaranteed to be true in any model: an event that contains only a dog barking and nothing else makes the sentence true. So we need a way to rule out this reading: Landman's SDP stipulates that the negation must take scope outside of existential closure, as in (51).

(51) $\neg\exists x[girl(x) \& \exists e[left'(e) \& Ag(e,x)]]$

In this way the SDP guarantees that we get the right interpretation of (49) in event semantics.

Another motivation for the SDP is that it is necessary in order to be consistent with the unique role requirement. Take the sentence in (52), with the putative interpretation in (53).

(52) Every girl walked
(53) $\exists e[walk(e) \& \forall x[girl'(x) \rightarrow Ag(e,x)]]$

If the interpretation in (53) were licensed, then the unique role requirement would tell us that this sentence could only be true in a model with one girl, because an event of walking can have only one agent. But this is obviously wrong. The SDP rules it out by requiring that the universal quantifier over girls take scope over the existential quantifier over events, as in (54).

(54) $\forall x[girl(x) \rightarrow \exists e[walk(e) \& Ag(e,x)]]$

Landman's version of the SDP allows that definite and indefinite NP's/DP's may

optionally take scope outside of the scope domain. However, we have already seen evidence against this idea. It is not clear to me why Landman assumes that we need optional QR for non-quantificational DPs,[6] but in any event here we will stick to the assumption that they do not QR. Hence indefinites and definites are not subject to the SDP.

4.2.3 A generalized D operator for event semantics

Now we have a general framework of event semantics, and we need to adapt our generalized D operator to event semantics. Again, Lasersohn (1998a) has already shown how such a D operator will work. Since the generalized D operator depends on generalized conjunction, we will have to define generalized conjunction for event semantics. We will adopt Lasersohn's treatment of generalized conjunction in event semantics. Note that crucially conjunction in event semantics involves a part/whole relation on events. I take the relevant part/whole relation to be $\subseteq$, following the discussion of the previous sections.

The definition of a conjoinable type, and of generalized conjunction, must be revised to take the type of events into account. These revisions are given in (55) and (56) (adapted from Lasersohn 1998a:87).

(55) a. <v,t> is a conjoinable type.
b. If <**a**,**b**> is a type and **b** is a conjoinable type, then <**a**, **b**> is a conjoinable

---

[6] It appears that for Landman, we need it to account for examples like (i), with two indefinites.
(i) Three boys invited four girls

The literature on these types of examples is very large, and I won't go into all of the issues that an example like (i) brings out (see, for example, Scha 1981, Roberts 1987, van der Does 1993). But the relevance of this example is that it appears that Landman uses QR to generate the reading of (i) in which, possibly, each boy invites four different girls, so there are up 12 girls involved — four girls per boy.

It seems to me that it would also be possible to account for this reading simply by inserting a D operator on the VP, which would take scope over the indefinite *four girls*. However, this is just speculation because I have not given a theory here of whether or when D operators are licensed with indefinites.

type.

(56) a. If $X \subseteq \mathbf{D}_{<v,t>}$, then $\sqcap X$ is that function f: $\mathbf{D}_v \rightarrow \mathbf{D}_t$ such that for all $e \in \mathbf{D}_v$:
$f(e) = 1$ iff $\forall f' \in X\ \exists e' \subseteq e\ f'(e') = 1$.
b. If $X \subseteq \mathbf{D}_{<a,b>}$ (where <**a**, **b**> is a type and **b** is a conjoinable type), then $\sqcap X$ is that function $f \in \mathbf{D}_{<a,b>}$ such that for all a, $f(a) = \sqcap\{ f'(a) \mid f' \in X \}$.

Our generalized D operator now does not have to change very much, except that the type of individuals is now type i instead of type e. Otherwise, (57)-(58) are exactly like (10) and (12) from section 4.1.2.

(57) *Distributable types*:
If **a** is a conjoinable type, then <i,**a**> is a distributable type.

(58) Where α is an expression of some distributable type <i,**a**> and x is any individual (ie, $x \in \mathbf{D}_i$): $[\![^{D}\alpha]\!]^{M,g}(x) = \sqcap\{[\![\alpha]\!]^{M,g}(y) \mid y \subseteq x \ \&\ y \in [\![Cov_i]\!]^{M,g}\}$

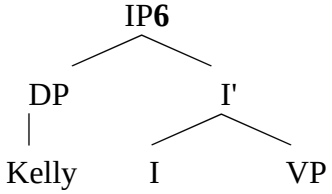
And again, as we did earlier, we adapt the definition in (58) so that our D operator can apply directly to a predicate.

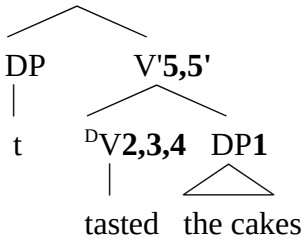
(59) Where Z is a variable of type <i,**a**> (ie a distributable type),
$^{D}=_{df} \lambda Z \lambda x \sqcap\{Z(z) \mid z \subseteq x \ \&\ z \in [\![Cov_i]\!]\}$

Let us apply this to an example similar to the one we saw earlier: I will change the indefinite *a boy* to the definite description *Kelly* to make our explanation below a bit simpler. If we apply this D operator to the verb in the sentence below, the effect of event semantics is not immediately obvious. We will need to do some more "unpacking" at the end of the derivation.

(60) Kelly tasted the cakes

(61)





1. the.cakes'
2. λxλyλe[tasted'(e) & Ag(e,y) Th(e,x)]
3. λRλx⊓{R(z) | z⊆x & z∈⟦Cov$_i$⟧}
4. λx⊓{λxλyλe[tasted'(e) & Ag(e,y) & Th(e,x)] (z) | z⊆x & z∈⟦Cov$_i$⟧}
5. ⊓{λxλyλe[tasted'(e) & Ag(e,y) & Th(e,x)] (z) | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}
   =⊓{λyλe[tasted'(e) & Ag(e,y) & Th(e,z)] | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}

5'. (=5) λy'[y'∈ ⊓{λyλe[tasted'(e) & Ag(e,y) & Th(e,z)] | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}
6. ⟦K⟧ ∈ ⊓{λyλe[tasted'(e) & Ag(e,y) & Th(e,z)] | z⊆⟦the.cakes'⟧ & z∈⟦Cov$_i$⟧}

The expression in line 6 has the events "buried" down inside the meaning of the function that takes the subject as its argument. We would like to translate this expression into something that looks more like the convention in event semantics, with an existentially-closed event variable having the highest scope.

To do this we need to see just what the function is that the generalized conjunction operator gives us. The ⊓ operator takes as its argument the set of functions *F* that we get by plugging each individual part of the denotation of *the cakes* as the theme argument of *tasted*. Let us suppose that the denotation of *the cakes* is the set containing a chocolate cake, a lemon cake, and a streusel cake. Then *F* could alternately be written as in (62).

(62) { λyλe[tasted'(e) & Ag(e,y) & Th(e,the.choc')], λyλe[tasted'(e) & Ag(e,y) & Th(e,the.lem')], λyλe[tasted'(e) & Ag(e,y) & Th(e,the.stru')] }

When the generalized conjunction operator takes this set of functions as its argument, it yields a function *f* of type <i,<v,t>>. This function can take *Kelly* as its argument, and will assign to *K* the value of a function *f'* of type <v,t>, which is also defined

by using generalized conjunction (because ⊓ is recursively defined).

The function *f'* is, according to our example, the generalized conjunction of the set of functions in (63). Here our *F* is a set of functions of type <v,t>.

(63) {λe[tasted'(e) & Ag(e,K) & Th(e,the.choc')], λe[tasted'(e) & Ag(e,K) & Th(e,the.lem')], λe[tasted'(e) & Ag(e,K) & Th(e,the.stru')] }

Now that our *F* is a set of functions of type <v,t>, the definition above tells us what ⊓ applied to (63) should look like. I repeat the definition as (64) for the reader's convenience.

(64) If $X \subseteq \mathbf{D}_{<v,t>}$, then $\sqcap X$ is that function f: $\mathbf{D}_v \rightarrow \mathbf{D}_t$ such that for all $e \in \mathbf{D}_v$:
$f(e) = 1$ iff $\forall f' \in X \; \exists e' \subseteq e \; f'(e') = 1$.

In other words, this function will yield the value true for some event *e* only if for every function *f* in the set of functions in (63), *e* has a subevent that would yield the value true for *f*.

Now, we can replace the universal quantification over functions *f* with a notation that expresses the fact that these functions are gotten by universally quantifying over the parts of *the cakes*; namely, we can quantify over the cakes directly. First we will use the set of events notation (because here we are dealing with functions from events to truth values) and then we will apply existential closure to come up with a proposition that is the meaning we want to assign to the derivation in (61).

(65) $\lambda e \forall x \exists e'[x \subseteq [\![\text{the.cakes}']\!] \; \& \; x \in [\![\text{Cov}_i]\!] \rightarrow \text{tasted}(e') \; \& \text{Ag}(e',M) \; \& \; \text{Th}(e',x) \; \& \; e' \subseteq e]$

(66) $\exists e \forall x \exists e'[x \subseteq [\![\text{the.cakes}']\!] \; \& \; x \in [\![\text{Cov}_i]\!] \rightarrow \text{tasted}(e') \; \& \text{Ag}(e',M) \; \& \; \text{Th}(e',x) \; \& \; e' \subseteq e]$

In other words, (66) means the same thing as the last line of (61) (modulo existential closure). In the examples we will look at in the rest of this chapter, we will follow this

pattern in looking at the derivation. We will first derive a sentence straightforwardly as in (61), and then "unpack" its meaning so that we can convert it into an equivalent statement that looks more like the event semantics we are used to.

The derivation with an indefinite subject would look quite similar. If we replaced *Kelly* with an indefinite like *a boy*, then we would assert that there exists a boy who is in the domain of a function *f* such that *f* assigns to that boy a property of events of the following sort: the property is true of an event if for each cake in *the cakes*, that event has a subevent of that boy eating that cake.

At this point, we are finally ready to look at some of the interesting phenomena involving *all* and *every* that were mentioned in section 4.1.5. First we will turn our attention to collectivizing adverbials.

4.3 Collectivizing Adverbials

The distribution of *all* and *every* differ with respect to a subclass of collectivizing adverbials, as shown in (67)-(68).

(67) All the planes landed together/in formation/as a group/at once
(68) *Every plane landed together/in formation/as a group/at once

By "collectivizing" I mean an adverb that takes a predicate that applies to atomic individuals, such as *land*, and appears to turn it into a predicate that will apply only to pluralities: a single plane can land, but it cannot land together.

The difference between (67) on the one hand, and (68) on the other, will be explained as a difference between the scope of distributivity and the scope of a quantified expression. Distributivity can take low scope, over just the verb. This means it can take scope inside the 'scope domain'. A quantified expression, on the other hand, always takes scope outside

the scope domain. We will see that this means that the expression *every plane* is forced to take scope that is too wide to give a sensible interpretation when it is combined with a VP containing one of these collectivizing adverbials.

To see how this works, we first need an analysis of collectivizing adverbials. I propose that *together* involves a kind of quantification over individual parts of a plural DP; this is contrary to a proposal by Lasersohn (1995) in which the quantification introduced by *together* is over events. However, I borrow from Lasersohn the idea of events that "overlap". The idea is that a sentence with *together* is true if the event that the sentence picks out is an event that has the right kind of subevents, and those subevents overlap along some dimension. For example, we may say that an event of John and Mary sitting is a 'together' event if the subevents of that event, namely the event of John sitting and the event of Mary sitting, overlap in space and time. Lasersohn discusses several different ways that events can overlap, including space, time, and a "social accompaniment" reading of *together*. For now let's suppose that the relevant notion of overlap for our planes example is overlap in terms of space and time, which I'll write as τ. Then we can give the following translation for VP-attached *together*.

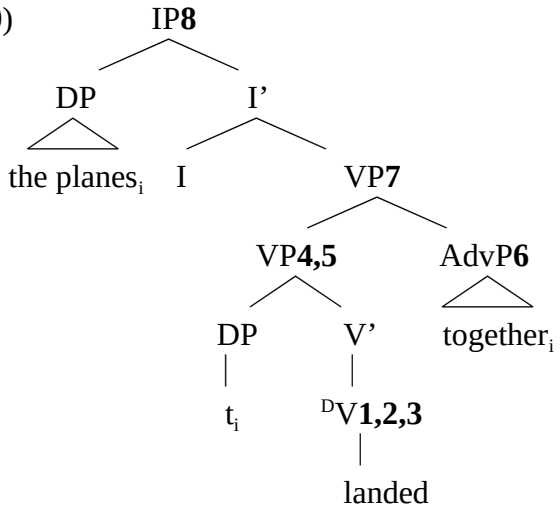
(69) $\lambda P \lambda x \lambda e[P(x)(e) \& \forall y,y',e',e''[y \in x \& y' \in x \& y \neq y' \& P(y)(e') \& P(y')(e'') \& e' \subseteq e \& e'' \subseteq e \rightarrow \tau(e') \circ \tau(e'')]]$

This definition is very long. To make our derivations easier to read I will adopt the convention of starting a new line for the meaning of *together* in the derivations.

Now let's apply our interpretation of *together* to a sentence like (68). We have already said that distributivity can apply to a VP. If we take the structure in (70), and apply distributivity at the level of the lowest VP, then the derivation can proceed as given in (71).

(Since *land* is intransitive, we could equivalently apply D to V.) (Note that I assume that the trace t_i contributes a variable x_i to the derivation, and that *together* is coindexed with the DP, licensing lambda abstraction over that variable. This is not crucial but it makes things simpler.)

(70)



(71)

1. λxλe[land'(e) & Ag(e,x)]
2. λPλx⊓{P(z) | z⊆x & z∈⟦Cov$_i$⟧}
3. λx⊓{λxλe[land'(e) & Ag(e,x)] (z) | z⊆x & z∈⟦Cov$_i$⟧}
   =λx⊓{λe[land'(e) & Ag(e,z)] | z⊆x & z∈⟦Cov$_i$⟧
4. ⊓{λe[land'(e) & Ag(e,z)] | z⊆x$_i$ & z∈⟦Cov$_i$⟧}
5. λx$_i$ [⊓{λe[land'(e) & Ag(e,z)] | z⊆x$_i$ & z∈⟦Cov$_i$⟧}]
6. λPλxλe[P(x)(e) & ∀y,y',e',e"[y∈x & y'∈x & y≠y' & P(y)(e') & P(y')(e") & e'⊆e & e"⊆e → τ(e') o τ(e")]]
7. λxλe [⊓{λe[land'(e) & Ag(e,z)] | z⊆x & z∈⟦Cov$_i$⟧}](e) &
   ∀y,y',e',e"[y∈x & y'∈x & y≠y' & land'(e') & Ag(e',y) & land'(e") & Ag(e',y') & e'⊆e & e"⊆e → τ(e') o τ(e")]]

8. λe[⊓{λe[land'(e) & Ag(e,z)] | z⊆⟦the.planes'⟧ & z∈⟦Cov$_i$⟧}](e) &
   ∀y,y',e',e"[y∈⟦the.planes'⟧ & y'∈ ⟦Cov$_i$⟧ & y≠y' & land'(e') & Ag(e',y) & land'(e") & Ag(e',y') → e'⊆e & e"⊆e & τ(e') o τ(e")]]

To see in more detail what this means, first notice that the contribution of the D operator and the contribution of *together* can be treated separately, because the meaning of

together is conjoined to the meaning of the VP. So first we will pay attention just to the distributive "part" of line 8 (namely, the first line).

First we need to consider what set *F* the generalized conjunction operator takes as its argument. Here it will be the set of functions *f'* that you get by substituting every individual plane that is part of ⟦the.planes'⟧ in for the Agent argument of *land*. So *F*={λe[land'(e) & Ag(e,plane1), λe[land'(e) & Ag(e,plane2)]...} and so on (modulo pragmatic weakening, to repeat the familiar refrain).

So this means that the function yielded by the generalized distributivity operator applied to this set could be equivalently written as (72) (recall that ⊓ of sets of functions of type <v,t> introduces existential quantification over parts of an event).

(72) $\sqcap\{\lambda e[land'(e) \& Ag(e,z)] \mid z \subseteq [\![the.planes']\!] \& z \in [\![Cov_i]\!]\} =$
$\lambda e'''\forall x \exists e'[x \subseteq [\![the.planes']\!] \& x \in [\![Cov_i]\!] \rightarrow land'(e') \& Ag(e',x) \& e' \subseteq e''']$

If we use this equivalence in line 8 of the derivation above, we get the expression in (73).

(73) $\lambda e[\lambda e'''[\forall x \exists e'[x \subseteq [\![the.planes']\!] \& x \in [\![Cov_i]\!] \rightarrow land'(e') \& Ag(e',x) \& e' \subseteq e'''] (e) \&$
$\forall y,y',e',e''[y \in [\![the.planes']\!] \& y' \in [\![the.planes']\!] \& y \neq y' \& land'(e') \& Ag(e',y) \& land'(e'') \& Ag(e',y') \& e' \subseteq e \& e'' \subseteq e \rightarrow \tau(e') \circ \tau(e'')]]$

Lambda conversion turns (73) into (74).

(74) $\lambda e[\forall x \exists e'[x \subseteq [\![the.planes']\!] \& x \in [\![Cov_i]\!] \rightarrow land'(e') \& Ag(e',x) \& e' \subseteq e \&$
$\forall y,y',e',e''[y \in [\![the.planes']\!] \& y' \in [\![the.planes']\!] \& y \neq y' \& [land'(e') \& Ag(e',y) \& land'(e'') \& Ag(e',y') \& e' \subseteq e \& e'' \subseteq e \rightarrow \tau(e') \circ \tau(e'')]]$

Finally, we apply the familiar operation of existential closure to arrive at the proposition in (75).

(75) $\exists e \forall x \exists e'[x \subseteq [\![the.planes']\!] \& x \in [\![Cov_i]\!] \rightarrow land'(e') \& Ag(e',x) \& e' \subseteq e \&$
$\forall y,y',e',e''[y \in [\![the.planes']\!] \& y' \in [\![the.planes']\!] \& y \neq y' \& land'(e') \& Ag(e',y) \& land'(e'') \& Ag(e',y') \& e' \subseteq e \& e'' \subseteq e \rightarrow \tau(e') \circ \tau(e'')]]$

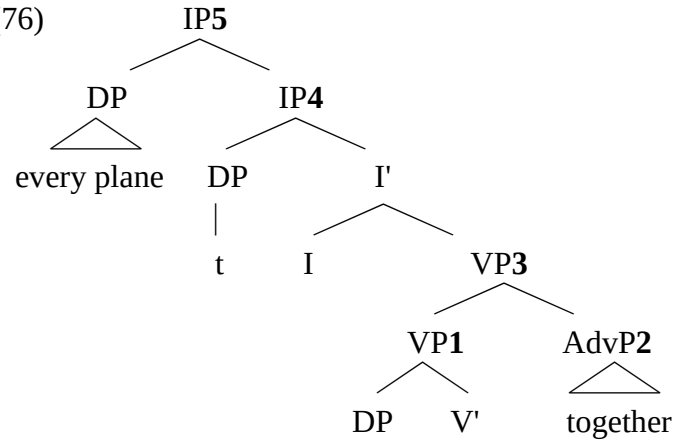
The expression in (75) says that there is an event which has as its subevents a landing event for each plane (modulo pragmatic weakening), and that these subevents are in the right sort of relation (spatio-temporal overlap) to make the big event e an event that is "together".

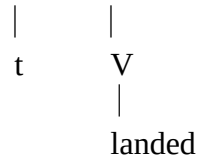
The derivation of (75) has two instances of universal quantification over the parts of ⟦the.planes'⟧: one introduced by the D operator, and one introduced by *together*. But crucially, neither of them has scope over the other, so neither of them "interferes" with the other.

If, on the other hand, the universal quantifier *every plane* is subject to the SDP, then it must QR and it will take wide scope over the universal quantifier(s) introduced by *together*. The result will be that *together* fails to have the right kind of object to quantify over, and hence the sentence is ill-formed.

Take the structure of (68) to be (76), where *every plane* has undergone QR so that it can take scope outside of the scope domain. Then the derivation will work as given below in (77).

(76)





(77)
1. λxλe[land'(e) & Ag(e,x)]
2. λPλxλe[P(x)(e) & ∀y,y'e',e''[y∈x & y'∈x & y≠y' & land'(e') & Ag(e',y) & land'(e'') & Ag(e',y') & e'⊆e & e''⊆e → τ(e') o τ(e'')]]

3. λxλe[land'(e) & Ag(e,x) &
∀y,y',e',e''[y∈x & y'∈x & y≠y' & land'(e') & Ag(e',y) & land'(e'') & Ag(e',y') & e'⊆e & e''⊆e → τ(e') o τ(e'')]]

4. ∃e[land'(e) & Ag(e,x) &
∀y,y',e',e''[y∈x & y'∈x & y≠y' & land'(e') & Ag(e',y) & land'(e'') & Ag(e',y') & e'⊆e & e''⊆e→ τ(e') o τ(e'')]]

5. ∀x[plane'(x) → ∃e[land'(e) & Ag(e,x) &
∀y,y'e',e''[y∈x & y'∈x & y≠y' & land'(e') & Ag(e',y) & land'(e'') & Ag(e',y') & e'⊆e & e''⊆e → τ(e') o τ(e'')]]

Now the problem should be clear. *Every* quantifies over the set of planes, but since *together* is inside its scope *together* is forced to quantify over a singleton domain.

This problem does not arise with *all* because, as we have seen, the scope of the distributivity operator is low enough that it does not interfere with the quantification introduced by *together*.

I suggest that other collectivizing adverbials, such as *as a group*, *at once*, and *in formation* have a similar meaning. Their contribution to the meaning of a sentence is to quantify over parts of an individual and parts of an event.

I have argued that the difference between the sentences in (67), with collectivizing adverbials and *every*, and the sentences in (68), with collectivizing adverbials and *all*, is really a difference between the scope that a distributivity operator can take and the scope

that *every* can take. The scope of *every* has to be outside the scope domain and hence outside the VP where the adverbial is adjoined.

There is no obvious account for this difference between *all* and *every* if we treat *all* as a universal quantifier. But it follows from the proposal I have already made to capture the maximizing effect of *all*.

In the next section we will return to predicates that are lexically collective, and see how we can explain the differences between *all* and *every* with these predicates. We will see again that scope plays a role. This in turn will lead us to an explanation for Taub's generalization.

4.4 Lexically Collective Predicates

By 'lexically collective predicates' I mean predicates that do not need a collectivizing adverbial to be interpreted collectively. These include predicates like the following:

(78) The students gathered in the hallway
(79) The boys are a big group
(80) The girls built a raft (*allows a collective reading*)

These predicates cannot be combined with *every*; however, some of them can be combined with *all*.

(81) a. ?Every student gathered in the hallway
b. All the students gathered in the hallway

(82) a. *Every boy is a big group
b. *All the boys are a big group

(83) a. Every girl built a raft (*distributive only*)
b. All the girls built a raft (*collective or distributive*)

As we discussed in Chapter 2, the distribution of *all* with collectives is somewhat curious because it can occur with some collective predicates but not others. There I

proposed an analysis of how collective predicates can occur with *all*, but it failed to predict that some collective predicates cannot. Here I will propose a different analysis of these cases, and we will crucially use Taub's generalization to point the way toward an explanation. Recall that Taub's generalization is (84).

(84) **Taub's Generalization**: The collective predicates that allow *all* are the collective activities and accomplishments; collective states and achievements do not allow *all*.

We have seen the data that Taub gives in support of this generalization before, but I repeat it below for the reader's convenience.

Collective states:
(85) *The boys are all a big group
(86) *The trees are all dense in the middle of the forest (*on collective *dense*)

Collective activities:
(87) All the boys carried the piano around for an hour

Collective accomplishments:
(88) All the students gathered in the hallway
(89) All the girls built a raft

Collective achievements:
(90) *All the senators passed the pay raise
(91) *All the students elected a president

Recall also that, as I pointed out in Chapter 3, *except*-phrases show the same distribution as *all* in this respect, as shown in (92)-(95). This gives us additional evidence for the significance of Taub's generalization.

(92) *The boys are a big group, except for Jason
(93) The boys carried the piano around for an hour, except for Alex
(94) The students gathered in the hallway, except for Sarah
(95) *The senators passed the pay raise, except for Lautenberg

Finally, we also saw that the predicates of Taub's generalization do not allow pragmatic weakening.

(96) *The boys are a big group, but not Jason/but not Jason and Alex
(97) The boys carried the piano around for an hour, but Alex was sleeping
(98) The students gathered in the hallway, but not Sarah
(99) *The senators passed the pay raise, but not Lautenberg/ but not Lautenberg and Kennedy

Clearly, then, the best account of Taub's generalization for *all* would also provide an explanation for why it also holds for the distribution of *except* phrases and for pragmatic weakening.

So the question is, how can we account for Taub's generalization? To answer this question, I think we have to ask what activities and accomplishments on the one hand, or states and achievements on the other, have in common. And the literature on lexical aspect gives us a pretty clear answer to the first question.

4.4.1 Lexical aspect: Do

Many authors (Dowty 1979, Mittwoch 1982, Pustejovsky 1991, Grimshaw & Vikner 1993 and others) propose that activities and accomplishments have in common a subcomponent of their meaning that is, in a sense that is made precise differently in different theories, an "activity." This component has various names, including "activity", "process" and "DO". Here I will adopt the term DO. Following Mittwoch (1982) I treat DO as a predicate (rather than, for example, an operator or a modifier, two options considered in Dowty 1979).

I treat DO as a kind of bleached out predicate that applies to processes of all sorts: many things can count as a DOing. Some predicates more strictly lexically specify what activities can count as their DO part. For example, a predicate like *sweep the floor* will pretty much only allow moving a broom back and forth across the floor as part of its DO.

On the other hand, a predicate like *build a raft* has to allow a large variety of things to count as DOings: hammering, sawing wood, etc, because all of these things help make up the process part of building a raft. This difference is, I take it, simply part of the lexical meaning of the verb.

Accomplishments and activities differ in that accomplishments are postulated to contain in addition to DO a subcomponent that is a stative predicate — a predicate naming a state that results from DOing the activity named by the verb. So, for example, the predicate *build a house* has two parts, which can be schematized as in (100).

(100) build a house
/ \
DO state

Many of the proposals for instantiating this idea include reference to a BECOME operator or a CAUSE operator (or both) (see, for example, Dowty 1979, Pustejovsky 1991). However, since the existence or nonexistence of these operators does not bear on the issues to be addressed here we will put this possibility aside. The important issue for us here is the existence of the activity component of the verb.

One frequently cited argument in favor of introducing the structure in (100) is the *almost* test. This argument notes that the sentence in (101) is ambiguous.

(101) Karen almost built a house.

The sentence can mean either that Karen was thinking about building a house but never actually did anything, or it can mean that she actually did do some building — maybe put down a foundation and built a frame — but that the house did not get finished. Proponents of the structure in (100) point out that this structure gives us a way to account

for this ambiguity: *almost* can take scope over the entire predicate, *build a house*, leading to the first reading; or it can take scope just over the result state part, leading to the second reading. Without this internal structure we would not be able to give this analysis for the ambiguity.

The argument for including DO as a subpart of activity predicates comes from the fact that activities and accomplishments show similar behavior with respect to a wide variety of aspectual tests. Here is just one of them: cooccurrence with durative adverbials. Consider the sentence in (102).

(102) Kate was building a house for three years

The adverbial phrase *for three years* does not modify the result state: this sentence does not mean that the house existed for three years. Rather, it means that the process of building went on for three years. So durative adverbial phrases modify processes, or that part of the predicate that we are calling DO.

Durative phrases also modify activities, as shown in (103)

(103) Kate was singing for twenty minutes

If we postulate that durative phrases modify processes or DO, then the simplest way to account for (103) is to say that it, too, has a process component to its meaning. (For more discussion on the similarities between activities and accomplishments, see e.g Mittwoch 1982, McClure 1994.)[7]

---

[7] There is a well-known phenomenon in which the classification of a sentence as an activity or an accomplishment can be affected by the form of the verb's direct object. Bare plural objects (and mass nouns) license an iterative reading that behaves like an activity.

(i) John built houses (for three years)

The existence of this phenomenon does not bear on our concerns here, since the alternation is between activity and accomplishment and we are interested only in what activities and

4.4.1.1 Projecting aspect into syntax

Here I will propose that the DO component of activities and accomplishments is projected into the syntax as a kind of aspectual head that takes the phrase projected by the lexical head of the verb (for instance, *build*) as its sister. This proposal is actually a version of an idea that has seen many versions. With respect to the aktionsart literature, Pustejovsky (1991) makes a similar proposal but his structures are at the level of lexical conceptual structure, not syntax. Dowty (1979) cites a paper by Ross (1972), in which Ross proposes that "every verb of action is embedded in the object complement of a two-place predicate whose phonological realization in English is *do*." (Ross, p.70, quoted from Dowty, p.111), couched in the Generative Semantics framework.

In addition, the sort of syntactic structure I am proposing has been widely used in the syntactic literature. I probably couldn't do justice to the many uses to which the idea has been put but here is a sampling. Hale and Keyser (1987, 1993) use an empty verbal head in their analysis of middle constructions. Larson (1988) proposed the "VP shell" analysis of English double object constructions, in which an empty verbal head dominates the phrase projected by the lexical head. Speas (1990) proposes that agentive predicates project an empty V head. Johnson (1991) uses the idea to account for some properties of adverb placement with transitive (ie single-object) verbs.

The proposal I will make here is probably closest to the syntactic proposals in McClure (1994). McClure proposes two aspectual projections above VP that are licensed

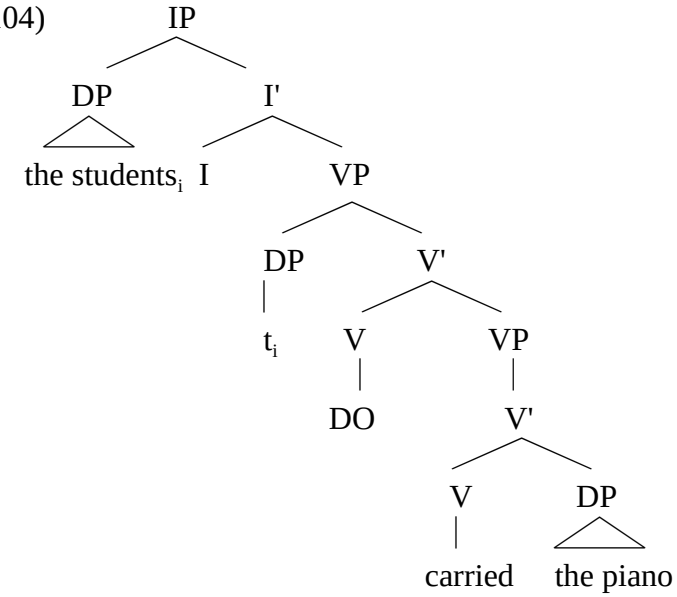
---

accomplishments have in common. For more discussion of this phenomenon see Tenny (1987) and Krifka (1992).

(in part) by the lexical content of the verb: for him DO occupies the higher of these two projections. Likewise, I propose that DO is a kind of aspectual head that is licensed by the semantics of the lexical head (although the meaning of my DO is quite different from the meaning of McClure's DO, which is described in terms of a different theory of aspect).

The structure I propose for activities and accomplishments is one in which the verb projects two heads into the syntax, and the higher head is an aspectual head that contains the DO portion of the predicate. The higher head is phonologically empty, but it contributes the meaning of DO to the verb (and to the sentence) and it is the predicate that takes the subject (usually, but not always, the Agent) as its argument. The structure of a sentence like *the students carried the piano*, then, looks like this:

(104)



Two syntactic issues should be immediately addressed. First, in my examples I show that the only functional head above DO is IP. However, I assume that in general DO is below any other functional projections that dominate VP, and the above structure can be

adapted to a syntactic theory in which the functional structure of the clause is more articulated (ie, where AgrOP or TP, for example, are hypothesized to be present) without affecting the basic points I intend to make here.

Secondly, one may wonder whether *carry* incorporates into DO, either at the level of overt syntax or at LF. Certainly it is the case that in most if not all of the analyses cited above, the lexical verb does incorporate into the head of the verbal "shell". For Larson's analysis it is crucial that the verb should raise because otherwise his analysis predicts the wrong word order, and he motivates the movement in terms of case theory. Many of the other analyses do not have the word-order imperative and so the reasons for postulating head movement are basically theory-internal (eg, again motivated by case considerations). It is not crucial to us whether head-movement occurs or not; what is crucial is that even if head movement occurs we treat DO and the lexical verb as distinct (even if incorporated) objects. So in the derivations I won't show incorporation but we could allow it if there were good reasons to.

I propose that the lexical semantics of the predicate *carry*, for example, is divided between the two verbal heads as follows. Each verbal head has two equivalent translations (just as we saw earlier that *sing* has two equivalent translations).

(105) DO = λxλe'[DO(e) & Ag(e,x)]

(106) carry' = λxλe[carry'(e) & Th(e,x)]

Given the structure proposed in (104), these two predicates must combine in some way after *carry* combines with its theme argument. The method of combining them that I will propose here is inspired by Kratzer's (1994) operation of "event identification".

4.4.1.2 Event identification

Based on the evidence from many theories of argument structure (see, e.g Williams 1981 and Grimshaw 1990) that the external argument of a predicate has a special status, Kratzer proposes a "severing" of the external argument from its verb. She argues that verbs in fact do not have an external argument, and so the predicate *kiss*, for example, takes as its arguments only its Theme argument and event argument. The external argument is introduced by an independent functional head which Kratzer dubs Voice (and its phrase VoiceP) which is combined with the predicate by means of secondary predication.

The process of secondary predication is "event identification". In event identification, the open event argument places of two predicates are "identified"; the predicates "fuse" in a way to form a single expression with a single open event argument slot. (The term is supposed to reminiscent of Higginbotham's 1985 "theta identification".) Kratzer's definition of event identification is given in (107).

(107) Event Identification:

f	g	->	h
<i<v,t>>	<v,t>		<i<v,t>>
			λxλe[f(x)(e) & g(e)]
λxλe[Ag(e,x)]	λe[wash'(clothes')(e)]		λxλe[Ag(e,x) & wash'(clothes')(e)]

Event identification in the form Kratzer proposes won't work for us. To see why, refer to the structure given in (104). Whereas Kratzer's "f" is a function of the form λxλe[Ag(e)(x)], our "f," which is the DO part of the predicate, also includes the predicate of events DO, as in λxλe[DO(e) &Ag(e,x)]. If we combine this expression with an expression like λe[carry'(e) & Th(e,piano')], then we would end up claiming that the DOing

and the carrying are the same thing: event identification would give us λxλe[carry'(e) & Th(e,piano) & DO(e) & Ag(e,x)] as the meaning of the higher VP in (104).

But this does not seem right. It's not the case that a DOing is a carrying, DOing is a part of a carrying. We would like our semantics to express this fact. So I will propose that "subordinating" the DO part of the event is part of the operation of event identification. Given this perhaps it would be more appropriate to call the operation I will use here event composition. Event composition would work as follows.

(108) **Event composition:**

f		g		h
<v,t>		<i,<v,t>>		<i,<v,t>>
λe[carry'(e) & Th(e,piano)]		λxλe[DO(e) & Ag(e,x)]	->	

λxλe[carry'(e) & Th(e,piano') & ∃e'[DO(e') & Ag(e',x) & e'◂e]

In fact since we are not "identifying" the events any more, we could also define event composition as an operator that can be used whenever we want to combine functions of the right type. The operator would be defined as follows, where Q is a variable over objects of type <v,t>.

(109) λQλPλxλe[Q(e) & ∃e'[P(x)(e')] & e'◂e]

It doesn't really matter here which option we choose. The important point is that event composition is a process that combines two expressions of the right type. Like Kratzer does, I will simply stipulate that event composition is available whenever we want to combine two expressions of the types corresponding to "f" and "g".

Given the definition of the event composition operator in (109), we might wonder whether it is better to do away with a separate operator, and treat one or the other of the verbal heads as having "event composition" as part of its meaning. For example, when we

combine the meaning of *carry the piano* with the operator in (109), we get the expression in (111).

(110) λQλPλxλe[Q(e) & ∃e'[P(x)(e')] & e'◂e] (λe[carry'(e) & Th(e,piano)]) =

(111) λPλxλe[carry'(e) & Th(e, piano) & ∃e'[P(x)(e')] & e'◂e]

We could in principle start out by proposing that (108) (with x instead of *piano* as Theme) is the meaning of *carry*. Alternatively, we could combine the meaning of DO from (118) with the event composition operator (by switching the order of the λ-operators), and propose that that is the lexical meaning of DO. Why shouldn't we choose one of these options, instead of proposing a separate operation of event composition?

There are two reasons. One is that if we were to do so, it would change the empirical predictions of the proposal I will make shortly in such a way that it wouldn't correctly account for the facts. The other is that I would like to keep as close as possible to the spirit of the idea proposed here, which is that the meaning of an activity and an accomplishment predicate is composed from two separate syntactic heads. If we hypothesize that the lowest verbal head already contains a "space" for the higher verbal head, then we would seem to rob the proposal that there are two separate heads of much of its interest.

The idea that the DO head introduces the external argument of a verb is certainly consistent with the spirit of Kratzer's proposal. She explicitly suggests that the head that introduces the external argument may carry aspectual information.

On the other hand, some of her arguments for severing the external head from the verb might be problematic for the proposal I've made here. Kratzer argues that the external argument is unlike true lexically specified properties of the verb in that it doesn't necessarily

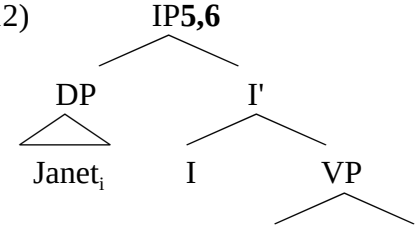
have to be there whenever we use the verb. One kind of evidence for this claim comes from gerunds like *Maria's reading of* Pride and Predjudice *was quite good*. Kratzer points out that the agent of the reading does not have to be specified. (Note that it doesn't have to be Maria: if Maria and Kelly went to two different readings; then "Maria's reading" can be the reading that Maria attended.)

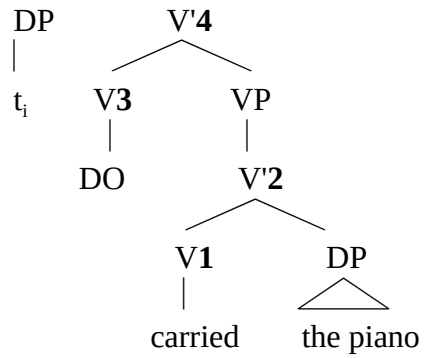
The apparent problem for me is that the external argument and the process part of the event are contributed together, so if I'm right and Kratzer is right this would seem to suggest that the event denoted by *Maria's reading* has no process part. Preliminary evidence suggests that this is wrong; for example, we can say *Maria's reading took two hours*. But I won't take this as knock-down evidence that I have to abandon the proposal I've made here. For one thing, it is clear that the event of Maria's reading does have some understood agent — the agent isn't completely absent (in this way it is similar to the passive); so the process isn't necessarily completely absent either. For now I will just say that these issues warrant future investigation.

4.4.1.3 A new analysis of "lexically collective" predicates

At this point we have enough tools in hand to see how the proposed derivation of a sentence with an activity or accomplishment predicate, with the structure hypothesized in (104), will proceed. For now I will use an example with singular DPs to keep things simpler.

(112)





1. λxλe[carry'(e) & Th(e,x)]
2. λe[carry'(e) & Th(e,the.piano')]
3. λxλe[DO(e) & Ag(e,x)] *(2 and 3 combine via event composition)*
4. λxλe[carry'(e) & Th(e,the.piano') & ∃e'[DO(e') & Ag(e',x) & e'◂e]]
5. λe[carry'(e) & Th(e,the.piano') & ∃e'[DO(e') & Ag(e',J) & e'◂e]]
6. ∃e[carry'(e) & Th(e,the.piano') & ∃e'[DO(e') & Ag(e',J) & e'◂e]]

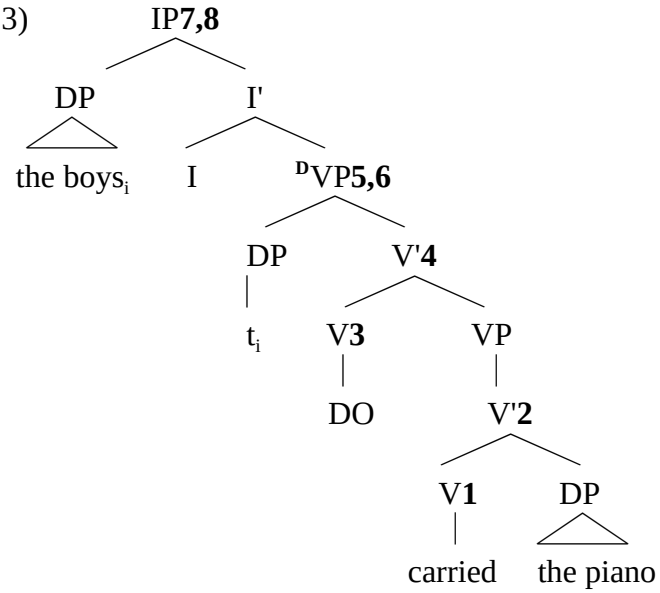
The expression in line 6 of (112) gives us the truth conditions that we want. It says that there is an event of carrying the piano in which Janet is the agent of its DO subpart. Notice that the notion of an event for a proposition does some work for us here. Without it, this sentence might count as true by virtue of an event in which Janet and Bob carry the piano — because it would still be true that Janet is agent of some DO subpart of some carrying-the-piano event. But because we adopt the notion of an event for a proposition, the event of Janet and Bob carrying the piano cannot make this sentence true because the event of Bob being agent of a DO subpart is not part of a minimal event of Janet being agent of a DO subpart. Thus only if there is an event in which Janet did all the carrying herself will the proposition be true.

Next we can see how our D operator interacts with the syntax and semantics we have just proposed for activity and accomplishment predicates. We will see that the distinction between distributive and collective readings is captured by the two possible insertion sites

for a D operator: on DO, and on the VP dominating DO. (Note that the VP dominating the lexical predicate is not a possible insertion site for D because it does not take an individual as its argument.) The reading that we have been calling "collective" will turn out to actually contain a sort of hidden distributivity.

First let us look at the derivation where D is on the highest VP. It proceeds almost exactly like the derivation we have just seen, with a singular subject.

(113)



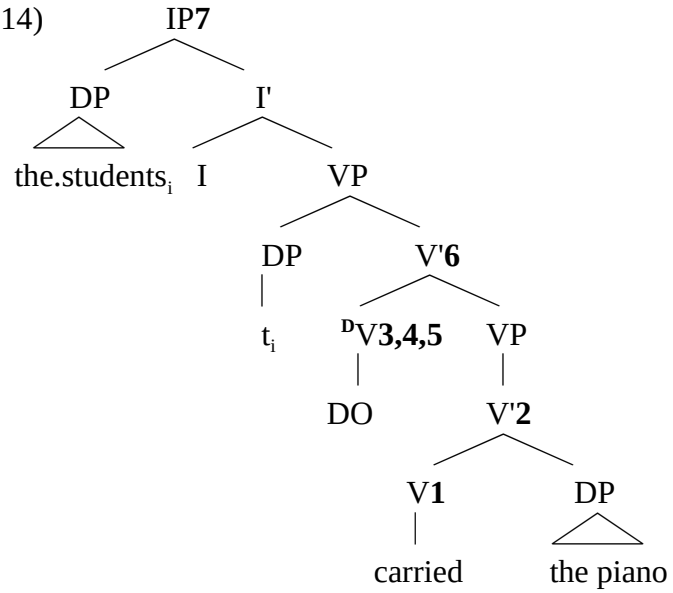
1. λxλe[carry'(e) & Th(e,x)]
2. λe[carry'(e) & Th(e,the.piano')]
3. λxλe'[DO(e') & Ag(e',x)] (*2 and 3 combine via event composition*)
4. λxλe[carry'(e) & Th(e,the.piano') &∃e'[DO(e') & Ag(e',x) & e'◂e]]
5. λPλx⊓{P(z) | z⊆x & z∈⟦Cov$_i$⟧}

6. λx ⊓{λxλe[carry'(e) & Th(e,the.piano) &∃e'[DO(e') & Ag(e',x) & e'◂e]] (z) | z⊆x & z∈⟦Cov$_i$⟧}
7. ⊓{λxλe[carry'(e) & Th(e,the.piano) &∃e'[DO(e') & Ag(e',x) & e'◂e]] (z) | z⊆⟦the.boys'⟧ & z∈⟦Cov$_i$⟧}
7' = ⊓{λe[carry'(e) & Th(e,the.piano') &∃e'[DO(e') & Ag(e',z) & e'◂e]]| z⊆⟦the.boys'⟧ & z∈⟦Cov$_i$⟧}
8. =λe∀x∃e''[z⊆⟦the.boys'⟧ & z∈⟦Cov$_i$⟧→ carry'(e'') & Th(e'',the.piano) &∃e'[DO(e') & Ag(e',z) & e'◂e'' & e''⊆e]
8.' =∃e∀x∃e''[z⊆⟦the.boys'⟧ & z∈⟦Cov$_i$⟧→ carry'(e'') & Th(e'',the.piano) &∃e'[DO(e') & Ag(e',z) & e'◂e'' & e''⊆e]

This sentence asserts the existence of a separate carry-the-piano event for each one of the boys (modulo pragmatic weakening). This, of course, is the "distributive" reading of the sentence.

The more interesting case is the case where the D operator is inserted on DO. This will yield what we have been calling a collective reading.

(114)



1. λxλe[carry'(e) & Th(e,x)]
2. λe[carry'(e) & Th(e,the.piano')]
3. λxλe[DO(e) & Ag(e,x)]
4. λPλy⊓{P (z) | z⊆x & z∈⟦Cov$_i$⟧}
5. λy⊓{λxλe[DO(e) & Ag(e,x)](z) | z⊆y & z∈⟦Cov$_i$⟧}

5' = λy⊓{λe[DO(e) & Ag(e,z)] | z⊆y & z∈⟦Cov$_i$⟧}

to make event composition easier, we will use the following notational variant of 5':

5" = λyλe[e ∈ ⊓{λe[DO(e) & Ag(e,z)] | z⊆y & z∈⟦Cov$_i$⟧}]

2 and 5" combine via event composition

6. λyλe[carry'(e) & Th(e, the.piano') & ∃e' ∈ ⊓{λe[DO(e) & Ag(e,z)] | z⊆y & z∈⟦Cov$_i$⟧}] & e'◂e]
7. ∃e[carry'(e) & Th(e, the.piano') & ∃e' ∈ ⊓{λe[DO(e') & Ag(e',z)] | z⊆⟦the.students'⟧ & z∈⟦Cov$_i$⟧} & e'◂e]

Now, as we have done in previous derivations, we will "unpack" the function given

by the generalized conjunction operator into something that looks, notationally, more like what we are used to in event semantics.

Let us suppose that the students are Harry, Bill, and Tom. Then the set *F* that the ⊓ takes as its argument is the set in (115), namely, the set of functions we get by substituting each one of the students for the Agent argument of DO.

(115) {λe[DO(e) & Ag(e,H)], λe[DO(e) & Ag(e,B)], λe[DO(e) & Ag(e,T)] }

Because the generalized conjunction operator introduces existential quantification over events when it applies to sets of functions of type <v,t>, when it applies to the set in (115) it will yield the following function.

(116) λe'"∀x∃e"[x⊆⟦the.students'⟧ & x ∈ ⟦Cov_i⟧ → [DO(e") & Ag(e",x)] & e"⊆e'"]

Now we can plug this back in to the formula in line 7 of the derivation in (114).

(117) line 7:
∃e[carry'(e) & Th(e,the.piano') & ∃e' ∈ ⊓{λe[DO(e') & Ag(e',z)]
| z⊆⟦the.students'⟧ & z∈⟦Cov_i⟧} & e'◂e]]

= ∃e[carry'(e) & Th(e,the.piano') & ∃e' ∈ λe'"∀x∃e"[x⊆⟦the.students'⟧ & x ∈ ⟦Cov_i⟧ → [DO(e") & Ag(e",x)] & e"⊆e'" & e'◂e]]

= ∃e[carry'(e) & Th(e,the.piano') & ∃e'∀x∃e"[x⊆⟦the.students'⟧ & x ∈ ⟦Cov_i⟧ → [DO(e") & Ag(e",x)] & e"⊆e' & e'◂e]]

Again, the last expression in (117) gives us the truth conditions we are looking for. It says that there is an event of carrying the piano, which has a complex DO subpart: its DO subpart is actually a plural event consisting of a separate DOing event for each one of the students (modulo pragmatic weakening).

So what we have found is that when we introduce distributivity on DO, we get a reading that is essentially equivalent to what we have been calling the collective reading.

The difference is that we assert that there is a separate event of DOing for each individual student. The use to which I have put DO here should sound quite reminiscent of Dowty's idea of distributive subentailments, discussed in section 1.4.1 of Chapter 1. The difference is that here we have not relied on the presence or absence of *all* as a test for the presence of subentailments; subentailments here are attributed to the presence of a DO predicate as a subpart of the lexical representation of the accomplishment predicate.

What I have done here is to use Taub's generalization to point the way toward a better-motivated proposal about what 'distributive subentailments' could be. Dowty proposed distributive subentailments but he couldn't give evidence for them; Taub provides evidence that the lexical aktionsart of a predicate plays a role in licensing *all* but doesn't give an account of how the two could be semantically connected. Here we are using the structure independently proposed for activities and accomplishments to give some concreteness to the idea of distributive subentailments: a distributive subentailment is just an event of DOing.

But what does all of this have to do with the distribution of *all*?

On the proposal I've given here, the kind of collectivity that we observe with an activity or an accomplishment predicate is quite similar to the kind of collectivity we observed with collectivizing adverbials like *together*; it actually contains a sort of hidden distributivity. This predicts that *all* should be possible with collective accomplishments and activities, because *all* is dependent on the D operator for its functioning. Of course we already know that this is true.

4.4.2 All *and lexically collective predicates*

Showing how *all* can combine with activity and accomplishment predicates to yield a distributive or a collective reading is at this point trivial. We have already seen how a D operator is implicated in both readings. Since *all* is dependent on a D operator, the derivation of a distributive or a collective reading with *all* will be exactly the same as the derivations for distributive and collective readings that we saw in the previous section.

For example, I argued above that when D is inserted on the higher VP of a sentence like the one in (118), the derivation will yield a distributive reading, represented by (119).

(118) The students carried the piano

(119) $\exists e \forall x \exists e''[z \subseteq [\![\text{the.boys}']\!] \ \& \ z \in [\![\text{Cov}_i]\!] \rightarrow \text{carry}'(e'') \ \& \ \text{Th}(e'', \text{the.piano}') \ \& \exists e'[\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \blacktriangleleft e'' \ \& \ e'' \subseteq e]]$

When D is inserted on the aspectual head DO of a sentence like (118), the derivation will yield a collective reading, represented by (120).

(120) $\exists e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \forall x \exists e''[x \subseteq [\![\text{the.students}']\!] \ \& \ x \in [\![\text{Cov}_i]\!] \rightarrow \text{DO}(e'') \ \& \ \text{Ag}(e'', x) \ \& \ e'' \subseteq e' \ \& \ e' \blacktriangleleft e]]$

We expect the same ambiguity to be available for *all*, since the only effect *all* has on the semantics of the sentence is to force the value of *Cov* to be a good fit.

So when D is inserted on the higher VP of a sentence like (121), with *all*, the derivation will yield a distributive and "maximized" reading, represented by (122).

(121) The students all carried the piano

(122) $\exists e \forall x \exists e''[z \subseteq [\![\text{the.boys}^{gf}]\!] \ \& \ z \in [\![\text{Cov}_i]\!] \rightarrow \text{carry}'(e'') \ \& \ \text{Th}(e'', \text{the.piano}) \ \& \exists e'[\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \blacktriangleleft e'' \ \& \ e'' \subseteq e]]$

Of course this same sentence can be interpreted collectively and "maximally", if D is inserted on DO.

(123) $\exists e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \forall x \exists e''[x \subseteq [\![\text{the.students}^{gf}]\!] \ \& \ x \in [\![\text{Cov}_i]\!] \rightarrow \text{DO}(e'')$

& Ag(e",x) & e"⊆e' & e'◂e]]

Thus the "maximizing effect" of *all* with the collective reading comes about because each individual student is asserted to an agent of a DO-ing event that is part of carrying event.

4.4.2.1 The contrast between all *and* every

From this discussion it should be clear why *every* doesn't allow a collective reading with accomplishment and activity predicates that are ambiguous between distributive and collective readings. It's impossible for *every* to have scope low enough to affect just the DO portion of the predicate. The SDP says that *every* must take scope outside of existential closure. So only the distributive reading is possible, which is shown in (125).

(124) Every girl built a raft

(125) ∀x[x∈⟦girl'⟧ → ∃e[built'(e) & ∃e'[DO(e') & Ag(e',x) & e'◂e] & Th(e,a.raft')]]

4.4.2.2 The "anti-distributive" use of together

Since we have talked about the meaning of *together* in section 4.3, there is another aspect to its meaning that should be discussed here. *Together* has a use that has been called "antidistributive" (Schwarzschild 1992b); that is, when *together* combines with the activity/accomplishment predicates that are ambiguous between a collective and a distributive reading, *together* forces the collective reading, as shown in (126).

(126) The students built a house together

It is not clear to me why *together* should have this effect.

On the account of the distributive/collective alternation of accomplishment

predicates that I have just proposed, we would expect the "antidistributive" reading to arise if the D operator and *together* have "scope" over DO. But we also expect that a distributive, spatiotemporal overlap reading should be possible if the D operator and *together* are inserted at the level of the higher VP. The spatiotemporal overlap reading would mean something like, each of the students built a house, but their house-buildings took place "together", ie, in spatiotemporal proximity (maybe they were working on a housing drive for Habitat for Humanity).

Yael Sharvit (p.c.) has tried to convince me that the spatiotemporal overlap reading is possible, if one works hard enough in a context like the Habitat for Humanity housing drive context mentioned above. But I think even if it is possible, it is clear that at least it is very strongly disfavored.

The analysis I have given shares with the analysis of Lasersohn (1995) the unfortunate feature that it predicts that "antidistributive" and ordinary distributive-but-overlapping readings should be equally available. (In Lasersohn's analysis this is the price to be paid for an analysis of *together* that accounts for its many uses, which we have not been concerned with here.) Since the analysis I have given does predict that the antidistributivity use of *together* should at least be allowed, I will suppose here that the spatiotemporal overlap use of *together* is so strongly disfavored for pragmatic reasons. Perhaps the possibility of overlap of the antidistributive sort is so salient that it is somewhat misleading to assert that there is overlap but it is of the perhaps less interesting spatiotemporal sort.

4.4.3 The other half of Taub's generalization: distributivity and economy

Now we can return to the question of why collective states and achievements do not allow *all*. I will argue that in most cases, collective states and achievements do not license any sort of distributivity, and hence give *all* nothing to operate on.

Let us consider how we might represent the sentence in (127).

(127) The boys are a big group
(128) ∃e[big.group'(e) & Pt(e,the.boys')]

The predicate *be a big group* is a state predicate. Stative predicates do not, by consensus in the aktionsart literature (see, e.g. Dowty 1979, Pustejovski 1991, McClure 1994), have any subparts like the DO part of activities and achievements. Therefore, if distributivity is going to apply, then it will have to apply to the VP.

If we try to put a D operator on the VP, however, we get results that are clearly wrong. (129) shows the result of putting a D operator on the VP of (128).

(129) $\exists e \forall x \exists e'[x \subseteq [\![\text{the.boys}']\!] \,\&\, x \in [\![\text{Cov}_i]\!] \rightarrow \text{big.group}'(e') \,\&\, \text{Pt}(e',x) \,\&\, e' \subseteq e\,]$

(129) says that there should exist a complex event whose subevents are a separate be-a-big-group event for each boy. But of course this is nonsense, because an individual boy cannot be a big group.

Of course, we might take the option of supposing that the value assigned to Cov contains a set that is equal to the denotation of *the boys*. But this wouldn't help us, because then we would have universal quantification over a singleton domain, which we argued led to ill-formedness in the *together* examples with *every*.

I propose that insertion of a D operator is prohibited here, because it violates the principle of economy: we can interpret (127) without distributivity, hence the use of a D operator adds superfluous extra structure. I will adopt an economy-inspired condition on

adding a D operator to a representation:

(130) Distributivity is permitted only when it is necessary.

If (130) holds, then if we don't need distributivity to interpret (127), we can't have it. So we conclude that (129) is not a licit way to interpret the sentence in (127), only (128) is.[8]

If this is the case, then we expect that *all*, which depends on distributivity, should not be possible with (127). This is of course the case, as we have already seen, and as Taub's generalization predicts.

(131) *All the boys are a big group

So we predict that *all* should not be possible with collective states, because collective states are genuine cases of 'group' predication. There is no distributivity present for *all* to operate on.

A similar analysis holds for the collective achievements. Suppose that the interpretation of (132) is as represented in (133).

(132) The students elected a president
(133) ∃e[elect'(e) & Pt(e,the.students') & Th(e,a.president')]

Again, abstracting away from such matters as the possible presence of a BECOME operator with achievements, (133) is a reasonable approximation; and if we insert a D

[8] Strictly speaking, there is one possibility I am glossing over here. If we interpret (127) with a D operator, where the value for Cov is ill-fitting, but still "collective" (ie many boys in one cell) it is arguable that this D operator would be doing some work for us in interpreting (127) (it would exclude some boys a la pragmatic weakening). Nevertheless, this is apparently not possible. There must be some other principle that is operating here to rule out this possibility. It might be that there is an additional assumption (in the form of a presupposition, perhaps) that a D operator with a good-fitting cover must be both possible and non-superfluous. I will leave this as a matter for further investigation.

operator on the V or the VP, we get (134) as a result.

(134) $\exists e \forall x \exists e'[x \subseteq [\![\text{the.students}']\!] \ \& \ x \in [\![\text{Cov}_i]\!] \rightarrow \text{elect}'(e') \ \& \ \text{Ro}(e',x') \ \& \ \text{Th}(e,\text{a.president}') \ \& \ e' \subseteq e]$

(134) says that we have a complex event whose parts are individual electing-of-a-president events by individual students. But this is not what the sentence in (132) means. The sentence means that the group, as a whole, elected a president. So again it appears we are dealing with genuine group predication, and distributivity is not licensed.

If distributivity is not licensed then we expect, as we have already seen, that *all* should also not be licensed, and we know that this is the case. So the unavailability of *all* with collective states and achievements is explained: there is no distributivity in the representation of a sentence with these predicates, not even the "hidden" kind that we saw with collective activities and achievements. Since *all* is dependent on distributivity, it is not permitted with collective state and achievement predicates. Furthermore, the absence of pragmatic weakening with these predicates is also explained, since pragmatic weakening is due to the presence of a D operator.

4.4.3.1 "Meaning Shift"

It is probably worth pointing out that I am not claiming that only activities and accomplishments can be ambiguous between a collective and a distributive reading. What I am claiming is that of those predicates that show this ambiguity, only activities and accomplishments will allow *all* on their collective reading. For example, (135) is a stative predicate that is ambiguous between a distributive and a collective reading.

(135) The bottles are too heavy to carry

This sentence can be interpreted to mean that each individual bottle in the set of

contextually salient bottles is too heavy to carry (modulo pragmatic weakening, of course). It can also mean that the bottles are too heavy to carry as a group, in virtue of their collective weight.

The proposal I have made here makes the prediction that if we combine this predicate with *all*, only the distributive reading should be possible. This prediction is correct.

(136) The bottles are all too heavy to carry.

As expected, (136) can only be interpreted distributively. Since the collective reading has gone away, I call this phenomenon "meaning shift".

Meaning shift has, in fact, already been noticed (with some puzzlement on the part of the authors) in the literature. For example, Dowty (1987) discusses the examples in (137)-(138).

(137) The trees are (all) denser in the middle of the forest *(att. to B. Partee)*
(138) The students (all) voted in favor of the proposal *(att. to B. Ladusaw)*

The sentences in (137)-(138), without *all*, are ambiguous. For example, (137) without *all* can mean either that the individual trees in the middle of the forest are thicker than the ones at the outside, a distributive reading; or that the trees are closer together in the middle of the forest than they are at the outside, a collective reading. But when *all* is added, the sentence can only have the distributive reading. Similarly, (138) without *all* has a collective reading roughly synonymous with 'pass the proposal', but when *all* is added that reading disappears in favor of a distributive reading (which could be paraphrased as 'cast an individual vote in favor').

The significant fact here is that *be dense* is a stative predicate, and *vote for* is an

achievement predicate. So the fact that the collective reading vanishes when *all* is used is predicted by the proposal I have made here.

4.4.3.2 Some exceptions

4.4.3.2.1 Unexplained exceptions

There are some exceptions to Taub's generalization that are not predicted by my analysis. For some reason all of the exceptions that I can find are predicates having to do with ownership, including stative predicates like *own* and *have*, and achievement predicates like *inherit* and *buy* (actually it's not quite clear whether *buy* is an achievement or an accomplishment). These predicates are all ambiguous between collective and distributive readings, and unfortunately for me, the collective readings with *all/both* are not as bad as I would predict.

For example, (139) allows a collective reading; this reading is even more salient in (140).

(139) John and Mary both own a house
(140) John and Mary both own that house

I think it is a little harder to get a collective reading with *all* than it is with *both*, but it is clearly still possible (which is a notable contrast with the examples like *be dense* that we saw in the previous section).

(141) The students all own that house
(142) The grandchildren all inherited that house

I don't have an explanation for these exceptions. However, the fact that they all have to do with possession suggests that the place to begin looking for an explanation would be

to look more closely at possession. If the general approach to *all/both* and collectivity that I have proposed here is correct, then we would expect to find that there is more internal structure in the meaning of verbs of possession that would play a role in licensing *all*.

4.4.3.2.2 The exceptions that prove the rule

On the other hand, there are other exceptions to Taub's generalization that, on close inspection, actually provide evidence in favor of the proposal I have made here. These are cases where the predicates of Taub's generalization do accept modification by *all*; but in these cases the context plays a crucial role because it provides information that we are talking about subpluralities of pluralities. (Thanks to Veneeta Dayal, p.c. for pointing these out to me.)

For example, suppose we are teachers at Wading River Elementary School. The students are holding elections for class president, so each grade will elect its own president. In this context, it is possible to say (143).

(143) All the students elected a president

But what does it mean? It doesn't mean that each individual student elected a president. It means that the third graders elected a president, the fourth graders elected a president, and so on. This being the case, we can suppose that a D operator inserted on the VP will not be ruled out by economy considerations, because in this context the D operator actually does do some work for us. The sentence in (143) would be interpreted as in (144).

(144) $\exists e \forall x \exists e'[x \subseteq [\![\text{the.students}']\!] \ \& \ x \in [\![\text{Cov}_i]\!] \rightarrow \text{elect}(e') \ \& \ \text{Ro}(e',x) \ \& \ e' \subseteq e]$

Here the context provides us with cells in the cover that are equal to the students in a particular grade. Since distributivity is permitted here, we have a Cov variable so we

expect that modification by *all* should be possible.

Of course, since pragmatic weakening is also due to the presence of a D operator, we expect that (143) should license pragmatic weakening as well. It does, as (145) shows.

(145) The students elected a president, but not the fourth graders because nobody would run.

What this tells us is that Taub's generalization is really about distributivity; the fact that it applies to aktionsart classes is a consequence of the fact that activities and accomplishments, even when collective, actually do involve distributivity, while the collective states and achievements usually do not. But if we can make a collective state or achievement distributive with a rich enough context, we get the same behavior we get with any other distributive predicate.

4.4.3.3 Exception phrases

This explanation for Taub's generalization also explains why it extends to exception phrases.

Exception phrases are similar to *all* in that they are phrases that "do something" to quantification that is already present elsewhere in the sentence. In the analysis of von Fintel (1994 and see references cited there) except phrases are domain subtractors, that is, they subtract things from the domain of quantification of a quantifier that is present elsewhere in the sentence. For example, take the sentence in (146).

(146) Every girl went to the gym, except for Jackie

Here *except* subtracts Jackie from the set of girls, which is the domain of quantification of *every*. *All* of course has a similar function: as we said in Chapter 3, one way to think of the good fit requirement is that it makes the domain of quantification as

strong as it can be with respect to the subject DP. So what *all* and *except* have in common is that they are dependent on quantification that comes from elsewhere.

In the case of collective states and achievements, I have argued that there is no quantification elsewhere, because these predicates do not require the presence of distributivity in order to be interpreted with a plural argument. Since there is no distributivity, and hence no quantification, the ill-formedness of except clauses with these predicates is predicted:

(147) *The boys are a big group, except for Jason
(148) *The students elected a president, except for Mary

In these sentences *except*, like *all* above, has nothing to operate on. However, in the previous section I showed a class of examples that appear to be exceptions to Taub's generalization. I argued that these are cases where a D operator is licensed by a context in which groups of individuals are very salient (like the separation of schoolchildren by grade). If a D operator is licensed in these cases, then we expect that exception phrases should be possible. And in fact they are, as shown by (149). Recall again our context where the students are holding elections for class president by grade.

(149) The students elected a president, except for the fourth graders.

Thus the fact that exception phrases and *all* have the same distribution is predicted by this analysis.

4.5 Summary

In this chapter I have expanded the empirical coverage of the idea, introduced in Chapter 3, that *all* is a modifier that interacts with distributivity. We have seen that the two key ingredients of the hypothesis, the idea that *all* is not a quantifier and that *all* is

dependent on distributivity, together explain a wide range of facts about the distribution and interpretation of *all*.

We have explained the scopal differences betewen *all* and *every*. We have also explained why *all* is compatible with collectivizing adverbs (again, in contrast to *every*).

Finally, I have proposed that *all*, pragmatic weakening, and exception phrases are all sensitive to the presence of quantification in a sentence. My proposal about the structure of activities and accomplishments, along with the theory of the contextual component of distributivity, allows us to predict when a D operator will be licensed. Hence the distribution of *all* (especially with collective predicates), pragmatic weakening, and exception clauses is also predicted.

We will conclude this chapter with comparisons to two other recent proposals about the meaning of *all*.

4.6 Comparisons to other approaches

4.6.1 Winter

In work in progress, Winter (1998a, 1998b) makes a very different proposal about the meaning of *all*. He proposes that *all* is in fact a universal quantifier, with a core meaning that is the same as *every*. The difference between *all* and *every* is in the kind of constituents they select as their first argument: while *all* can take a plural argument, *every* selects a singular argument. His proposal (like the one I have made here) is part of a general analysis of plurality and quantification.

Winter proposes a basic distinction between "atom predicates" and "set predicates". In the nominal domain, the distinction corresponds to the morphological singular/plural

distinction. Singular common nouns denote atom predicates, and plural common nouns denote set predicates (that is, sets of sets; we have seen this idea in our discussion of Landman 1989 in Chapter 2).

In the verbal domain, Winter proposes the following diagnostic for distinguishing the two types of predicates: predicates for which *all* and *every* yield equivalent truth conditions are atom predicates; predicates for which *all* and *every* yield different truth conditions are set predicates.

The distinction can be exemplified with the following predicates.

(150) **Atom predicates:** *sleep, laugh, be hungry, be a big group, outnumber, vote*
(151) **Set predicates:** *meet, gather, be similar, carry the piano (together)*

There are two things about this classification that should immediately be pointed out. First, note that Winter treats some collective predicates, such as *be a big group* and *outnumber*, as atom predicates. Winter adopts the idea (from Landman 1989 and others, see Chapter 2) that pluralities can be mapped on to groups that behave like atoms. Secondly, note that these predicates fall under Taub's generalization: they are collective, stative predicates. Significantly, every collective predicate that Winter categorizes as an atom predicate is a state or achievement.

Without any further stipulations, the framework Winter proposes makes two correct predictions. If we take the basic meaning of quantifier/determiners to be relations between atom predicates, then we expect that *every* and *all* should have identical truth conditions with atom predicates, and we also expect the ill-formedness of *every* with set predicates. But the challenge that remains is to explain how *all* can combine with those collective predicates that it can combine with, which have been classified as set predicates. Winter

adopts a proposal from Scha (1981) for numerals and broadens its application to include *all*. Quantification over set predicates is interpreted as follows.

(152) In a quantificational structure *D* **A B** where **A** and **B** are set predicates, interpret *D* **A B** as *D* (∪**A**)(∪(**A**∩**B**)).

Note that due to the conservativity of natural language determiners, for atom predicates, *D* **A B** is always equivalent to *D* (**A**) (**A**∩**B**). The rule in (152) stipulates that we may use (**A**∩**B**) when computing the second argument of *D;* nevertheless, this rule can be viewed as making a rather elegant proposal about the similarities between singular (atomic) quantification and plural quantification.

The rule of set quantification in (152) effectively takes set predicates and "boils them down" into first-order sets to serve as arguments for the quantifier. (Because $\cup Y =_{df} \{x:\exists Z [x \in Z \ \& \ Z \in Y]\}$.) So a sentence like (153) will be interpreted as in (154)[9] (using rather informal notation).

(153) All (the) students gathered in the hallway

(154) all (∪(students')) (∪(students' ∩ gathered'))

This derives the right truth conditions for most cases (for more detailed discussion of some more exotic cases, see Winter 1998a & b). So the analysis explains how *all* can combine with some collective predicates, which as we know is something that must be acounted for by any theory of what *all* means.

The theory that Winter gives is rather simple and elegant. In addition, it has a

[9] Winter ignores the distinction between *students* and *the students*. Recall that we saw in Chapter 1 that *all NP* can in fact only be interpreted as a generic. However, it is not clear how the use of *all* with generics would figure in to Winter's proposal (or mine, for that matter).

slightly different empirical domain than the proposal I have made here: Winter is concerned with all sorts of quantifiers that take plural arguments, so in addition to the facts about *all* it also accounts for the interaction of numeral plural determiners (*exactly three, at least nine,* etc.) with set predicates. This is a phenomenon that I have not made any attempt to address here.

Despite the appeal of its simplicity, however, I see two problems with Winter's proposal. The first problem is the following. Because Winter uses the distribution of *all* as a diagnostic for his classification of predicates into set predicates and atom predicates, he has not given any explanation for why *all* has that distribution. While it does follow from his account of plural quantification that only set predicates should be felicitous with *all*, the account is somewhat unsatisfying because it was the distribution of *all* that motivated the distinction between set predicates and atom predicates in the first place. Without independent motivation for the distinction between set predicates and atom predicates, Winter has not explained why *all* cannot occur with (verbal) atom predicates.

In light of the proposal I have made here this is a drawback of Winter's proposal, because in fact there is independent evidence for the distinction between "atom predicates" and "set predicates", but Winter doesn't use it. The independent evidence, of course, is Taub's generalization. So while Winter's proposal essentially takes Taub's generalization as a starting point, I have tried here to give an explanation for Taub's generalization.

Winter's proposal also inherits an empirical problem of a sort that we have already seen, in Chapter 2. Winter adopts a groups approach to nonmaximality (that is, he treats nonmaximality as a kind of collectivity). So his account inherits the problems of that

approach; he predicts that (155) on a distributive reading (one-sandwich-per-boy) is synonymous with (156).

(155) The boys ate a sandwich
(156) All the boys ate a sandwich

Although Winter has given a worked-out and believable semantics for *all*, it nevertheless fails to predict the difference between (155) and (156), because he takes a groups approach to nonmaximality.

We might suppose, then, that to fix this Winter could choose to adopt the approach to nonmaximality I have proposed here, and combine it with the semantics for *all* that he proposes.

But in Winter's conceptual framework there is no connection between *all* and nonmaximality. So even if he were to adopt the approach to nonmaximality that I have proposed here, the fact that *all* and pragmatic weakening (not to mention *except* phrases) are possible with the same classes of predicates would have to be considered an accident.

4.6.2 Lasersohn

In a recent paper Peter Lasersohn (1998b) independently develops an approach to *all* which is much closer in spirit to the approach I have proposed in this thesis, although it is quite different in its execution.

Lasersohn, too, takes the problem I have called "nonmaximality" to be crucial to an understanding of the meaning of *all*; and his proposal about the meaning of *all* is quite similar to my proposal in that its proposed function is to prevent nonmaximality. But his account of nonmaximality is quite different from the one I have proposed here.

Lasersohn's account of the strengthening effect of *all* is part of a more general

proposal which includes other expressions like *exactly* and *perfectly*. His idea is that people often say things which are recognizably false but close enough to true that we are willing to let them count as true; in other words, that speakers grant one another a certain amount of "pragmatic slack". Expressions like *all*, *exactly*, and *perfectly*, then, are "slack regulators:" they signal to a listener that the speaker wants to limit the amount of leeway that would normally be expected in an interpretation of the utterance.

The formal system that Lasersohn proposes for representing this "slack" introduces and defines the idea of a "pragmatic halo" for the denotation of a natural language expression. The halo consists of things that are "sufficiently close" to the actual denotation of the expression.

This is perhaps best illustrated by an example. Lasersohn proposes that the difference between the sentences in (157)-(158) is due to the slack regulating effect of *exactly* on the temporal expression *three o'clock*.

(157) Polly arrived at three o'clock
(158) Polly arrived at exactly three o'clock

To capture this, we construct a halo of meanings associated with *three o'clock*. This will be a set of things of the same type as the denotation of *three o'clock*, hence it will be a set of times. This set may be ordered (totally or partially) in a way that reflects their closeness to three o'clock; so it will capture the fact that 2:58 is closer to 3:00 than is 2:42.

Lasersohn allows that any natural-language expression may in principle have a halo. These halos can compose to form more complex halos. To keep things simple, though, in our example we will assume that every word in (157) has a trivial halo except for the expression *three o'clock*. If we build the meaning of the sentence compositionally, then, it

too will have a halo. Its halo will consist of propositions of the form *Polly arrived at 2:59, Polly arrived at 2:58*, and so on.

So the sentence will be literally true if there is an event in which Mary arrived at precisely three o'clock. However, if we are allowing for some pragmatic slack, then the sentence will count as "true enough" if there is an event that makes one of the propositions in its halo true.

The function of *exactly* in (158) is essentially to "shrink" the halo of *three o'clock*. This will have the effect of shrinking the halo around the sentence, and hence the sentence will only be "true enough" if Polly's arrival is right at three o'clock (perhaps in this case we are willing to give a few milliseconds of slack).

All works in essentially the same way. Lasersohn supposes that the pragmatic halo of a definite plural like *the students* consists of sets that contain nearly the same members as ⟦the.students'⟧. For example, if ⟦the.students'⟧ = {a,b,c}, then its halo might be {{a,b}, {a,c},{b,c}}. So a sentence like *the students arrived* might count as true even if only Alan and Bill arrived, since the proposition that Alan and Bill arrived is in the halo. As a slack regulator, *all* would shrink the halo of *the students*, so the halo of the sentence *all the students arrived* would be smaller (at least with respect to *the students*) than the halo of *the students arrived*, and it would not count as true if only Alan and Bill arrived.

Lasersohn's proposal for the meaning of *all* is somewhat more general than the proposal I've made here, and interesting because it relates the meaning of *all* to other words like *exactly* and *perfectly*, whose similarity to *all* is perhaps otherwise not so obvious. He does not account for the limited distribution of *all* with collective predicates although he

does explicitly recognize that it is an issue; his putting that issue aside is perhaps justified since he is concerned with more than just plurality in his proposal.

However, one thing that I find intriguing about Lasersohn's proposal is that the idea of pragmatic slack works best for things that can (more or less naturally) be quantified. The distance between times, and the number of students, are two examples of this.

For a case like *perfectly spherical*, which Lasersohn discusses, the issue arises an be defined in terms of degrees of roundness. So in this case Lasersohn makes use of the scalar predicate *round* to introduce degrees on a scale, in order that we can again work with things that can be quantified.

It seems to me that the farther we get from things on which some kind of quantificational scale can be imposed, the less evidence we see that speakers use pragmatic slack. For example, what would be the halo of a predicate like *sneeze*, or *arrive*? Although for *sneeze* we might be able to imagine some notion of getting nearly all the way through a sneeze or something, it seems that this is the kind of thing for which people rarely employ pragmatic slack.

This suggests to me that quantification is lurking somewhere and that pragmatic slack is really related to domain of quantification effects, as I've proposed here for *all*.

Chapter 5

The Syntax of Floating Quantifiers

5 Introduction

Perhaps the most well-studied aspect of floating quantifiers is the syntactic question of how 'floating' comes about. The earliest answers to this question in the generative literature, including work by Postal (1974), Kayne (1975), and Maling (1976), was that a rule of "Quantifier Float" applied to this class of quantifiers and moved them rightward, away from the NP. Dowty and Brodie (1984) introduced the idea that floating quantifiers are not really moved quantifiers at all, but adverbial elements that are base-generated in their VP-adjoined positions. Baltin (1995) proposed that floating quantifiers are members of a small syntactic category he calls "preverbs".

In this chapter I will defend the idea that *all* and *both* are adverbs when they appear in floated position, and degree words/adjectives when they appear in prenominal position. Thus there is no operation of quantifier float. I will show that this analysis, when combined with the semantic proposals made in the previous chapters, accounts for the apparent "subject-orientedness" of floated quantifiers, for their syntactic distribution in the verbal extended projection, and for their syntactic distribution in the nominal extended projection.

5.1 The syntax of floating quantifiers

5.1.1 Floating as stranding

Perhaps the most influential proposal about floating quantifiers in the last decade

is Sportiche's (1988) proposal. Traditional approaches to quantifier float, such as those cited above, treated quantifier float as a transformation that moves a quantifier away from the subject DP. Sportiche turned this idea on its head: he argued that floating quantifiers appear to have moved rightward off of the DP because they have actually been stranded by the DP's having moved leftward.

This proposal has been especially appealing in light of the increasingly-accepted idea that subjects originate in VP-internal position (see, for example, Koopman and Sportiche 1985, 1991) because it gives a very natural account of how floated quantifiers can appear to the right of auxiliary verbs. It also offers an explanation for the (apparently) anaphoric link between the quantifier and the DP: the quantifier is sister to a trace of the DP, which has moved up to a higher position in the phrase structure.

Sportiche's proposal is based on an analysis of French floating quantifiers, but he argues that it extends as well to English. His proposal means that the structure of (1) is something like (2) (abstracting away from some different assumptions Sportiche makes about clausal structure which don't concern us here).

(1) All the boys should have eaten

(2) [[All [the boys]] [should [t have [t eaten]]]]

Since each trace in (2) represents a landing site for the subject as it moves through the clause, each one also represents a potential "stranding" site for *all*. Hence the possibilities in (3) are expected.

(3) (All) the boys should (all) have (all) eaten.

This proposal has since been adopted by many authors, for several languages,

including Déprez (1989), Giusti (1990) for German, Shlonsky (1991) for Hebrew, and others.

5.1.1.1 Problems for the stranding analysis

However, there are several problems for the analysis, many of which were noticed by Sportiche himself, and others which have been pointed out in the literature since. We will briefly review those here.

As Sportiche notes, one problem for the stranding hypothesis is that it leads us to expect that floated quantifiers should be able to appear after a passive verb, since the surface subject of a passive is supposed to have originated as the object. But this is not possible in English (and it is arguable whether it is even possible in French, as Sportiche claims (Viviane Déprez, p.c.))[1].

(4) *The boys were arrested all.

Sportiche offers an explanation for the ungrammaticality of (4) which in part involves a serious reanalysis of the passive, and it is not clear whether that reanalysis is independently motivated.

Another problem with the stranding hypothesis, pointed out by Pollock (1989, ftn. 5) and Doetjes (1992), is that it fails to capture a cross-linguistic generalization about adverbs and floating quantifiers. The position between the subject and the first auxiliary in English allows both adverbs and floating quantifiers, as shown in (5)-(6).

---

[1] Sportiche claims that a stranded *tous* ('all') after a passive verb is acceptable, and improves with an intonation break between the verb and the *tous*. This judgment is apparently subject to significant variation among speakers; Déprez says that it is impossible without a very prominent intonational break, suggesting that *tous* is a parenthetical in that case.

(5) The students all have left
(6) The students probably have left

However, neither adverbs nor floating quantifiers are permitted in that position in French (the examples are from Pollock).

(7) *Les enfants tous vont partir
the children all are going to leave

(8) *Les enfants beintôt vont partir
the children soon are going to leave

On the stranding hypothesis this apparent similarity between adverbs and floating quantifiers has to be considered an accident, since they do not appear in the same syntactic position (adverbs are, presumably, adjoined to a maximal projection, while stranded quantifiers are in the specifier of a functional projection).[2]

5.1.2 Arguments that all *is not an adverb*

Another kind of argument, however, that Sportiche makes in favor of the stranding hypothesis is to argue that floating quantifiers couldn't be adverbs. One part of his argument is to cite the fact that in some languages (Moore, see Tellier 1986; and Kilega, see Kinyalolo 1986) the distribution of floating quantifiers and adverbs do not overlap: in these languages floating quantifiers may occur between IP and VP but

[2] A recent analysis of adverbs by Cinque (to appear), in which adverbs occupy specifier positions, might at first appear to resolve this problem. Cinque proposes that adverbs have a close relationship with certain kinds of functional heads, and argues that rather than being freely adjoined, adverbs must occupy the spec of a particular functional head. (To give just one example, he supposes that modal adverbs such as *probably* and *possibly* occupy the specifier of a modal epistemic projection, whose head can host modal auxiliaries like *must* and *may*.)

However, in the same work, he treats floated quantifiers as occupying the specifier of "nominal-related" functional heads (presumably agreement/clitic heads). Since adverbs do not occupy the same specifier positions that DPs move through, the correlation between floating quantifier distribution and adverb distribution noted by Pollock still fails to be accounted for.

adverbs may not.

Secondly, he argues that floating quantifiers could not be adverbs because quantifiers are of a different semantic type than adverbs. But there are two problems with this argument. First, it is not true that adverbs cannot be quantifiers; adverbs of quantification like *always* and *usually* have since Lewis (1975) been analysed as quantifiers over situations or events. Secondly, we have argued extensively in Chapters 2 and 3 that *all* is not a quantifier anyway, at least not a determiner-quantifier.

Finally, Sportiche argues specifically that *all* could not be an adverb because it doesn't match the distribution of other subject-oriented adverbs such as *willingly*. In making this argument, Sportiche hypothesizes that adverbs should obey a generalization he calls the "Adjunct Projection Principle" (p.429) which says that modifiers must be adjacent to (the head or XP of) the thing they modify. For subject-oriented adverbs, this means that they must be "adjacent" (in some way; he just gives a sketch of the proposal) to both the subject and the propositional content of the clause. If floated quantifiers are subject-oriented adverbs then we expect some difficulty to arise in a sentence with both a subject-oriented adverb and a floating quantifier, since they can't both be adjacent to both constituents. However, it is quite possible to combine them in one sentence, as shown in (9).

(9) The girls all willingly left

From this, Sportiche concludes that floating quantifiers cannot be subject-oriented adverbs.

5.1.2.1 Déprez's arguments

In later work Déprez (1989) augmented Sportiche's arguments with the following examples, from McConnell-Ginet (1982). McConnell-Ginet observes that the sentence in (10) is ambiguous.

(10) The children have been willingly taught by their parents

The subject-oriented adverb *willingly* in (10) can modify either the syntactic subject *(the children)* or the underlying agent (*their parents*). (Note also that this gives us some reason to regard the Adjunct Projection Principle, at least in the use it is put to by Sportiche, with suspicion.) Déprez points out that floated quantifiers behave differently from subject-oriented adverbs in this environment. The ambiguity in (10) is not possible with a floated quantifier, as (11) shows.

(11) The children have all been taught by their parents

All in (11) can only be interpreted as being related to the DP *the children*.

Déprez makes a similar point with implied subjects and subject-oriented adverbs; although subject-oriented adverbs can modify implied subjects, this is impossible for floating quantifiers. For floating quantifiers, the antecedent must always be overt.

(12) This book$_j$ has been deliberately$_i$ damaged
(13) *This book$_j$ has all$_i$ been damaged

Déprez's arguments are, I think, more convincing than Sportiche's. But at this point, the conclusion that we are apparently forced to is only that *all* is not a subject-oriented adverb.

Both Sportiche and Déprez assume that if *all* were an adverb, it would be a subject-oriented adverb. This move is quite plausible, since it is clear that there must be

some syntactic relationship between *all* and the subject. However, it is not a necessary assumption. In section 5.1.5 I will propose an account of the different kind of "orientation" shown by *all* and other adverbs (such as *together*) that will capture the fact that it is syntactically linked to the subject (and that it is not a subject-oriented adverb).

But before I provide an account of how floated *all* becomes "oriented" towards the subject, I first want to argue that there are good reasons for supposing that floated *all* is an adverb, and in fact a kind of speaker-oriented adverb. The evidence will come from the similarity in distribution between *all* and the speaker-oriented adverbs.

5.1.3 Types of Adverbs

In his discussion of the syntax and semantics of adverbs, Jackendoff (1972) divides adverbs into three semantic classes: speaker-oriented adverbs such as *probably, allegedly, thankfully*; subject-oriented adverbs such as *willingly, intentionally, carefully* (on one reading); and manner adverbs such as *quickly, carelessly* (on one reading) and *quietly*. Jackendoff's classification has since then been articulated into more subtypes (most notably by Bellert 1977), but the basic categories he proposed have not been changed. For now we will keep things simpler by considering only these three classes.

Syntactically, Jackendoff argues, the speaker-oriented and subject-oriented adverbs behave as a class and what differences there are in their distribution are due to meaning. He shows that speaker- and subject-oriented adverbs are S-attached adverbs, and manner adverbs attach to VP, with the following data.

Jackendoff claims that S-attached adverbs, and not VP-attached adverbs, can occur before a single auxiliary or modal, as in (14). (The discussion through example

(17) is adapted from Jackendoff 1972, pp 75-76.) However, I find the subject-oriented adverbs to be somewhat odd in this position (he uses only the speaker-oriented adverb *probably* in his examples).

(14) The kids probably;??willingly/*completely have read the book
will lose their minds

After a single aux or modal, adverbs from either syntactic class can occur.

(15) The kids have probably;willingly/completely read the book
were under water
will lose their minds

Between two auxiliaries, VP adverbs are somewhat odd but S adverbs are perfect.

(16) The boys have probably;willingly/*completely been under water
will be ruined by the tornado
are being uncooperative

After two auxiliaries, Jackendoff says, VP adverbs are much better than S adverbs.

(17) The boys are being *probably;*willingly/completely uncooperative
will be ruined by the tornado
will have read the book

Turning to floating quantifiers, we find that their distribution closely matches the distribution of S adverbs.

(18) (from (14)) The boys all have read the book
were under water
will lose their minds

(19) (from (15)) The boys have all read the book
were under water
will lose their minds

(20) (from (16)) the boys have all been under water
will be ruined by the tornado
are being uncooperative

This suggests that *all* must be a speaker-oriented or a subject-oriented adverb. But we have already seen good evidence that *all* is not a subject-oriented adverb. In the next section we will see that, consistent with the claims of Sportiche and Déprez, the distribution of *all* differs from the distribution of subject-oriented adverbs, suggesting that it is not a subject oriented adverb. On the other hand, its distribution is similar to that of speaker-oriented adverbs.

5.1.3.1 Adverbs after two auxiliary elements

Jackendoff claims that S-adverbs are not found after two auxiliaries, but I think this is not quite correct. First, it appears that the S-adverbs do not behave as a uniform class in this position: the distribution of the speaker-oriented adverbs is different from the distribution of the subject-oriented adverbs.

Speaker-oriented adverbs are in fact possible after two auxiliaries, but only if the second aux is *have*. This pattern is obscured in Jackendoff's data by a kind of semantic mismatch between certain modals and certain adverbs, and can be improved if we combine them in the right way. Observe the variation in (21)-(22).

(21) The children will have supposedly read the book
The boys must have certainly gone home
The children would have probably read the book
The girls should have allegedly finished by now.

(22) *The children will be certainly reading the book
*The boys must be certainly going home
*the children might be probably reading the book
??The girls should be allegedly finishing by now

Overall, the sentences in (22), where the second aux is *be*, are worse than the sentences in (21), where the second aux is *have*.

In contrast, it appears that subject-oriented adverbs are possible after two auxes, regardless of whether the second aux is *have* or *be*. The data is shown in (23)-(24). (23)

The children will be willingly reading the book
The boys must be intentionally failing the exam
The children might be intentionally leaving by now

(24) The children will have willingly read the book
The boys must have intentionally failed the exam
The children might have willingly left by now

So in fact it appears that the second auxiliary makes a difference: speaker-oriented adverbs are permitted after two auxiliaries if the second aux is *have*, but not if the second aux is *be*. Subject-oriented adverbs are permitted with either auxiliary as the second one. Turning to *all*, the data in (25)-(26) show that it behaves more like a speaker-oriented adverb than like a subject-oriented adverb. Note that (27) shows that if *be* is followed by an adjective that allows *all* as a modifier then the sentence can only be interpreted with *all* modifying the adjectival predicate.

(25) The children will have all read the book
The boys must have all failed the exam
The children might have all left by now

(26) *The girls have been all sleeping
*The boys must be all going home
*The children might be all reading the book.

(27) The boys have been all dirty
The girls might be all wet

These data suggest, first, that Sportiche and Déprez were correct to argue that *all* is not a subject-oriented adverb. Secondly, the data suggest that *all* could be a speaker-oriented adverb, since it has the same distribution.

However, if we are going to claim that *all* is a kind of speaker-oriented adverb,

we must provide evidence that this is a plausible way of looking at things. In the sections to follow, I will make two arguments for this hypothesis: a conceptual argument, and an empirical argument.

5.1.4 All *as a speaker adverb*

First, the conceptual argument. The speaker-oriented adverbs have, since Jackendoff's work, been further articulated into several classes (Bellert, 1977, Ernst 1998). The subclasses include the following (adapted from Ernst 1998).

(28) a. speech-act: *frankly, honestly, simply, briefly*
b. modal/ evidential: *probably, surely, possibly/ clearly, apparently, obviously*
c. evaluative: *luckily, amazingly, oddly, curiously, ideally*

We will address shortly the question of where *all* belongs in this classification. What we are interested in now is what these classes have in common, that merits their being considered subclasses of a single class. Although I know of no formal definition that tells us what should count as a speaker-oriented adverb, most of the authors who discuss these adverbs seem to rely on an intuitive characterization like the following: the speaker-oriented adverbs are adverbs that give a speaker a way to modulate an assertion.

I will give just a brief discussion of this idea here. Rather than simply asserting that *p* or that *not p*, using a speaker-oriented adverb a speaker may assert *probably p*, *luckily p*, or *obviously p*. In the analysis of Ernst (1998) (simplifying somewhat) the modal/evidential and the evaluative adverbs are treated as adjectival predicates that map the proposition *p* (or the event that it can be taken to denote) on to a scale of "probable" or "lucky" events. By using the adverb *luckily*, then, a speaker not only asserts that *p* but also asserts that the occurrence of an event described by *p* is relatively high on the scale

of lucky things.

The “maximizing” effect of *all*, which I have accounted for in the proposals of Chapters 2 and 3, can be viewed in a similar way: it says that a proposition *p* is true in the strongest way possible (with respect to distributivity). It, too, can be viewed as mapping a proposition *p* on to a scale; a speaker who uses *all* asserts that *p* and furthermore that *p* is true in the strongest possible way (with respect to distributivity). In this way it is a "proposition modulator" of a sort similar to other speaker-oriented adverbs.

For example, take the sentence *the girls all left*. We can take *p* to be the proposition that [the girls left]. Then given that *all* requires a good-fitting cover so that quantification by the D operator is maximized with respect to the girls, we might say that what *all* does is to express *maximally p*. That is, *p* is true in the strongest way it can be true.

Note that this picture is made possible by the proposal that *all* is <u>not</u> a determiner-quantifier. If it were, then there would be no way to separate the proposition *p* from the contribution of *all*; without the quantifier there simply wouldn't be a proposition. But the proposal I have made here gives us a way to see how the contribution of *all* can be separated from the propositional content of the rest of the sentence, and hence makes the idea that *all* might be an adverb much more plausible.

I should make one more note here. Later in this chapter I will argue for an analysis of adverb orientation in which the term "orientation" means something different from what it means here as part of the term *speaker-oriented*. So to avoid confusion

from now on I will refer to this class of adverbs simply as speaker adverbs.

5.1.4.1 Types of speaker adverbs

As I mentioned above, the class of speaker adverbs has been further subdivided since Jackendoff's work. Since I am arguing that *all* is a kind of speaker adverb, I must address the issue of what subclass of speaker adverbs the floating quantifiers belong to.

The work of Bellert (1977) and Ernst (1998) suggests that the speaker adverbs can be divided into three subclasses, shown below (the category names are taken from Ernst 1998).

(29) a. speech-act: *frankly, honestly, simply, briefly*
b. evaluative: *luckily, amazingly, oddly, curiously, ideally*
c. modal/evidential: *probably, maybe, surely,/clearly, apparently, obviously*[3]

The criteria for dividing the speaker adverbs into the categories a-c is, according Bellert (1977), mostly distributional and semantic. By her criteria, it appears that *all* does not fall into any of these categories.

Bellert identifies the speech act adverbs (she calls them "pragmatic adverbs") by the fact that they can always be paraphrased by a sentence with the word *speaking*, as in (30)-(31)

(30) I honestly did it myself
(31) Speaking honestly, I did it myself.

This kind of paraphrase is not possible with a floating quantifier: **Speaking all, the girls left.* So we can conclude that floating quantifiers are not speech-act adverbs.

---

[3] Actually, Ernst divides this class into two classes, modal and evidential (and the adjectives belonging to the respective classes is indicated by the /). However he doesn't give any evidence for the difference. Below we will see that they behave, at least with respect to the phenomena that we are interested in, like a single class, and Bellert treats them as belonging to a single class, so I have put them together in class c.

Bellert treats the modal and evidential adverbs as belonging to a single class which can be characterized by, among other things, the fact that they cannot be morphologically negated, as shown in (32), and the fact that they cannot occur in questions, as shown in (33).

(32) *Impossibly, John has arrived
(33) *Has John evidently arrived?

The evidence regarding *all* is not clear with respect to these tests. Although *all* cannot be morphologically negated (like the modal and evidential adverbs), it appears that it can be periphrasically negated.

(34) The girls didn't all leave
(35) Not all the girls left

On the other hand it is clear that *all* is perfectly compatible with questions, unlike the modals and evidentials.

(36) Have all the girls arrived?

I tentatively conclude from these facts that *all* is not a modal/evidential adverb.

The evaluative adverbs are, like the modal adverbs, incompatible with questions, as shown in (37).

(37) *Has John surprisingly arrived?

But they are distinguished from the modal adverbs because they are factive predicates: if (38) is true then (39) is necessarily true also.

(38) John has fortunately arrived
(39) John has arrived

The fact that *all* is compatible with questions, as we saw above, suggests that *all* is not an evaluative adverb, either.

So it appears that if I am to propose that *all* is a speaker adverb, then I will have to propose that it makes up a new class of speaker adverb.

As far as I know, there is no impediment to this from any theory that provides the essential or defining properties of speaker adverbs. As I discussed earlier, the justification given in the literature for classifying these adverbs together comes from intuitive characterizations about what it could mean to be "speaker-oriented". I argued there that these same kinds of characterizations plausibly apply to *all*. So I will assume that floating quantifiers comprise another subclass of speaker adverbs. In the next section we will see evidence in favor of this claim.

5.1.4.2 The scope of adverbs

The empirical evidence that *all* is a speaker adverb will come from *all*'s interaction with other adverbs. It is well-known that when several adverbs appear in a sentence, there are restrictions on the order in which they may appear. (See Jackendoff 1972, Ernst 1998, Cinque to appear, and others.) I will show that *all*, too, is subject to such restrictions, and that the restrictions are predicted by the hypothesis that *all* is a kind of speaker adverb.

There are various ways of accounting for the kinds of interactions I will describe below (see Ernst 1998, Cinque to appear). I will describe the interactions in terms of "scope," because it is a convenient term and because I find Ernst's (1998) theory of the scope of adverbs to be persuasive. However, I will not give a theory of what it means for an adverb to have scope; our goal here is not to give a theory of adverbs but to give evidence that *all* behaves like an adverb.

Adverbs can be organized into a hierarchy that captures their distributional properties. The hierarchy is given below, in (40) (adapted from Ernst 1998). The first three subclasses (a-c) comprise the speaker adverbs. Categories d-e are the subject-oriented adverbs, and f is the manner adverbs.

(40) a. speech-act: *frankly, honestly, simply, briefly*
b. evaluative: *luckily, amazingly, oddly, curiously, ideally*
c. modal/evidential: *probably, maybe, surely,/clearly, apparently, obviously*

d. agent-oriented: *politely, stupidly, cleverly, graciously, rudely*
e. mental-attitude: *happily, willingly, sadly, anxiously*

f. manner: *loudly, tightly, jerkily, blindingly*

In a string of two or more adverbs, adverbs of class *a* must have scope over (ie, appear to the left of) adverbs of classes *b-f*; adverbs of class *b* must have scope over the lower classes, and so on, as schematized in (41).

(41) a >> b >> c >> d,e >> f

This means that, for example, a manner adverb cannot appear in a position dominating a modal adverb, because the class of modal adverbs is higher in the hierarchy. Thus we expect that a sentence like (42) should be ungrammatical, which is indeed true.

(42) *The girls loudly probably talked about the movie

Sometimes when an adverb from a lower class dominates an adverb from a higher class, we get an effect that is different from ungrammaticality. We see in these cases a kind of "forced meaning" effect in which one of the adverbs takes on a different kind of meaning, one that is compatible with its syntactic position and the class of adverbs that may appear there. For example, *honestly* is a speech-act adverb, and so the sentence in

(43) is anomalous.

(43) #The police probably honestly left

However, to the extent that we can undestand this sentence, it seems to demand that we somehow imagine a manner interpretation for *honestly* (some kind of honest leaving?). So *honestly* is "forced" into a manner interpretation (or possibly a subject-oriented interpretation) because it has scope under *probably*.

I propose that the class of floated quantifiers occupies the same place in the hierarchy as the modal adverbs (which also, of course, belong to the larger class of speaker adverbs). In the next sections we will see that this predicts the distribution of *all* with respect to other adverbs and also the fact that it, too, can be involved in "forced meaning" interpretations.

5.1.4.3 Evidence from interaction with other adverbs

5.1.4.3.1 All *and lower-type adverbs*

If floating quantifiers belong in the same place in the adverb hierarchy as the modal adverbs, this predicts quite straightforwardly that they should show the same distribution. We expect, for example, that *all* cannot occur in the scope of manner adverbs and subject oriented adverbs.

Examples (44)-(45), modified from Jackendoff (1972), show that neither the subject-oriented adverb *bravely* nor the manner adverb *skillfully* can dominate the modal adverb *probably*. This is predicted by their relative places in the heirarchy. On the other hand, the reverse order is fine, as shown in (46)-(47).

(44) *Max bravely probably climbed the wall
(45) *Max skillfully probably climbed the wall

(46) Max probably bravely climbed the wall
(47) Max probably skillfully climbed the wall

All shows similar behavior in this respect, as shown in (48)-(51).

(48) ??The boys bravely all climbed the wall
(49) ??The boys skillfully all climbed the wall

(50) The boys all bravely climbed the wall
(51) The boys all skillfully climbed the wall

Although (48)-(49) are not quite as bad as (44)-(45), it is quite clear that something has changed. In addition to the fact that (48)-(49) are both more awkward than their counterparts in (50)-(51), we also see a meaning change in evidence, which appears to be similar to the type that we saw above in our discussion of "forced meaning." It appears that some kind of higher-type meaning is being forced upon *skillfully* and *bravely* which they are not accepting too gracefully.

For example, the adverb *skillfully* in (49) clearly cannot be interpreted as a manner adverb. To the extent that it can be understood, it means something like, it was skillful of the boys to all climb the wall. That is, they showed some kind of group organizational skills, rather than climbing in a skillful manner. Here's a context where this makes a difference. Suppose that the boys are a bunch of boy scouts engaged in a group-building exercise, which requires that they somehow manage to get every member of the group over a 10-foot wall. None of them has any climbing skills. However, through teamwork, they manage to devise a system whereby everyone helps everyone else to hoist his carcass over the wall, no matter how ungracefully (and some of them do go over rather ungracefully). In this situation, (49) is true (to the extent that it is grammatical) and (51) is false. It appears that the reading of (49) that is marginally

available is a type of agent-oriented reading.

Note also that the problem with (49) is not simply the difficulty of attributing skillfulness to a group agent (that is, to the boys as a whole) (as suggested to me by Tom Ernst, p.c.). The "group skillfulness" reading is available for the same sentence without *all*, as in (52).

(52) The boys skillfully climbed the wall

Although the "group skillfulness" reading is not the most salient reading out of the blue, it is perfectly sensible in the context described above for (49). So the idea of a 'group skillfulness' reading cannot be blamed for the awkwardness of (49). However, (49) is predicted by the hypothesis that *all* is an adverb that belongs in the hierarchy with the class c adverbs. It is not group skillfulness that is the problem here but the possiblity of treating *skillful* as a higher-type adverb (or *all* as a lower-type adverb).

5.1.4.3.2 All *and modal/evidential adverbs*

Another prediction we make with this hypothesis is that *all* and the modals should occur in any order. This is true also, as shown in (53)-(54).

(53) The boys probably all went home
(54) The boys all probably went home

The evidentials show exactly the same behavior.

(55) The girls all obviously left
(56) The girls obviously all left

This is predicted on the hypothesis that *all* is a speaker adverb that has the same

scopal properties as the class *c* adverbs of the hierarchy given in (41). [4]

5.1.4.3.3. All *and speech act, evaluative adverbs*

We also predict that *all* should not be able to occur with scope over speech act adverbs and evaluative adverbs because these are higher in the scopal hierarchy. This is true for the speech-act adverbs.

(57) The girls honestly all left
(58) *The girls all honestly left

Note that (58) is marginally permitted if we can come up with some way to have *honestly* have a manner interpretation, but this is exactly what we would expect.

On the other hand, it's not so clear whether this hypothesis makes correct predictions for the evaluative adverbs. The evaluatives are are higher in the heirarchy, so we expect that *all* should not be able to occur to the left of them. However, it appears that *all* may in fact dominate at least some evaluative adverbs (although there seems to be some variation here).

(59) *The girls all amazingly quit smoking
(60) The girls have all unfortunately left.

It's not clear to me why this should be possible, but I will leave this issue aside for now.

Overall, we have seen that the hypothesis that *all* is a speaker adverb of the same

---

[4] Roger Schwarzschild (p.c.) points out to me two other contexts where *all* and the modal adverbs show the same distribution, which lends some support to the hypothesis that they are similar.

(i) The boys left, probably/apparently/all in different cars
(ii) The boys left, *frankly/*amazingly/*carefully/*quickly in different cars
(iii) The documents, probably/apparently/all written in Greek, remained elusive
(iv) The documents, *frankly/*amazingly/*carefully/*quickly written in Greek, remained elusive

hierarchical status as the modal adverbs correctly predicts a wide range of its distribution with respect to other adverbs. We correctly predict that it should should be in free variation with modal and evidential adverbs, and be unable to take scope over speech-act (evaluative) adverbs.

We also predict the fact that *all* must take scope higher than manner and subject-oriented adverbs; and the kind of semantic anomaly that we see when the adverbs occur in the opposite scopal order is just the kind of semantic anomaly that we see generally when lower adverbs take scope over higher adverbs. This last fact is perhaps especially important because it is predicted by the hypothesis that *all* is a type of speaker adverb but it is not at all clear whether we would expect this kind of interaction if *all* were a quantifier.

Let us summarize what we have found. Earlier I proposed, on the basis of distributional evidence of the sort analyzed by Jackendoff (1972), that *all* is a type of speaker adverb. In this section we saw that this hypothesis holds up to further scrutiny when we look more closely at the interaction of adverbs with one another, hence we have provided additional evidence that *all* is a kind of speaker adverb.

5.1.5 Adverb orientedness

At this point we have hypothesized that floated *all* is a type of speaker adverb. But by itself that doesn't explain why the good fit requirement of floated *all*, in English, is always imposed on the subject. I will argue here that this syntactic fact is explained by an account of the "orientedness" of floating quantifiers that is related to, but not the same as, accounts that have been given for the subject-oriented adverbs (McConnell-Ginet

1982) and for another kind of "adverb orientedness" proposed by Schwarzschild (1994c) to account for some facts about *together*. We will look briefly at each of these in turn and then I will give a proposal that accounts for the fact that floated *all* must always be construed with the subject. (I thank Roger Schwarzschild for comments and suggestions that led me to pursue this approach.)

5.1.5.1 McConnell-Ginet's analysis of subject-orientedness

McConnell-Ginet is concerned with the fact, which we have already seen, that adverbs that are subject-oriented in an active sentence are ambiguous in the passive, between an "orientation" toward the surface subject and an orientation towards the agent (the underlying subject). This is shown in (61)-(62).

(61) Joan willingly instructed Mary *(Joan willing/ *Mary willing)*
(62) Mary was willingly instructed by Joan (*Joan willing/Mary willing)*

The first point that should be noted about this phenomenon is that the position of the adverb in the passive makes a difference: if the adverb precedes the auxiliary only (surface) subject orientation is possible. In other positions either kind of orientation is possible.

(63) Willingly, Mary was instructed by Joan (**Joan willing/Mary willing)*
(64) Mary willingly was instructed by Joan (**Joan willing/Mary willing)*
(65) Mary was instructed by Joan willingly (*Joan willing/Mary willing*)

One difficulty in giving an analysis of this phenomenon, as McConnell-Ginet's discussion makes clear, is in allowing *willingly* to orient towards the agent in the passive while also ruling out the possibility of allowing *willingly* to orient towards the object of the active sentence.

McConnell-Ginet proposes a semantics of subject-oriented adverbs (and most

other types of adverbs) in which they are not just predicate modifiers, but actually interact with the meaning of a verb to create a new meaning. She adopts the term Ad-Verb to reflect this idea. An Ad-Verb is an operator that applies to a verb to give it an extra argument place for an adverbial modifier (much like the argument place found in verbs that strictly subcategorize for adverbs, like *behave*), and then the Ad-Verb itself also fills that argument place.

Syntactically, this Ad-Verb may apply both to main verbs, like *instruct*, and to auxiliary verbs, like *was* in the passive. Since a passive sentence contains two possible verbs for *willingly* to attach to, and an active sentence contains only one, this is what accounts for the ambiguity of "orientation" of this kind of adverb in the passive but not in the active.

On this hypothesis, subject orientation in an active sentence is simply a by-product of what it means to "willingly instruct". *Willingly* has applied to the verb *instruct* to yield a different kind of instructing — willing instructing. For McConnell-Ginet, there is no sense in which the Ad-Verb *willingly* is "looking" for the subject. The fact that we understand the subject to be willing comes from the fact that it is predicated of the complex verb *willingly instruct*.

In the passive, surface subject orientation is due to the fact that the Ad-Verb attaches to *was* to yield something meaning roughly, *was-willingly*. The agent-oriented reading is due to the fact that the Ad-Verb applies to *instruct*, which, as we saw above, yields a constituent meaning something like *instruct-willingly*.

Since it is the attachment of *willingly* to *was* that derives the "orientation" toward

the surface subject, this also explains why the "agent-oriented" reading of the Ad-Verb in the passive is not available when the Ad-Verb is unambiguously adjoined to the S node. So on the analysis proposed by McConnell-Ginet, the phenomenon of subject orientation of adverbs has (at least) the following two properties. The fact that there is any "orientation" possible at all is attributed simply to the meaning of the particular Ad-Verb in question. Secondly, the direction of the orientation is determined in large part by the syntactic position of the Ad-Verb.

5.1.5.2 Another kind of adverb orientedness

In more recent work Schwarzschild (1992c) proposed a distinction between traditional subject-orientation of the sort described by McConnell-Ginet, and another sort of orientation that he observes. This kind of orientation is exhibited by the adverbs *together* and *individually*.

The first feature of this kind of adverb orientation is that it can be toward the subject or the object in active sentences.

(66) Alice and Max went shopping together
(67) Polly cooked the beans and the rice together

In the right kind of syntactic context, *together*, like *willingly*, can be ambiguous. Together can apparently be oriented toward the subject or the object.

(68) Alice and Max swallowed the pills and the alcohol together

This sentence can be true if Alice and Max, together, swallowed the pills on one occasion and the alcohol on another (ie *together*(A & M) not *together* (p & a)). On the other hand, it can also be true if Alice swallowed the pills and the alcohol together in her apartment on Spruce Street on Tuesday, and Max did the same on Easton Avenue on

Wednesday.

But there is one clear difference between the kind of orientation exhibited by *together* and the kind of orientation exhibited by *willingly*. While *willingly* can be oriented towards an implicit subject, *together* never can; it requires an overt argument.

(69) *The house was together destroyed
(70) The house was deliberately destroyed

Of course, this kind of contrast should look familiar. As part of Déprez's (1989) arguments that *all* is not a subject-oriented adverb, we saw that *all* also cannot be oriented towards an implicit subject. So our goal here will be to use Schwarzschild's ideas about orientation to point the way toward an account of *all*'s orientation.

Schwarzschild classified adverbs of the *together* type as "structurally-oriented" and adverbs of the *deliberately* type as "semantically-oriented". In light of our discussion of McConnell-Ginet's analysis of Ad-Verb orientation, in which structure plays a role, this terminology is perhaps not so felicitous. So I will continue to use "subject-oriented" in the traditional way, to refer to adverbs of the *deliberately* type, and refer to adverbs of the *together* type as "argument-oriented", for reasons that should become clear shortly.[5]

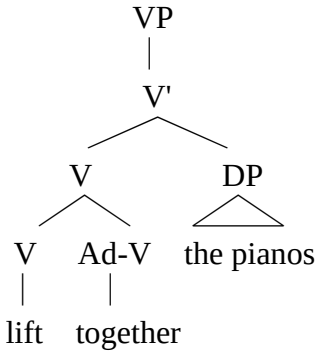
The analysis proposed by Schwarzschild is similar to McConnell-Ginet's analysis with respect to the two features we are most interested in: first, which argument it orients toward is determined by its position in the syntax. Secondly, the fact that it shows

[5] Lasersohn (1990,1995) also discusses the argument orientation of *together* and proposes an account of it. Schwarzschild argues against Lasersohn's approach, but we will not look at the details of the debate here because it is orthogonal to our concerns.

orientation is due to its meaning. Specifically, Schwarzschild proposes that adverbial *together* takes two arguments, a DP argument and a VP argument.[6]

Here is a sketch of how this works. Like McConnell-Ginet, Schwarzschild supposes that *together* can combine directly with a verb, as in (71). (Although note that for Schwarzschild *together* does not "operate" on the verb in the manner proposed by McConnell-Ginet.)

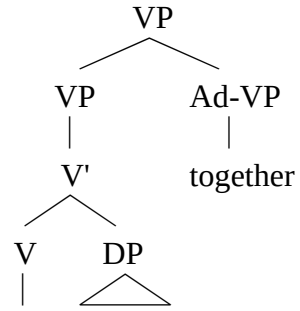
(71)



He proposes that the operation of right-wrap (see Bach 1984) then moves the object next to the verb to derive the surface word order.

Schwarzschild also proposes that *together* can adjoin to a (transitive) VP, as shown in (72).

(72)



[6] This idea is due to Lasersohn, who makes a similar proposal for *together* in his 1990 book.

lift the pianos

Given these two possible insertion sites, Schwarzschild defines *together* so that its "subject" (ie its DP argument) is the DP that combines with *together*'s V or VP argument. In other words, when *together* adjoins to a V its subject will be the verb's object, and hence *together* will be "oriented" toward the object. When *together* adjoins to VP, its subject will be the argument of the VP, and hence it will be "oriented" toward the subject.

If we assume that adjunction to V is always accompanied by right-wrap, then this hypothesis correctly predicts that the only position that licenses object-oriented *together* is the position to the right of the object. This is shown in (73)-(74).

(73) *Polly together cooked the beans and the rice
(74) [Alice and Max]$_i$ have together$_{i/*j}$ swallowed [the pills and the alcohol]$_j$

It should be pointed out that Schwarzschild leaves some facts about *together*, particularly adnominal *together*, unexplained. But this will not concern us because we are not interested here in giving an analysis of *together*; what we want to do is see how Schwarzschild's analysis points the way toward an analysis of *all*.[7] For this reason too I will not go into the formal details of the rules Schwarzschild gives combining *together* with its arguments.

What we have learned from Schwarzschild's and McConnell-Ginet's analyses are the types of adverb-orientation to be found in English (at least subject-orientation and argument-orientation) and that the direction of orientation can be accounted for by the

[7] Although it would perhaps be fruitful at a future date to explore the similarities between *together* and *all* more thoroughly.

right kind of syntactic-semantic analysis.

5.1.5.3 The orientedness of all

In this section I will argue that *all* exhibits the kind of "argument orientation" exhibited by *together*. The main difference between *all* and *together* is that *all* cannot adjoin to the verb. It can only be adjoined at a higher point in the sentence structure: at least outside of VP (below I will argue that there are independent reasons for assuming this difference). This will explain why adverb *all* is never object-oriented, at least in English.

I have proposed in this dissertation that the contribution of *all* to meaning is not a truth-conditional one; rather, *all* interacts with the process of assigning a value to the domain variable *Cov*. This will make spelling out the derivations of sentences with *all* a bit more complicated. In what follows I will adopt the convention of placing the good fit requirement in between the symbols ‖ ‖ to orthographically mark that good fit is not evaluable as part of the truth conditions of the sentence, but is an operator that interacts with the context to limit the possible choices of *Cov*. I call this aspect of meaning the "domain-adjusting meaning".[8] The idea is that the domain-adjusting meaning of a

[8] An obvious question here is whether the idea of "domain adjusting meaning" bears any relationship to presupposition, another major source of non-truth-conditional meaning. In a dynamic approach to presupposition (Stalnaker 1973, Heim 1983), the "pre" suffix is taken to heart: we understand a presupposition to be a kind of precondition for a felicitous utterance. I don't have any strong intuition that *all* sets up that kind of precondition. On the other hand, in the work of Beaver (1996), presupposition is treated as a device for "context selection"; that is, just as speakers entertain a set of possible worlds as candidates for the actual one, speakers also entertain a set of possible contexts (a set of sets of possible worlds) as candidates for the "actual" context. On this account, presuppositions are used to rule out certain kinds of contexts, that is, as clues to the kind of context that a speaker has in mind. If this account is correct it seems to be much closer to what I have in mind for *all*. Nevertheless, I refrain here from calling the good fit requirement a presupposition because it would raise numerous new questions. For one thing, as

constituent (if it has one) is separate from its truth-conditional meaning and is derived alongside the truth conditions. Here I just sketch an account of how this works, but in light of the fact that other operators can arguably affect the resource domain variable of a quantifier (see Chapter 3, section 3.3.1), the system proposed here might plausibly have applications beyond the syntax/semantics of *all* and *both*.

I propose the following translation rule for *all*.

(75) ***all*-translation rule:**
all has no ordinary translation, and a domain-adjusting meaning of
$^{\|}\lambda x[gf(Cov)(x)]^{\|}$

The fact that *all* behaves like an adverb comes from the fact that its meaning is essentially conjoined to the meaning of the VP, and then it takes as its argument the same object that the VP takes as its argument. This is guaranteed by the following rules. The first rule combinines *all* and a VP, and the second combines the resulting VP with an NP argument. (76) and (77) are "mode of composition" rules, similar in that respect to the rule for Lasersohn's generalized D operator (Chapter 4, sec.4.1.2).

(76) ***all*-VP rule:**
Where α is an expression of type <e,τ> (where τ is any type ending in t) whose ordinary translation is α', then:
$[\![\text{all}'\ \alpha']\!] = \lambda z\ \alpha'(z),\ ^{\|}\ \lambda z[\lambda x[gf(Cov_n)(x)]](z)^{\|}$
where *n* is the index on the sister node to *all*.

(77) **composition rule for domain-adjusting meaning:**
Where α is an expression of type <e,τ> (where τ is any type ending in t) with a two-part translation **α**' whose parts consist of an ordinary translation α' and a domain-adjusting meaning δ; and β is an expression of type e with an ordinary

---

Roger Schwarzschild (p.c.) has pointed out to me, if the good fit requirement of *all* is a presupposition we might expect it to project. But the extremely local effect of *all* that will be encoded in the syntactic rules I will give in this chapter makes me wonder whether we would expect any evidence of this sort anyway, since the effect of *all* is so local.

translation β', then:
⟦**α'** β⟧ = ⟦α'⟧(⟦β⟧) , $^{\|}\delta^{\|}$ (⟦β⟧)

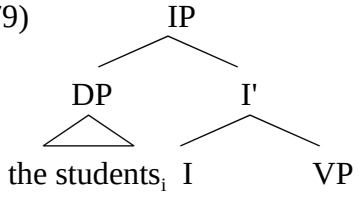
It will be clear how these rules work when we apply them in the course of a derivation. For now note that the combined effect of (76) and (77) will be that the NP argument of the VP will also serve as the argument of the "domain adjusting meaning" introduced by *all*.

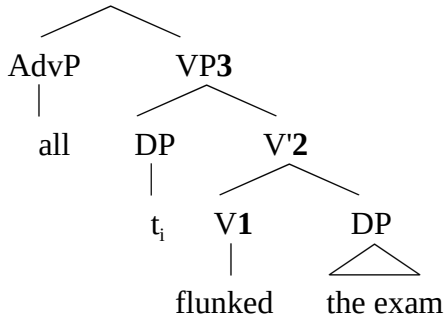
In addition, note that the definition in (77) makes reference to an index *n* on the sister of *all*. But up until now, we haven't been assuming that there is any index on the sister of *all* (that is, on the VP). But let us suppose that the index on *Cov* is inherited from an index on the D operator. (This is close to Schwarzschild's (1996) original rule for the Part operator.) Since the D operator is present in the syntax, its index will be visible in the syntax as well. The rule in (76) ensures that the value of *Cov* in the good fit operator and the value of *Cov* in the restriction of the D operator are the same.

Finally, there are some additional syntactic details that need to be considered. In the derivations I have been using, a VP is not in fact typically of type <e,t>. It is usually of type t, because its argument slots have been filled either by full arguments, or by variables if the syntactic position is filled by a trace. So, for example, a sentence like (78) has the structure in (79), and according to the half-finished derivation below it, the translation of the VP node is an expression of type t.

(78) The students all flunked the exam

(79)





1. λxλy[flunked'(x)(y)]
2. λy[flunked'(the.exam')(y)]
3. flunked'(the.exam')(x_i)

Our rule in (76) requires that *all* applies to an expression that takes an e-type argument. Clearly, the way that line 3 of (79) can be turned into such an expression is by λ-abstracting over the free variable x_i. But I follow Bittner (1994) in assuming that we are not free to bind x_i whenever we want: we can only bind it when its index is "compositionally visible." An index is compositionally visible on a node α if it is the index of α or of α's sister.

To make the index visible, I will propose that *all* is governed by (and hence coindexed with) the nearest DP that dominates it. It should be clear as we proceed that this is just the mechanics we need to capture the idea that *all* takes as its argument the first DP that the VP takes as its argument.

5.1.5.3.1 Nominal government

I will assume that the requirement that *all* be governed by the nearest DP is a kind of antecedent government, as instantiated in the theory of relativized minimality proposed by Rizzi (1990). I will call it nominal government, because the governor must belong to a nominal category (ie NP or DP; we will see when we look at prenominal *all*

that the governor can be an NP).

Rizzi's rules for antecedent government and relativized minimality, (pp.7-8) with *all* substituted in the appropriate places, are given in (80) and (81). In addition, I adopt an m-command requirement, rather than Rizzi's original c-command requirement.

(80) *X* antecedent-governs *all* where
(i) X and *all* are coindexed
(ii) X m-commands *all*
(iii) no barrier intervenes
(iv) relativized minimality is respected

(81) X α-governs Y only if there is no Z such that
(i) Z is a typical potential α-governor for Y
(ii) Z c-commands Y and does not c-command X.

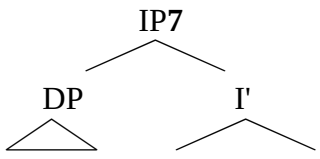
For us a "typical potential governor for Y" is any nominal category. (As we proceed it will become clear that if any other category is allowed to be a governor for *all*, the derivations wouldn't work because *all* requires a DP argument.)

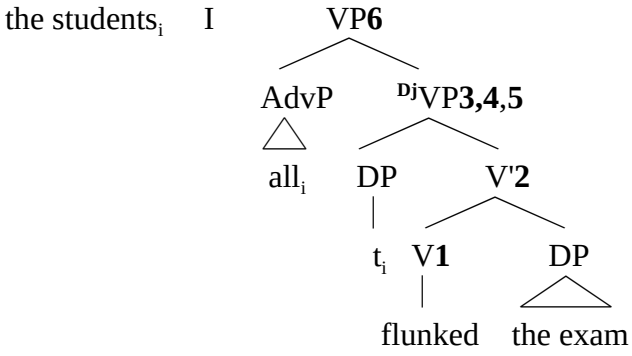
We also need a definition of m-command. I use the following definition, from Chomsky 1986 (essentially the same definition that Rizzi uses).

(82) α m-commands β iff α does not dominate β and every γ that dominates α dominates β, where γ is a maximal projection.

Now we can put this all together. Assuming that a D operator has an index, and using the idea that *all* is coindexed with its DP-governor, the derivation that we started in (79) can be finished as in (83). (Note I am using "eventless" semantics because events are not crucial here.)

(83)





1. λxλy[flunked'(x)(y)]
2. λy[flunked'(the.exam')(y)]
3. flunked'(the.exam')(x_i)
4. λx_i[flunked(the.exam')(x_i)]
5. ^{Dj} λx_i[flunked(the.exam')(x_i)]
6. λz [^{Dj}λx_i[flunked(the.exam')(x_i)]] (z) , ‖λz[λx[gf(Cov_j)(x)](z)‖
6'. λz [[^{Dj}flunked(the.exam')](z)] , ‖λz[gf(Cov_j)(z)]‖
7. ^{Dj}[flunked(the.exam')](the.students'), ‖gf(Cov_j)(the.students')‖
8. ∀x[x⊆⟦the.students'⟧& x∈Cov_j→flunked(the.exam')(x)] ,
‖gf(Cov_j)(the.students')‖

Note that I assume without explicitly formulating a rule that once the meaning of *all* is fully "filled in", the domain adjusting meaing is simply passed up in the derivation without interacting any further with the derivation of the ordinary meaning.

5.1.5.3.2 All *oriented toward the subject*

Thus, on this analysis, the fact that *all* is oriented towards the subject follows from the fact that the subject is the first DP argument of the VP, so it also serves as the argument of *all*. We used a relativized minimality approach to government to help us carry out this idea, which works for us because the first DP that the VP takes as its argument (after combining with *all*) will also govern *all*.

If adverb *all* cannot combine with any constituent lower than VP, then we predict that the only way *all* can be object-oriented is if the object is displaced in such a way that

it can be the first argument to combine with the VP.

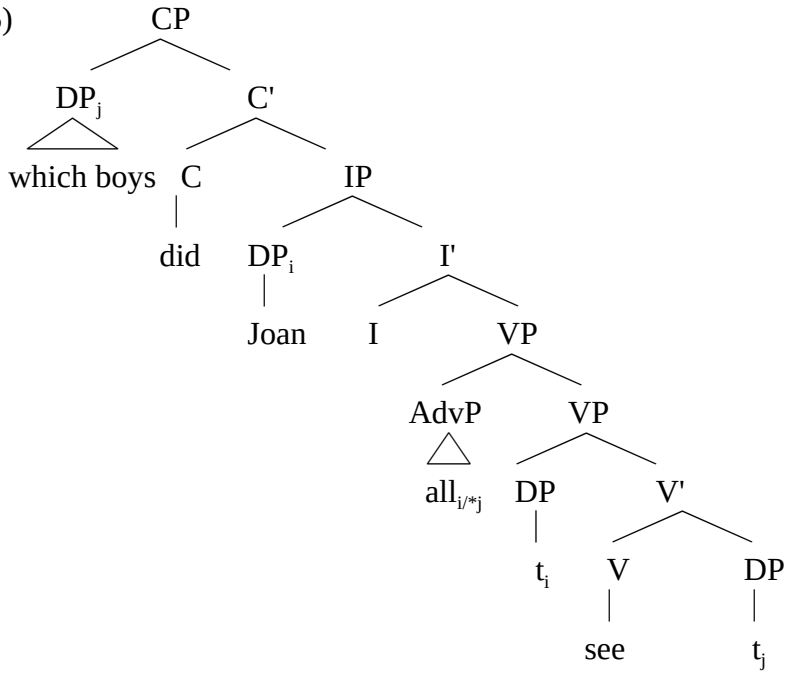
In English, this is completely impossible. Besides NP movement, which *all* is compatible with, objects can be displaced only by *wh*-movement or by topicalization, and in both cases the object is not the first argument to combine with the VP. Thus we expect that *all* could not be oriented towards the object in these cases.

This effect is captured by the fact that *all* must be coindexed with its DP governor. Because we adopt a relativized minimality approach to government, the DP governor of *all* will always be the first DP argument of the VP that *all* combines with.

I show how government from a *wh*-object is blocked in English in (85).

(84) *Which boys did Joan all see?

(85)



All cannot be DP-governed by *which boys* because *Joan* is a closer DP-governor,

according to the definition of relativized minimality in (81). Note that this has the welcome effect of blocking λ-abstraction over the variable x_j at the VP level; if we were to abstract over that variable then *Joan* would incorrectly be "fed in" as the object of *see*. The DP *Joan* is a legitimate DP governor for *all*, but I assume that this derivation crashes because *all* is dependent on distributivity over *Joan*, but a D operator on the VP is not licensed here (for economy reasons) since *Joan* denotes a singular entity.

5.1.5.3.3 All *in sentence-peripheral position*

The requirement that *all* be governed by a nominal category has another welcome consequence (which I thank Viviane Déprez for pointing out to me). There is one aspect of *all*'s distribution that makes it different from other speaker adverbs: speaker adverbs like *probably* and *surely* can appear at the periphery of sentences, but *all* can never appear in either position.

The sentences in (86)-(87) show that the speaker adverb *probably* can appear with a pause at the right edge, and without a pause at the left edge, as shown in (86)-(87)

(86) Probably Max and Susan like crab cakes
(87) Max and Susan like crab cakes, probably

This is completely impossible for *all*, as shown in (88)-(89)

(88) *All Max, Susan, and Emily like crab cakes
(89) *Max, Susan, and Emily like crab cakes, all

If we suppose that in both the right-adjoined and the left-adjoined cases the adverb is adjoined to IP node above the subject, then *all* will fail to be governed by a nominal category, and hence will be unlicensed. So the requirement that *all* be governed rules out these sentences as ungrammatical.

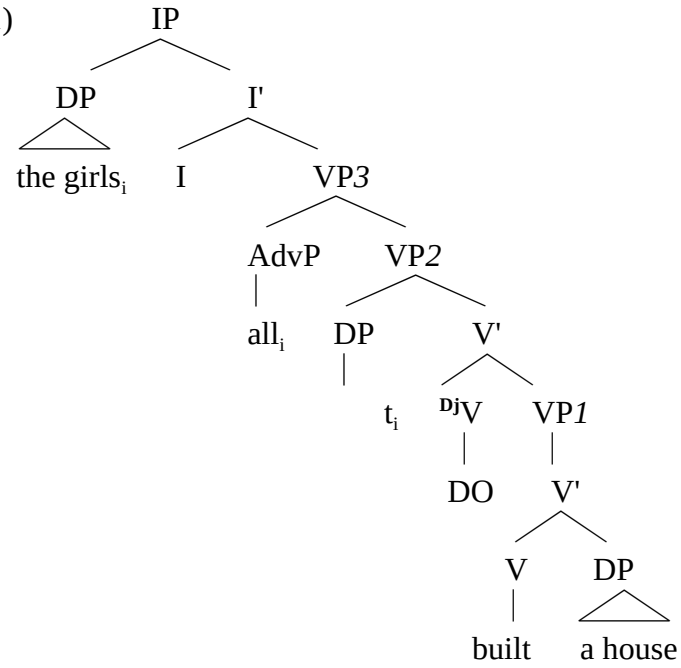
5.1.5.3.4 The D index

The presence of the index on the D operator in (83) was crucial to making the rule in (76) work. But we saw in Chapter 4 that there are cases where the D operator is not conveniently placed on the sister of floated *all*. These are cases like (90).

(90) The girls all built a house

By hypothesis, on the collective reading of (90) the D operator is positioned on DO, the head of the higher verbal projection. If the structure of (90) is (91), then it's not clear how an index comes to be on the sister of *all*.

(91)



One possible way to solve this problem is to suppose that the index on the D

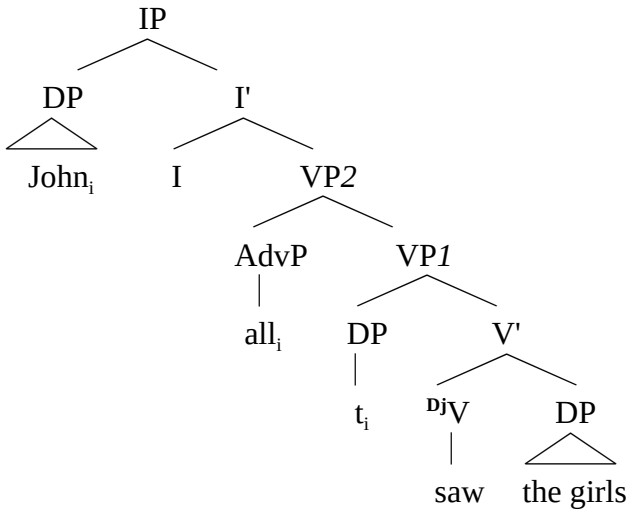
operator may percolate up to VP*2*, where it is visible to *all*. This is what I will ultimately propose.

But there is a kind of example that shows that we must put limits on how far this percolation may go. This is shown by the example in (92). (Thanks to Viviane Déprez and Roger Schwarzschild for pointing this example out to me.)

(92) *John$_i$ all$_i$ ^{Dj}[saw] the girls

The situation we want to avoid is one where the index *j* on the D operator on *saw*, (over *the girls*), is allowed to percolate up to VP where it can provide the index for the occurrence of *Cov* for the good fit operator according to the rule in (76).

(93)



If we allowed the index *j* to percolate up to VP*1* in (93), then we would incorrectly predict the sentence to be grammatical. This is because we have been supposing that what rules out floated *all* in a sentence like (92) is an economy condition:

D is not licensed here because of the singular subject, so *all* has nothing to "operate on" (ie, no *Cov* variable to be coindexed with). But if there is a plural object and we do not restrict the percolation of the index *j* in (93), then the *Cov* variable inside the good fit operator could be coindexed with Cov_j, which is clearly not the right result.

It is intuitively clear what we need to do to avoid this: we need to make sure that the *Cov* variable introduced by *all* gets coindexed with the "right" DP — that is, the DP most local to it. To handle this, I will introduce a "D index percolation rule". The idea that this rule is intended to capture is that a D index cannot percolate beyond the point where a distributed node combines with its distributee.

(94) **D-index percolation rule**: The index of a D operator can percolate up to the first higher node that directly dominates a nominal category, and no farther.

Since the first node up from a D operator that directly dominates a nominal category will always be the node that dominates *the* nominal category that the D operator distributes over, this rule essentially instantiates the local relation between the D operator and the DP it distributes over. This will in turn have the right results for floated *all* because *all*, too, must be in a local relationship with a DP. Therefore a local relationship between a D operator and a DP (reflected in part by the D index percolation rule) and a local relationship between a DP and *all* (instantiated in terms of government) conspire to ensure the right kind of locality between *all* and the D operator.

With this rule, the reader can verify that the D operator index *j* can percolate up to VP*2* in (91), but only up to V′ in (93). Thus we can apply the rule in (76) to get the right results.

5.1.5.3.5 All *oriented toward the object*

In section 5.1.5.3.2, we saw that the proposed account of the "orientation" of *all* correctly predicts that floated *all* can only be construed with the subject in English. The explanation given there also predicts that object-orientation should be possible in a language in which it is possible for the object DP to be in a position to DP-govern *all*. It appears that this prediction is borne out by French, because French allows object-orientation for floated *all*. I will not attempt a full accounting of the syntax of French floating quantifiers here (see Kayne 1975, Sportiche 1988, and Doetjes 1992 for discussion), but I will give a sketch of what we would expect on the account I have proposed.

A significant fact about French floating quantifiers is that although they allow object orientation, it is never permitted for an in-situ object, as in (95). The object must be displaced either by clitic movement (96) or by *wh* movement (97).

(95) *Jean ont tous$_j$ vu les enfants$_j$ (in-situ object)
John has all seen the children

(96) Jean les$_j$ ont tous$_j$ vous (clitic mvt)
John them-CL has all seen

(97) Les enfants$_i$ que tu as tous$_i$ grondés sont partis en pleurant (*wh*-mvt)
the children who you have all scolded have left crying

In the case of (96) it seems fairly clear that the object clitic *les* is in a position to DP-govern *tous*. But the facts get more complicated rather quickly, because in (97) *tu* (you) appears to intervene to block DP-government of *tous* by the *wh*-moved *les enfants* (or by comp).

However, using (96) as a guide, it seems plausible to suppose that French *wh*-movement of objects might involve, at least optionally, an intermediate landing site —

perhaps specifier of AgrO. This kind of analysis has been proposed by Sportiche (1992) for clitic movement with participles. If the definition of DP-government were expanded so that DP traces could be DP governors, then this would account for (97).

In general then, we would expect that DP government of the sort we find in (97) would be licensed by the presence of DP traces in intermediate positions. The difference between French and English would be that French makes much more "active" use of positions in the Agr functional projections (especially for objects) in overt syntax because it has overt verb movement (see, eg, Pollock 1989, Sportiche 1992). The fact that French only licenses object oriented *tous* when the object has moved is then expected.

Let us summarize the account we have given of *all*'s orientation. We have proposed that *all*'s orientation toward a subject (or an object in French) follows from the fact that adverb *all* takes as one of its arguments the first type e argument that its sister VP combines with; and the question of which argument *all* will be oriented towards is answered by the syntax. In English objects can never govern *all*, but in French, we have speculated that the possibility of clitic movement and use of the AgrO position in the overt syntax makes it possible for displaced objects (or object traces) to DP govern *all*.

5.1.5.3.6 Why all *can't adjoin to V*

The explanation I've given so far for the orientation of *all*, and why its orientedness differs from *together*'s, depends in part on the fact that *all* can adjoin only to VP (or higher), but *together* can adjoin to V or VP. In this section I will show that this difference does not have to be stipulated: there is independent evidence for it.

There are basically two kinds of evidence in support of the hypothesis that

together can occur "lower" in the syntax than *all* can. The first kind of evidence is the evidence from distribution. We saw earlier that *all* is classified as an S-adverb according to the criteria discussed by Jackendoff. By these same criteria, *together* can arguably be classified as a VP adverb.

The only position that is unambiguously a VP-adverb position is sentence-final position without a pause. If S-adverbs appear in sentence-final position, a pause must be insterted. *Together* appears in sentence-final position without pause, which suggests that it is a VP adverb.

(98) Jack and Jill left together
(99) Jack and Jill left quickly
(100) Jack and Jill left, probably

I argued in section 5.1.3.1 that both VP adverbs and S adverbs can appear after two auxiliaries if the second auxiliary is *have*. *Together* seems to me to be a bit awkward in that position, shown in (101). But since the position allows both S adverbs and VP adverbs, I take it that this fact doesn't have any bearing on our argument about which class of adverbs *together* belongs to and it must be ruled out by something else.

(101) The boys will have together fried the tomatoes

Jackendoff says that VP adverbs are "odd" between two auxiliaries, but S adverbs are permitted there. We find that *together* is, as predicted, odd there.

(102) The boys have together been frying the tomatoes.

Another position which is supposed to allow only S adverbs is the position before two auxiliaries. But we can't tell for sure whether or not adverb *together* can appear there because of the existence of adnominal *together*.

(103) Jack and Jill together were frying the tomatoes

So we find that *together* has the distribution we would expect if it were a VP adverb: the most unambiguous evidence for this is the fact that it can occur in sentence final position without a pause. But *all* has the distribution of an S adverb.

Another kind of evidence for treating *together* as being attached lower than *all* comes from a comparison of their respective semantics.

In Chapter 4 I treated *together* as a modifier of events, following the proposal given by Lasersohn (1995). Recall our definition of *together* from Chapter 4.

(104) **together':** $\lambda P \lambda x \lambda e[P(x)(e) \& \forall y,y',e',e''[y \in x \& y' \in x \& y \neq y' \& P(y)(e') \& P(y')(e'') \& e' \subseteq e \& e'' \subseteq e \rightarrow \tau(e') \circ \tau(e'')]]$

The requirement that particular subevents "overlap" is basically the core of the semantic contribution of *together*.

In general there is a correspondence between the syntactic position of an adverb and what it is taken to modify. Parsons (1990) argues that the adverbs that occur closest to the verbs are best analysed as modifiers of events. This idea also plays a role in the theory of the scope of adverbs proposed in Ernst (1998).

So the evidence from Jackendoff's tests that *together* is attached relatively low in the syntax is corroborated by the semantics of *together*. In fact it is probably more accurate to turn that statement around: the fact that *together* is a modifier of events leads us to expect that it should occur in positions closer to the verb.

On the other hand, *all* does not say anything about events: it affects the assertion that is made. Of course, if we take assertions to be "about" events in some sense, then *all* will have an indirect effect on events. But then, so does a modal adverb like

probably; if I assert *probably S* then I assert that whatever event is picked out by *S* probably happened. Still, it is not the case that *probably* is a modifier of events.

So if the syntactic position of an adverb depends at least in part upon its meaning (that is, what it modifies), then the fact that *all* occurs higher in the syntax (at least above VP) is expected.

In other words, we can conclude that it is predictable from their respective semantics that *all* should occur higher in the syntax than *together*.

5.2 The syntax of prenominal **all**

5.2.1 Comparing prenominal and floated all

Since we have just proposed a syntax for floated *all*, perhaps the first question that arises when we turn our attention to prenominal *all* is, what is the relationship between floated *all* and prenominal *all*? The answer that I will give is close to the account proposed by Lasersohn (1990, 1995), for the relationship between adnominal and adverbial *together*.

Lasersohn's definition of *together* is such that it takes one argument of type *e* and one argument of type <e,t>. He argues that adnominal and adverbial *together* have the same core meaning, and the difference is simply a matter of which argument gets "fed in" first. That is, abstracting away from the details, adverbial *together* is defined as in (105) and adnominal *together* is (106).

(105) λPλx[together'(P)(x)]
(106) λxλP[together'(P)(x)]

I also will propose that the core meaning of *all*, whether adverbial or prenominal, is constant. The effect of the rules in (105) and (106) is similar to the effect we would get

if we simply added λP to the definition of *all* when it combines with a VP, that is, it makes it possible for *all* to combine with a verbal argument. This is similar to the effect Lasersohn achieves simply by swapping the order of λP and λx above.

It should perhaps be pointed out that this proposal by Lasersohn has been criticized by Schwarzschild (1994a) on the grounds that there are some differences between adnominal *together* adverbial *together* that it doesn't explain. Likewise, there is some evidence from the interaction of prenominal *all* and adverbs that there are some differences. It comes from the distribution of manner and agent-oriented adverbs with an object DP containing *all*.

Based on the fact that floated *all* cannot appear in the scope of an agent-oriented or manner adverb, we might expect that these adverbs in combination with prenominal *all* in object position would lead to some difficulty. However, these kinds of sentences are perfectly grammatical.

(107) Jack skillfully climbed all the beanstalks.
(108) Evelyn bravely fought all the lions

A manner interpretation for *skillfully*, and an agent-oriented interpretation of *bravely*, both seem to be perfectly fine here. It's quite clear that the oddity we found with (48)-(49) are gone.

In spite of this fact, the overwhelming majority of the data seems best handled by treating prenominal and adverbial *all* on a par. For this reason, I will propose that their meaning is exactly the same, as we will see shortly.

But before we look in detail at how prenominal *all* combines with a DP, I will first look at two closely related issues: the category of prenominal *all*, and the syntactic

position of prenominal *all*. I will argue that *all* is a kind of adjective/degree word, and that it occupies the specifier of DP.

5.2.2 The category of prenominal all/both

I propose that prenominal *all* (and also *both*) has the features of degree words and adjective words, and hence is some kind of mixed degree/adjective category. The evidence for this is mostly circumstantial, but nevertheless intriguing.

Haspelmath (1995) reports that *all* is historically speaking derived from the adjective *whole*. *All* is clearly not a simple adjective anymore (**the all girls*) but nevertheless the fact that it is historically derived from one lends some plausibility to my proposal that it is at least part adjective.

Furthermore, by supposing that prenominal *all* is somewhat "adjectival," this gives us a plausible explanation for why *all* can occur both prenominally and as an adverb. The path in English from adjective to adverb is a very old and well-traveled one: *slow-slowly, deliberate-deliberately, possible-possibly*, et cetera. Although floated *all* lacks the usual adverbializing suffix *-ly*, it is nevertheless plausible to suppose that the adjective-adverb route is what explains the existence of both forms of *all* in modern usage.

One way to look for evidence for the claim that *all* is adjectival (and adverbial on its floating incarnation) would be to compare the structure of adjective and adverb phrases and see if prenominal and floated *all* share the same kinds of similarities. However, this kind of approach would be unlikely to turn up any decisive evidence for the adverb/modifier hypothesis of *all* over the quantifier hypothesis of *all*. The reason

is that it has been shown by many authors (Bresnan 1973, Jackendoff 1977, Abney 1987) that the most obvious similiarities between the structure of adjective phrases and adverb phrases, namely the degree modifier system, is also shared by quantifier phrases, by and large. Of course, the idea that adverbs, adjectives, and quantifiers are syntatically very similar is interesting; but it also would stymie our attempts to find decisive evidence that *all* is an adjective and an adverb and not a quantifier.

I also think that assigning a degree/adjective category to *all* is plausible in light of the semantic analysis of *all* that I proposed in Chapters 2 and 3. The idea that it operates on the context in some way is consistent with its being a functional category, but the idea of *good fit* could also been seen as a kind of lexical meaning.

I will offer one more circumstantial piece of evidence for the hypothesis that *all* is a degree/adjective category, but this will come after we have looked at the syntactic position of prenominal *all* in the next section.

5.2.3 The position of prenominal all

What exactly is the structural position occupied by prenominal *all*? We have basically three choices. *All* could be adjoined to DP; *all* could take DP as its complement; or *all* could be in the specifier of D. I will argue that *all* occupies the specifier position of D. The evidence comes from the non-occurrence of *all* with DPs that arguably lack a specifier position, and from the fact that, pace Sportiche, there is no evidence that *all* can be 'stranded.'

There are three types of constructions that show this. For every construction to be discussed below, a floated *all* is acceptable, indicating that the meaning of the DP in

question is perfectly compatible with *all*. Note also that the acceptability of floated *all* with these constructions means that they pose a problem for the stranding hypothesis.

5.2.3.1 Evidence from pronouns

The first fact is that *all* cannot occur to the left of a pronoun. This was observed for object positions by Maling (1976), but it is also true of nominative pronouns.

(109) *Jack saw all them/all us/ all you
(110) Jack saw them all/us all/ you all

(111) *All they/all we/ all you left
(112) They/you/we all left

5.2.3.2 Evidence from PRO

As we might expect given the facts above, it also appears to be the case that *all* cannot occur immediately to the left of PRO, as the following sentences show (taken from Baltin 1995).

(113) *All to leave would be a shame
(114) To all leave would be a shame

5.2.3.3 Evidence from conjunction

All is quite bad in combination with a conjunction of singular terms, despite the fact that these constituents are semantically interpreted as plural individuals.

(115) *All Peter, Paul and Mary got strep throat.
(116) Peter, Paul, and Mary all got strep throat.

(117) *All the butcher, the baker and the candlestick maker got married
(118) The butcher, the baker, and the candlestick maker all got married

So now the question is, what do these kinds of things, pronouns, PRO, and conjunctions, have in common? One obvious possibility is: that none of these kinds of constituent, arguably, have the usual internal structure of complement and specifier. This

suggests that the syntactic position of *all* is in the specifier of D. This would quite straightforwardly account for the failure of *all* to appear in these environments.

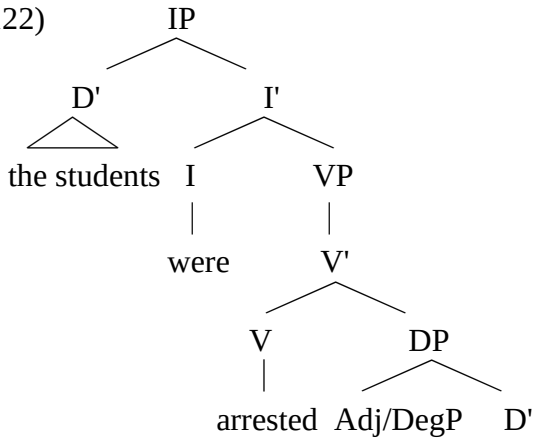
Note also that our other two options, namely that *all* takes a DP-complement or that *all* is adjoined to DP, would fail to explain the distribution of *all* that we have just seen.

If we assume that NP movement, topicalization, and *wh* movement cannot target D', this proposal predicts that it should not be possible to 'strand' *all*, contrary to Sportiche's hypothesis. In fact there is no decisive evidence that *all* can occur immediately to the left of a trace of *wh*-movement or NP movement. The following would be convincing evidence if grammatical, but they are both ungrammatical (in English, at least).

(119) *The students were arrested all
(120) *Who did you see all?
(121) *The plants, I like all

The ungrammaticality of the sentences in (119)-(121) is expected on the hypothesis that *all* occupies spec, DP. The sentence in (119), for example, would require that passive movement be able to target D', as shown in (122).

(122)



| |
all t

If we were to suppose that *all* is adjoined to DP, or takes DP as its argument, we would have to say something extra about why (119) is not grammatical. On the hypothesis that *all* is in specifier of D, it follows from general assumptions about the nature of phrasal movement.

5.2.3.4 Interesting circumstantial evidence

There is circumstantial evidence from one construction in English that makes it plausible to suppose that an adjective/degree category can appear in the specifier of D. Examples of this construction are given in (123).

(123) too large a house
how heavy a box
as nice a man

Note that these phrases behave like noun phrases. For example, they can appear in argument positions.

(124) John bought too large a house
How heavy a box can you carry?
Yesterday I met as nice a man as you could hope to meet

These kinds of phrases are not completely free as argument noun phrases; see Bresnan (1973) for some discussion. But the well-formedness of the sentences in (124) suggests that they are in fact noun phrases, so the AP *too large*, for instance, must be either adjoined to or in the specifier of the DP.[9] I have not been able to find any evidence

[9] Abney (1987), on the other hand, argued that the head of the phrase is really the adjective, which somehow inherits a substantivizing feature from the noun, so it can behave like a noun

for one or the other syntactic position, but since I am arguing that *all* occupies the specifier of D it seems reasonable to suppose that the APs in (123) may occupy the same position.

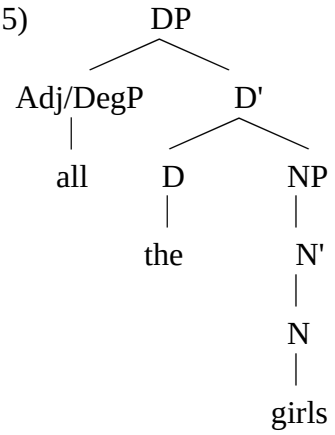
The presence of the degree words in the APs in (123) is of course crucial, but this fits nicely with my argument that *all* is a degree/adjective.

Note also that we don't need to be concerned that *all* itself can be modified by degree words like *almost* and *nearly*, since it is well known that the degree system is recursive.

5.2.4 The syntax-semantics of prenominal all

The arguments of the previous section lead us to conclude that the structure of a DP like *all the girls* is the structure shown in (125), with *all* occupying spec,DP.

(125)



One consequence of this structure is that we can maintain the requirement,

---

phrase.

introduced in the section on floated *all*, that *all* must be governed by a nominal category (thanks to Viviane Déprez for pointing this out to me). This is then another point of similarity between prenominal and floated *all*: because we use m-command in our definition of nominal government, the NP *girls* governs *all*.

Let us see now how to interpret this determiner phrase, and then how to interpret the DP in a sentence. Below is rule we have already seen for interpreting *all*, repeated from section 5.1.5.3.

(126) ***all*-translation rule:**
all has no ordinary translation, and a domain-adjusting meaning of
$^{\shortparallel}\lambda x[gf(Cov)(x)]^{\shortparallel}$

To interpret the DP in (125), we need to combine *all* with the meaning of the D'. We can do this with an adjustment to our "composition rule for domain-adjusting meaing". The rule, also from section 5.1.5.3, was originally written to facilitate composition of a domain-adjusting reading when a VP with *all* and an NP combine. It is repeated below. (127) **composition rule for domain-adjusting meaning (in VP):**

Where α is an expression of type <e,τ> (where τ is any type ending in t) with a two-part translation α' whose parts consist of an ordinary translation α' and a domain-adjusting meaning δ; and β is an expression of type e with an ordinary translation β', then:
$[\![\boldsymbol{\alpha}' \beta]\!] = [\![\alpha']\!]([\![\beta']\!])\ ,\ ^{\shortparallel}\delta^{\shortparallel}\ ([\![\beta']\!])$

It needs to be adjusted so that it can apply in the environment we are concerned with here, inside DP. I make the needed adjustments in the rule below.

(128) **composition rule for domain-adjusting meaning (in DP):**
Where α is *all*, with a domain adjusting meaning δ, and β is an expression of type e with an ordinary translation β', then:
$[\![\boldsymbol{\alpha}' \beta]\!] = [\![\boldsymbol{\beta}']\!]\ ,\ ^{\shortparallel}\delta^{\shortparallel}\ ([\![\boldsymbol{\beta}']\!])$

The similarities between the two rules should be clear. In each case, the ordinary

denotation of an NP becomes part of the domain-adjusting meaning introduced by *all*. Perhaps it would be possible to devise a single, more general rule that captures the effect of these two rules. I refrain from doing so, however, because of the risk that the rule would be too general and allow for the generation of unwanted domain-adjusting meanings. For now I want to ensure that the "cross-translation" function application is limited to those environments where the domain adjusting meaning of *all* has an open argument slot. These rules generate the following translation for the DP *all the girls* in (125).

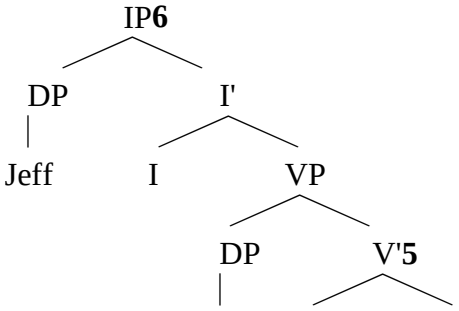
(129) the.girls' , $\|gf(Cov)(the.girls')\|$

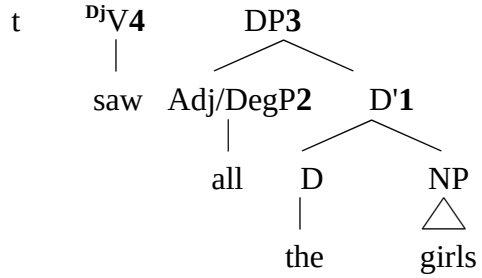
We still need one more composition rule, which will guarantee the *Cov* variable in the meaning of the DP *all the girls* will get the right index when it combines with a VP. This rule is given below.

(130) **NP VP rule:**
Where α is an expression of type <e,τ> (where τ is any type ending in t), with translation α' and β is an expression of type e, with an ordinary translation β' and a domain-adjusting meaing $\|gf(Cov)(\beta')\|$, then:
$[\![\alpha'(\beta')]\!] = [\![\alpha'(\beta')]\!]$, $\|gf(Cov_n)(\beta')\|$
where *n* is the index on α.

We can apply this rule to a sentence like (131), which is derived according to (132).

(131)





(132)
1. the.girls'
2. ‖λx[gf(Cov)(x)]‖
3. the.girls' , ‖gf(Cov)(the.girls')‖
4. λx ⊓{ λyλy'[saw(y)(y')] (z) | z⊆x & z∈ ⟦Cov$_j$⟧}
5. λx ⊓{ λyλy'[saw(y)(y')] (z) | z⊆x & z∈ ⟦Cov$_j$⟧ } (the.girls') ,
   ‖gf(Cov$_j$)(the.girls')‖
   = ⊓{λy'[saw(z)(y')] | z⊆⟦the.girls'⟧ & z∈ ⟦Cov$_j$⟧} , ‖gf(Cov$_j$)(the.girls')‖
6. ⟦J⟧ ∈ ⊓{λy'[saw(z)(y')] | z⊆⟦the.girls'⟧ & z∈ ⟦Cov$_j$⟧} , ‖gf(Cov$_j$)(the.girls')‖

Note that once *all the girls* has combined with *saw*, we have a meaning with the same kind of structure that we saw earlier in section 5.1.5.3: there is an "ordinary" meaning and a denotation adjusting meaning.

5.3 Other Issues

5.3.1 Possessives

There are some apparent problems for the proposal that *all* occupies spec,DP. The first is that *all* is perfectly compatible with possessives, including pronominal possessives.

(133) All David's children are dentists
(134) All his children are dentists

The pronominal case is simpler, so let us deal with that one first. In this case there's no reason to assume that anything occupies the spec of *his*, so there is space there for *all*. I don't believe there is any contradiction in saying that pronouns do not license a specifier position but posessive pronouns do; if a possessive pronoun is some kind of

contraction of the possessive marker and a pronoun, then it could be that the possesive "part" of the contraction licenses the specifier position.

On the other hand, (133) is a bit harder to handle, because it appears that *David* is occupying the specifier position of the possessive marker. I don't think it is a viable option to suppose that *all* is in the specifier of *David* for two reasons: first, because proper names arguably also project directly to phrases, and so lack a specifier. Secondly, this would lead us to expect the wrong constituency relations anyway.

So instead I will propose that a possessive can license *two* specifier positions, if necessary, to accommodate both *all* and the possesor DP. Notice that our generalization that led us to propose the spec,DP position for *all* said that *all* cannot appear as a constituent with nominal phrases that completely lack internal structure. This being the case, we would not expect them to be able to 'sprout' structure just to accommodate *all*. But the possessive marker is different, because it clearly does license structure inside the phrase, so it is possible to suppose that it can sprout more structure if necessary. The idea would be that the fundamental difference is between heads that license phrasal structure, and heads that do not.

Note furthermore that the order of the two specifiers is as we would expect. Since the possessor is selected by the posessive head, it is expected that it should appear in the specifier position closest to the head.

5.3.2 *Conjunction (redux)*

There are some issues related to the conjunction cases that need to be addressed. The hypothesis that conjoined DP's lack a specifier position accounts neatly for the

ungrammaticality of phrases like **all Peter, Paul, and Mary*. However, conjunctions of plural elements do allow *all*, although it is very unclear what the syntactic position of *all* might be in these constructions.

(135) All the students and the teachers pitched in to make the picnic a success
(136) All the students, the teachers, and the parents pitched in to make the picnic...

Of course it is possible in both cases that *all* is in the specifier of only the first DP in the conjunction, *the students.* But the reading we are interested in is one where *all* seems to be related to the entire conjunction. On this reading (136) seems to me not quite as good as (135), but still it's not crashingly bad as we saw was the case with conjunction of singular terms.

The only suggestion I can make at this point is that in fact *all* really does sit in the spec of the first DP in the conjunction, where it is perfectly legitimate. The reading where it seems to belong to the whole conjunction is perhaps due to some kind of accomodation process, where *all* sort of "spreads out" to cover the whole conjunction. In support of this notice that it's awfully hard to get *all* <u>not</u> construed with the whole conjunction in (135). If you mean specifically all the students, you almost have to change the order so it's clear: *the teachers and all the students*. If the difference between the "one conjunct" and the "whole conjunction" reading of (135) were simply a matter of syntactic ambiguity, then it should be much easier to separate out the readings.

5.3.3 Both *and conjunction*

I have argued that *all* and *both* form a class, and this leads us to expect that prenominal *both* should also not be licensed with conjunctions. But the sentences below show that this is clearly not the case.

(137) Both John and Bill have read that book
(138) Both the butcher and the baker read the newspaper every day

However, this error of prediction is only apparent, because *both*, unlike *all*, has a homonym whose function is to "introduce" conjuctions (see, for example, Stockwell, Schachter and Partee 1973, Edmonson 1978). This *both* can cross-categorially introduce a conjunction of almost any type (note that *all* would be ungrammatical if substituted for *both* in every sentence below).

(139) Marie both sings and dances
(140) This painting is both stunning and understated
(141) Alex planted seeds both in the flowerbed and in the windowbox
(142) Pete fixed the plumbing both quickly and effortlessly

Furthermore, it appears that this conjunction-introducing *both* has a different meaning from the good-fit *both* that is related to *all*. Conjunction *both* apparently has a meaning that corresponds to an old-fashioned conjunction reduction transformation. In other words, a sentence with conjunction *both* is interpreted as if it was "reduced" from two sentences, which has the effect of yielding a strictly distributive reading. This would make (143)a, for example, equivalent to (143)b, and so on.

(143) a. Both Harry and James answered the telephone
b. Harry answered the telephone and James answered the telephone

(144) a. Pete fixed the plumbing both quickly and easily
b. Pete fixed the plumbing quickly and Pete fixed the plumbing easily

(145) a. Carol put a couch both in the livingroom and in the den
b. Carol put a couch in the livingroom and Carol put a couch in the den.

Note that (145)a can't mean that Carol installed a couch that was big enough so that half of it was in the livingroom and half in the den (as we might expect if conjunction *both* allowed a kind of "collective" reading here). It can only mean that two

different couches were put in.

Once we have clarified the difference between good fit *both* and conjunction *both*, we have an explanation for the otherwise puzzling data that was noted in Brisson (1996) and in Chapter 3 (section 3.4). Good fit *both* has a freer distribution with certain collective predicates than does conjunction *both*, as shown in (146)-(147)

(146) a. Pete and Les both collided
b.*Both Pete and Les collided

(147) a. Jane and Sarah both left together
b. *Both Jane and Sarah left together

The (a) examples contain good fit *both*, and the (b) examples contain conjunction *both*. Since conjunction *both*, but not good fit *both*, imposes a strictly distributive reading, the difference between the (a) and the (b) sentences is explained. If we change conjunction *both* to prenominal good fit *both*, we expect that the sentences should get better, and they do. Although judgments at this point become a bit murky, I think there is a contrast between the sentences below and the (b) sentences of (146)-(147), and I have confirmed this with other speakers.

(148) Both (the) boys collided
(149) Both (the) girls left together

In addition, if the ungrammaticality of the (b) sentences in (146)-(147) is due not to the prenominal position of *both* but to the meaning of conjunction *both*, we predict that we should <u>not</u> find any prenominal/floated contrast with prenominal versus floated *all*, since *all* does not have a homonym corresponding to conjunction *both*. This prediction is correct.

(150) a. The cars all collided

b. All the cars collided

(151) a. The girls all left together
b. All the girls left together

The fact that conjunction *both* requires a strictly distributive reading leads us to expect that there should be a contrast between the (a) and the (b) sentences in (152) and (153), using familiar predicates from Chapters 2 and 3.

(152) a. Both the boys ate a sandwich
b. Both John and Bill ate a sandwich

(153) a. Both the professors built a house
b. Both Jane and Alan built a house

The (a) sentences should allow a collective reading, while the (b) sentences should not. Unfortunately judgments are not as sharp here as the analysis predicts they should be; but it does seem to be the case that a collective reading is at least easier for the (a) sentences than for the (b) sentences.

5.3.4 "Q-pro Flip"

The hypothesis that prenominal *all* must sit in the specifier of D explains why (154) is bad. However, it doesn't explain why (155), which represents the phenomenon Maling (1976) described as Q-Pro flip, is good.

(154) *Evelyn saw all them
(155) Evelyn saw them all

I really am not sure what to say about this phenomenon, except I do have one observation that I think has not yet been made. It appears that in the sentence that has undergone the "flip," *them all* is not a constituent (which suggests that "Q-Pro Flip" is a misnomer, since it implies constituency).

'Them all' is not a good answer to a question; and it can't be topicalized, in contrast to the presumably synonymous *all of them*.

(156) *them all, I like.
all of them, I like

(157) Which cookies did Rhonda eat?
*them all
all of them

This being the case, it would appear that we are forced to say that *all* is adjoined to a verbal projection. But then this environment (with an accusative pronoun) is the only environment that licenses *all* in that position; it's pretty bad when the object is a full DP, and it's also clearly not possible to construe a postverbal, post-pronominal *all* with the subject.

(158) *?Evelyn saw the boys$_i$ all$_i$
(159) *The girls$_i$ kissed him all$_i$

At this time I don't have any explanation for the existence of this phenomenon.

5.4 All/both *inside a partitive*

In Chapter 1 we saw that *both* is not permitted to appear as part of the lower phrase of a partitive. *All* is not grammatical there, either.

(160) *Each/one/most of both boys have eaten dinner
(161) *Each/one/most of all the girls have gone outside

The data with *both* was problematic for Barwise and Cooper's (1980) proposal that *both CN* forms a definite description equivalent in meaning to *the two CN*. This is what led Ladusaw (1983) to propose that *both* has a distributive component to its meaning.

The possibility of *all/both* occurring inside a partitive is ruled out by the proposal

I have made here. We have seen that *all/both* are dependent on the presence of a D operator in order to be licensed. But partitive phrases are incompatible with the presence of a D operator.

In partitive constructions like the ones in (160), *both*/*all* are dominated by a quantifier or an indefinite. Both a quantifier and an indefinite introduce a different kind of quantification into the sentence from the quantification of a D operator. If there is no D operator present, then *both* inside a partitive is not licensed.

In principle we might wonder whether it is possible to insert a D operator in a sentence where we already have a quantifier like *every girl*, or an indefinite like *two boys*. However this would yield ungrammaticality of the same sort that we saw in Chapter 4, section 4.3, with *every plane landed together,* where the same set is quantified over twice. So the proposal I have made here correctly predicts the ungrammaticality of *all/both* in downstairs position in a partitive.

Chapter 6

Conclusions

6 Introduction

The core ideas defended in this dissertation are that floating quantifiers are not determiner quantifiers, and they don't float. The first step toward our analysis was to develop an account of the meaning of *the CN* and *all the CN* which correctly predicts their differences. This analysis proposes that *all* is not a quantifier, but a modifier whose meaning interacts with the quantification introduced by a distributivity operator. With this analysis we predict a wide range of *all*'s behavior, including its puzzling distribution and the many differences in meaning and distribution between *all* and *every*. Much of this had been previously unaccounted for in the literature.

We then show that once we abandon the idea that *all* must be a quantifier, we don't need to postulate movement to explain how *all* can appear in postnominal position. The relationship between prenominal *all* and floated *all* turns out to be not much more complicated or surprising than the relationship between *slow* and *slowly* or between adnominal and adverbial *together*. The abandonment of the stranding hypothesis also leads to some welcome empirical results, in particular an improved account of *all*'s adverb-like distribution and semantic behavior.

6.1 Some consequences of the hypothesis

The theory proposed in the preceding five chapters has consequences for various issues in semantic and syntactic theory that are perhaps hard to see in the dense thicket of

argumentation for the various pieces of the proposal. Now that we are at the edge of the forest where the light is a bit better, we can take a look at some of these consequences.

6.1.1 Distributivity

The first and most obvious consequence for distributivity of the theory I have proposed here is that it exists as part of grammar. Secondly, the D operator must be a syntactic object. This is required by many different strands of the analysis in Chapter 4, including the scopal differences betwen *all* and *every* and between *the +CN* and *every*. Finally, insertion of the D operator must be free and optional, with the ill-formed cases ruled out by "other factors". The two "other factors" that we have seen are the semantic requirement that the D operator apply to a constituent that takes an argument of type e, and the requirement that insertion of the D operator not violate economy of derivation (this second fact is discussed below).

Except for the evidence that insertion of the D operator is subject to economy considerations, all of the foregoing properties of distributivity have been previously proposed and defended in the literature. The analysis I've given here provides additional evidence that each of them is correct.

6.1.2 Collectivity

The results of this dissertation imply at least two new consequences for a view of collectivity. First, we have found in this dissertation that collectivity is not a unified phenomenon, as is assumed in nearly every account of distributivity and collectivity that I know of. We have seen evidence that there are two different kinds of collectivity, one that involves hidden distributivity, and one that does not.

As early as Chapter 2 we saw arguments that collectivity must involve some kind of quantification, because nonmaximality was turning out to be a quantificational effect and nonmaximality is possible with (some) collective predicates. I used this as one of my arguments against the groups approach to nonmaximality, because the groups approach treats nonmaximality as a subcase of collectivity, and collectivity as just the absence of quantification over a plural.

The results of Chapter 4 give us another kind of evidence that collectivity involves quantification. There we saw that *all*, pragmatic weakening, and *except* phrases are all sensitive to something that is present in some sentences with collective predicates, and not in others (this was the data relevant to Taub's generalization). We proposed that that something is the quantification introduced by a D operator. Hence the second consequence of our analysis for collectivity is that collectivity is not simply the absense of quantification, as it is widely taken to be.

6.1.3 All *and nonmaximality*

This proposal takes the "maximizing" effect of *all* to be crucially related to the nonmaximality of plural definites. To the extent that it is successful, then, it constitutes an argument that any theory of the meaning of *all* must also take seriously the phenomenon of nonmaximality. This criterion is met by, for example, Lasersohn's (1998) proposal, but not by Winter's (1998a,b).

6.1.4 Plurality and the lexicon

The distribution of *all* with collective predicates is one of the most difficult aspects of *all*'s grammar to account for. I have proposed here that *all*'s distribution can be

explained by the novel theory of the interaction of plurality, syntax, and the lexicon proposed in Chapter 4, section 4.4. Thus this account provides new evidence for the hypotheses about lexical structure (which are themsleves not novel) employed in that theory. In particular, as pointed out to me by Fred Landman (p.c.), this account provides evidence for the semantic and syntactic existence of the hypothesized predicate DO.

6.1.5 Economy

The "exceptions that prove the rule" from Chapter 4, section 4.4.3.2.2, show that we must suppose that insertion of the D operator is optional in the grammar, subject to other well-formedness constraints. They also show that one of the relevant well-formedness constraints is the principle of economy of derivation.

We argued in section 4.4 that the predicates of Taub's generalization do not (normally) license any hidden distributivity, which explains their incompatibility with *all*. But the fact that distributivity is permitted with these predicates under the right contextual circumstances shows that insertion of a D operator must be optionally permitted.

If the D operator is permitted with these predicates under some circumstances, then we have to say what rules it out in the other circumstances, the cases that constituted the core of Taub's generalization. For this we appealed to economy: the idea that a D operator is licensed only if it contributes something to the interpretation of the sentence. Hence we have another kind of phenomenon that shows economy to be an important principle of grammar.

6.1.6 The stranding hypothesis

I have argued against the stranding hypothesis of floating quantifiers (Sportiche 1988) both explicitly, and, in proposing that *all* is an adverb, implicitly. Many authors have adopted the stranding hypothesis and used it as a building block for a theory of other phenomena. To take just one example, Giusti (1990) argues from the stranding hypothesis and the syntactic distribution of *all* in German that German is not a nonconfigurational language. Indeed for scrambling languages the position of floating *all* is frequently taken to mark the base position of the DP that *all* takes as its argument. If the hypothesis I have proposed here is correct then the position of floated *all* is most likely not a valid diagnostic for the base position. Pursuing this would require a detailed reanalysis of floating quantifiers in other languages in light of the proposal I have made here, which could have interesting consequences.

Bibliography

Abney, S. 1987. The English Noun Phrase in its Sentential Aspect. Doctoral dissertation, MIT, Cambridge.

Asher, Nicholas. 1993. *Reference to Abstract Objects in Discourse.* Dordrecht: Kluwer.

Bach, Emmon. 1984. Some Generalizations of Categorial Grammars. In *Varieties of Formal Semantics*, ed. F.Landman and F. Veltman, Dordrecht: Foris.

Baltin, Mark. 1995. Floating Quantifiers, PRO, and Predication. *Linguistic Inquiry* 26:199-248.

Barwise, Jon and Robin Cooper. 1981. Generalized Quantifiers and Natural Language. *Linguistics and Philosophy* 4:159-219.

Beaver, David. 1996. An Infinite Number of Monkeys. In *Acta Linguistica Hungarica (Special edition devoded to Proceedings of the 5th Hungarian Symposium on Logic and Language).*

Bellert, Irena. 1977. On Semantic and Distributional Properties of Sentential Adverbs. *Lingusitic Inquiry* 8: 337-351.

Bennett M.R. 1975. Some extensions of a Montague fragment of English. Doctoral dissertation, UCLA, Los Angeles, CA.

Bittner, Maria. 1994. Cross Linguistic Semantics. *Linguistics and Philosophy* 17: 57-108.

Bresnan, Joan. 1973. Syntax of the Comparative Clause Coonstruction in English. *Linguistic Inquiry* 4:275-343.

Brisson, Christine. 1996. Distributivity, Asymmetry, and *Both*. In *NELS 26*, ed. J. Beckman, Amherst, MA: GLSA.

Brisson, Christine. 1997. On Definite Plural NPs and the Meaning of *all*. In *Semantics and Linguistic Theory VII*, ed. A. Lawson, Ithaca, NY: CLC Publications.

Chierchia, Gennaro . 1991. Functional Wh and Weak Crossover. In *WCCFL 10.*

Chierchia, Gennaro . 1993. Questions with Quantifiers. *Natural Language Semantics* 1:123-148.

Chierchia, Gennaro. 1995. The Plurality of Mass Nouns and the Notion of Semantic Parameter. ms, University of Milan.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and Grammar: An Introduction to Semantics*. Cambridge: MIT Press.

Cinque, Guglielmo. to appear. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press. (ms dated 1996)

Cole, Peter, ed. 1978. *Synatx and Semantics vol 9: Pragmatics.* New York: Academic Press.

Cooper, R. 1983. *Quantification and Syntactic Theory*. Dordrecht: Reidel.

Dalrymple, Mary, Makoto Kanazawa, Sam Mchombo, and Stanley Peters. 1994. What Do Reciprocals Mean? In *Semantics and Linguistic Theory IV*. Ithaca, NY: CLC Publications.

Dayal, Veneeta. 1992. Two types of universal terms in questions. In *NELS 22* Amherst: GLSA.

Dayal, Veneeta. 1995. Licensing *any* in non-negative/non-modal contexts. In *Semanitcs*

and Linguistic Theory V, ed. M. Simons and T. Galloway, Ithaca, NY: CLC Publications.

Dayal, Veneeta. 1996. *Locality in Wh-Quantification.* Dordrecht: Kluwer.

Dayal, Veneeta. to appear. *Any* as Inherently Modal *Linguistics and Philosophy.*

Déprez, Viviane. 1989. On the Typology of Syntactic Positions and the Nature of Chains. Doctoral dissertation, MIT, Cambridge, MA.

Does, Jaap van der. 1993. Sums and Quantifiers. *Linguistics and Philosophy* 16: 509-550.

Doetjes, Jenny. 1992. Floating Quantifiers Float to the Left. *The Linguistic Review*, 9-4:313-332.

Dougherty, Ray. 1970. A Grammar of Coördinate Conjoined Structures: I. *Language*, 46: 850-898.

Dowty, David. 1979. *Word Meaning and Montague Grammar.* Dordrecht: Reidel.

Dowty, David. 1987. Collective Predicates, Distributive Predicates, and *All*.In *Proceedings of the 3rd ESCOL,* ed. F. Marshall, Ohio State U.

Dowty, David. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67: 547-619.

Dowty, David and Belinda Brodie. 1984. The Semantics of 'Floated' Quantifiers in a Transformationless Grammar. In *Proceedings of WCCFL 3.*

Edmondson, J.A. 1978. On How to get *Both* in Categorial Grammar. *Studies in Language* 2.3: 295-312.

Engdahl, E. 1986. *Constituent Questions.* Dordrecht: Kluwer.

Ernst, Thomas. 1998. The scopal basis of adverb licensing. ms, Rutgers University.

Fiengo, Robert and Howard Lasnik. 1973. The logical structure of reciprocal sentences in English *Foundations of Language* 9:447-68.

Gillon, Brendan. 1987. The Readings of Plural Noun Phrases in English . *Linguistics and Philosophy* 13: 477-485.

Giusti, Giuliana. 1990. Floating Quantifiers, Scrambling, and Configurationality. *Linguistic Inquiry,* 22: 632-641.

Grimshaw, Jane. 1979. Complement Selection and the Lexicon . *Linguistic Inquiry* 10: 279-326.

Grimshaw, Jane and Sten Vikner. 1993. Obligatory Adjuncts and the Structure of Events. In *Knowledge and Language, Vol. II,* ed. E. Reuland and W. Abraham, Dordrecht: Kluwer.

Groenendijk, J. and Stokhof, M. 1983. Studies on the Semantics of Questions and the Pragmatics of Answers. Doctoral dissertation, University of Amsterdam.

Hale, Ken and Samuel Jay Keyser. 1987. Explaining and Constraining the English Middle. In *Studies in generative approaches to aspect,* Lexicon Project Working Papers 10, ed. Carol Tenny, Center for Cognitive Science, MIT.

Haspelmath, Martin. 1995. Diachronic Sources of *all* and *Every*. In *Quantification in Natural Language,* ed Bach, Partee, Jelinek. Dordrecht: Kluwer

Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral dissertation, University of Massachussetts, Amherst.

Heim, Irene. 1983. On the Projection Problem for Presuppositions. In *Proceedings of*

WCCFL 2.

Heim, Irene, Howard Lasnik, and Robert May. 1991. Reciprocity and Plurality. *Linguistic Inquiry* 22: 63-101.

Higginbotham, James . 1981. Reciprocal Interpretation. *Journal of Linguistic Research* 1:97-117.

Higginbotham, James. 1985. On Semantics. *Linguistic Inquiry* 16:547-593.

Horn, Laurence. 1986. Presupposition, Theme and Variations in *Proceedings of CLS 22; Papers from the Parasession on Pragmatics and Grammatical Theory*. p. 168-192.

Horn, Laurence. 1989. *A Natural History of Negation*. Chicago University Press, Chicago.

Horn, Laurence. 1991. Given as new: When redundant affirmation isn't. *Journal of Pragmatics* 15: 313-336.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Jackendoff, Ray. 1977. *X' Syntax*. Cambridge: MIT Press.

Johnson, Kyle. 1991. Object Positions. *Natural Language and Linguistic Theory* 9: 577-636.

Kadmon, N. and F. Landman. 1993. Any. *Linguistics and Philosophy* 16: 353-422.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic*. Dordrecht: Kluwer

Kayne, Richard. 1975. *French Syntax*. Cambridge: MIT Press.

Keenan, Edward and Leonard Faltz. 1985. *Boolean Semantics for Natural Language*. Dordrecht: Kluwer

Koopman, Hilda, and Dominique Sportiche. 1985. Theta theory and extraction. *GLOW Newsletter* 14:57-58.

Koopman, Hilda and Dominique Sportiche. 1991. The position of subjects. *Lingua* 85: 211-259.

Kratzer, Angelika. 1994. On External Arguments. In *UMOP Occasional Papers 17 - Functional Projections.* ed. E. Benedicio & J. Runner, eds, Amherst: GLSA.

Krifka, Manfred. 1996. Pragmatic Strengthening in Plural Predications and Donkey Sentences. In *Semantics and Linguistic Theory VI*, Ithaca, NY: CLC Publications.

Krifka, Manfred. 1992. Thematic relations as Links between Nominal Reference and Temporal Constitution. In *Lexical Matters*, ed. Ivan Sag and Anna Szabolsci, Stanford University, Palo Alto.

Kroch, Anthony. 1975. The Semantics of Scope of English. Doctoral dissertation, MIT, Cambridge, MA.

Ladusaw, William . 1982. Semantic Constraints on the English Partitive Construction. in *Proceedings of WCCFL 1*.

Landman, Fred . 1989. Groups, I. *Linguistics and Philosophy* 12: 559-605.

Landman, Fred . 1996. Plurality In *The Handbook of Contemporary Semantic Theory*, ed. S. Lappin, Cambridge: Blackwell.

Landman, Fred. 1997. Class notes, LSA summer institute; taken from *The Jerusalem Lectures*, forthcoming.

Langendoen, D. Terence . 1978. The Logic of Reciprocity *Linguistic Inquiry* 9:177-

197.
Larson, Richard. 1988. On the Double Object Construction. *Linguistic Inquiry* 19:335-391.
Lasersohn, Peter. 1990. *A Semantics for Groups and Events.* New York: Garland Publishing.
Lasersohn, Peter. 1993. Lexical Distributivity and Implicit Arguments. In *Semantics and Linguistic Theory III*, Ohio State University.
Lasersohn, Peter. 1995. *Plurality, Conjunction, and Events*. Dordrecht: Kluwer
Lasersohn, Peter. 1998a. Generalized Distributivity Operators *Linguistics and Philosophy* 21: 83-93.
Lasersohn, Peter. 1998b. Pragmatic Halos. ms, University of Illinois.
Levinson. 1983. *Pragmatics*. New York: Cambridge.
Lewis, David. 1975. Adverbs of Quantification. In *Formal Semantics of Natural Language*. ed E. Keenan. Cambridge: Cambridge University Press.
Lewis, David . 1979. Scorekeeping in a Language Game. *Journal of Philosophical Logic* 8: 339-359.
Link, Godehard . 1983. The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach. In *Meaning, Use, and Interpretation of Language*, ed. Bauerle et al, Berlin: DeGruyter.
Link, Godehard . 1984. Hydras. On the Logic of Relative Constructions with Multiple Heads. In *Varieties of Formal Semantics*. ed Landman and Veltman, Dordrecht: Foris.
Link, Godehard. 1987. Generalized Quantifiers and Plurals. In *Generalized Quantifiers,*ed P. Gardenfors, Dordrecht: Reidel.
Link, Godehard. 1994. Plural. English translation, 1984. In *Handbuch der Semantik*, ed Wunderlich & von Stechow, Berlin: de Gruyter.
Lønning, J.T. 1987. Collective Readings of Definite and Indefinite Noun Phrases. In *Generalized Quantifiers*, ed. P. Gardenfors, Dordrecht: Reidel.
Maling, Joan. 1976. Notes on Quantifier Postposing. *Linguistic Inquiry* 7: 708-718.
Massey, G. 1976. Tom, Dick, and Harry, and All the King's Men. *American Philosophical Quarterly* 13:89-107.
May, Robert. 1985. *Logical Form: Its Structure and Derivation*. Cambridge: MIT Press.
McClure, William. 1994. Syntactic Projections of the Semantics of Aspect. Doctoral dissertation, Cornell, Ithaca, NY.
McConnell-Ginet, Sally. 1982. Adverbs and Logical Form: A Linguistically Realistic Theory. *Language* 58: 144-184.
Merchant, Jason. 1996. Scrambling and Quantifier Float in German. In *NELS 26* GLSA, Amherst.
Mittwoch, Anita. 1982. On the difference between *Eating* and *Eating Something*: Activities versus Accomplishments. *Linguistic Inquiry* 12: 113-122.
Moltmann, Friederike. 1997. *Parts and Wholes in Semantics*. Oxford: Oxford University Press.
Parsons, Terence. 1990. *Events in the Semantics of English*. MIT Press, Cambridge.

Partee, Barbara. 1995. Quantificational Structures and Compositionality. In *Quantification in Natural Language*, ed Bach et al. Dordrecht: Kluwer.

Partee, Barbara and Mats Rooth. 1983. Generalized Conjunction and Type Ambiguity. In *Meaning, Use, and Interpretation of Language*, ed. Bauerle et al, Berlin: DeGruyter.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20: 365-424.

Postal, Paul. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications*. Cambridge: MIT Press.

Pritchett, Bradley L.. 1990.. A note on Scope Interaction with Definite Plural NPs . *Linguistic Inquiry* 21:646-654.

Pustjovski, James . 1991. The syntax of event structure. *Cognition* 41: 47-81.

Quine, W.V.O. 1980. *Set Theory and its Logic*. Cambridge: Revised Edition, Harvard University Press.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge: MIT Press.

Roberts, Craige. 1987. Modal Subordination, Anaphora, and Distributivity. Doctoral dissertation, University of Massachusetts at Amherst

Rooth, Mats. 1985. Association with Focus. Doctoral dissertation, University of Massachussetts, Amherst.

Ross, J.R. 1970. Act. in *Semantics of Natural Language.* ed. D. Davidson, G. Harman, Dordrecht: Reidel.

Sadock, J. . 1978. On testing for conversational implicature. in P. Cole ed..

Safir, Ken and Tim Stowell. 1989. Binominal *each.* In *NELS 18*, Amherst: GLSA.

Schein, Barry. 1993. *Plurals and Events*. Cambridge: MIT Press.

Schwarzschild, Roger. 1991. On the Meaning of Definite Plural Noun Phrases. Doctoral dissertation, University of Massachusetts at Amherst

Schwarzschild, Roger. 1992a. Types of Plural Individuals. *Linguistics and Philosophy* 15: 641-675.

Schwarzschild, Roger. 1992b. *Together* as a Non-Distributivity Marker. In *Proceedings of the 8th Amsterdam Colloquium*. ed. P. Dekker & M. Stokhof, Institute for Logic, Language, and Computation, University of Amsterdam.

Schwarzschild, Roger. 1992c. Distinguishing oriented-adverbials from secondary predicates in the analysis of *together* Talk presented at IATL 1992, Bar-Ilan University.

Schwarzschild, Roger. 1994. Plurals, Presuppositions, and the Sources of Distributivity. *Natural Language Semantics* 2: 201-248.

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer

Selkirk, Lisa. 1977. Some Remarks on Noun Phrase Structure. In *Studies in Formal Syntax*, ed. A. Akmajian, P. Culicover, and T. Wasow, New York: Academic Press.

Sharvy, Richard. 1980. A More General Theory of Definite Descriptions. *The Philosophical Review* 89: 607-624.

Shlonsky, Ur. 1991. Quantifier Phrases and Quantifier Float. In *NELS 21*, Amherst: GLSA.

Speas, Margaret. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer

Sportiche, Dominique. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19: 425-449.

Sportiche, Dominique. 1992. Clitic Constructions. Talk presented at GLOW/ms, UCLA, Los Angeles.

Stalnaker, Robert . 1973. Presuppositions. *Journal of Philosophical Logic* 2: 447-457.

Stalnaker, Robert. 1978. Assertion. In P. Cole, ed..

Stockwell, Robert P., Paul Schacter and Barbara Partee. 1973. *The Major Syntactic Structures of English*. New York: Holt, Reinhart, and Winston.

Strawson, P.F. 1950. On Referring. *Mind* 61: 480-96.

Taub, Alison. 1989. Collective Predicates, Aktionsarten and *All*. In *Papers on Quantification*, ed. Bach, Kratzer, Partee. University of Massachusetts at Amherst.

Tenny, Carol. 1987. *Grammaticalizing Aspect and Affectedness*. Doctoral dissertation, MIT, Cambridge, MA.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Cornell University Press, Ithaca.

von Fintel, Kai . 1994. Restrictions on Quantifier Domains. Doctoral dissertation, University of Massachusetts at Amherst.

Westerståhl, Dag. 1984. Determiners and Context Sets. In *Generalized Quantifiers in Natural Language*, ed. Johan van Bentham and Alice ter Meulen, Dordrecht: Foris.

Wilkinson, Karina. 1991. Studies in the Semantics of Generic Noun Phrases. Doctoral dissertation, University of Massachusetts at Amherst.

Williams, Edwin. 1982. The NP Cycle. *Linguistic Inquiry* 13:277-295

Williams, Edwin. 1992. Reciprocal Scope. *Linguistic Inquiry* 22: 159-173.

Winter, Yoad. 1998a. Atom Predicates and Set Predicates: Towards a General Theory of Plural Quantification. To appear in *Semantics and Linguistic Theory VIII*.

Winter, Yoad. 1998b. Doctoral dissertation (in progress), UIL/OTS, Utrecht.

Yoon, Y. 1996. Total and Partial Predicates and the weak and strong interpretations. *Natural Language Semantics* 4:217-236.