

PLURALS, *ALL*, AND THE NONUNIFORMITY OF COLLECTIVE PREDICATION*

1. INTRODUCTION

In the literature on plural noun phrases it is widely observed that the sentence in (1) implies something like (2).

- (1) The girls jumped in the lake.
- (2) Every girl jumped in the lake.

This leads many authors to posit the existence of a “D operator” (see Link (1983), Landman (1989, 1996), Lasersohn (1995, 1998), Schwarzschild (1996) for several different versions, the differences between which do not concern us here). The purported implication from (1) to (2) is captured by an implicit distributivity operator on the VP (a D operator), which introduces universal quantification over the individual girls in the denotation of *the girls*, as shown in (3)–(4) (for some alternative formulations of this idea, see Lasersohn (1998)).

- (3) $D_{\text{jumped.in.the.lake'}}(\text{the.girls})$
- (4) $\forall x[x \in \llbracket \text{the.girls}' \rrbracket \rightarrow x \in \llbracket \text{jumped.in.the.lake}' \rrbracket]$

However, Landman (1989, 1996) and others (e.g., Lasersohn 1995) observe that (1) is not completely synonymous with (2). The difference between the sentences is that (2) requires that each and every girl jumped,

* The material in this paper is based in large measure on my dissertation, completed at Rutgers University in 1998. I must thank the Linguistics Department there for their nurturance and support; special thanks to Veneeta Dayal and Roger Schwarzschild. Thanks also to Viviane Déprez, Angelika Kratzer, Manfred Krifka, Fred Landman, Peter Lasersohn, Yael Sharvit, and audiences at SALT VII, the University of Connecticut, and the University of Delaware. I am also grateful to two anonymous reviewers who helped improve this paper with their comments. Errors that remain are my responsibility alone.



while (1) is weaker, allowing for exceptions.¹ Let us call this effect ‘nonmaximality’ (adapting a term from Dowty 1987). Because of nonmaximality with definite plural NPs, Landman proposes that (1) is interpreted simply as predication of the plural individual *the girls* of the predicate *jumped in the lake*. I’ll call this kind of approach to plurals (and nonmaximality) the *groups approach*. There are various instantiations of the groups approach, but the basic idea is that in (1), *the implication that every girl (or at least most of them) jumped comes from a meaning postulate*.² While this idea is appealing in its simplicity, it is, unfortunately, empirically inadequate, as we will see shortly.

In this paper I will propose an alternative to the groups approach to nonmaximality. I will argue that nonmaximality is due, not to the absence of a quantifier, but to the presence of a quantifier and a specific kind of domain-of-quantification effect. The proposal makes crucial use of Schwarzschild’s (1996) idea that a D operator has a domain variable in its restriction whose value is contextually specified.

A related issue is the meaning of *all*, which appears to undo the nonmaximality of definite plurals. Speakers might allow exceptions to (5), but they will not do so for (6).

(5) The girls jumped in the lake.

(6) The girls all jumped in the lake.

As a first pass we might assume simply that *all* is a universal quantifier, and that’s why it has this “maximizing effect”. But something else must be said, because unlike the universal quantifiers *every* and *each*, *all* can combine with collective predicates, as shown in (7)–(12).

(7) The girls gathered in the hallway.

(8) All the girls gathered in the hallway.

¹ In an earlier paper, Scha (1981) also observed that universal quantification is too strong to capture the meanings of sentences with definite plurals, although he was looking at different kinds of examples.

² Here I do not mean “groups” in the technical sense of, say, Landman (1989) and Lasnik (1995), although those theories do represent a particular way of formalizing the idea I have in mind here. While it would perhaps be better to use a different term, it seems to me that there is no other term available. In the large body of work on plurals, most of the terms that could be used to refer to this way of looking at plurals, including *plurals*, *sums*, *pluralities*, *bunches*, and *groups* have already been used in a very specific way by one author or another.

- (9) *Every girl gathered in the hallway.
- (10) The girls built a raft. (*ok on collective reading*)
- (11) All the girls built a raft. (*ok on collective reading*)
- (12) Every girl built a raft. (*distributive only*)

Observe that while (7) allows for exceptions, (8) means that every girl participated in the gathering. In other words, when *all* does combine with collective predicates, it has the same “maximizing effect” that it has with distributive predicates.

One of the difficulties in accounting for the maximizing effect of *all* with collective predicates comes from the fact that collectivity is widely assumed to be simply the absence of distributivity. In this paper I will argue that this is incorrect. We will see that collective predicates do not comprise a uniform class and that a subclass of the collective predicates actually involves quantification. This will in turn explain why some collective predicates show the same sort of nonmaximality that we observe with distributive predicates, and why *all* is compatible with those predicates.

The paper will proceed in essentially two parts. In the first part, sections 1–3, I will develop a way to capture nonmaximality in the semantics of definite plurals, and propose a meaning for *all* that interacts crucially with the theory proposed for capturing nonmaximality. In the second part, sections 4–6, I will spend some time developing a framework of event semantics and an analysis of collective predicates of certain *aktionsart* classes. As we go deeper into the paper, the connection between the first part and the second part will emerge: the analysis of distributivity and *all*, and the analysis of collective predicates, when taken together, accurately predict a wide range of subtle data related to collective predication and *all*. In addition, the analysis will be shown to make the right predictions for some new data that will be introduced along the way.

2. THE D OPERATOR ACCOUNT AND SOME OF ITS PROBLEMS

The D-operator account of distributivity adopted by the authors cited in the introduction captures several linguistic phenomena related to definite plurals. For example, it captures the intuition that predication of a definite plural is synonymous with universal quantification over the individuals that make up the plural. The examples in (13)–(15) are taken from Higginbotham (1981), Link (1983), and Yoon (1996), respectively.

- (13) The men are left-handed.

(14) The pigs died.

(15) 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. Of course we have already seen that it is questionable whether these sentences are really strictly universal but for now let us accept this judgment.

The D-operator approach to distributivity also accounts for the “many sandwiches” reading of (16).

(16) The girls ate a sandwich.

(16) has (at least) two possible interpretations. One is a “collective” reading: it is true in the case where the girls shared a single (large) sandwich. The other is a “distributive” reading: it is true in the case where each girl ate a separate sandwich. The latter reading gives us an argument for the D operator. If the D operator introduces a universal quantifier, then it can take wide scope over the existential quantifier associated with the singular indefinite *a sandwich*, making it possible to understand that many sandwiches were eaten.

However, the many-sandwiches reading of (16) also exposes a problem with the D-operator approach to distributivity. As we observed in the introduction with the predicate *jumped in the lake*, it is not clear that (16) strictly speaking requires that each and every girl ate a sandwich in order to be true; many speakers will judge (16) to be true even if one or two out of a large group of girls did not eat a sandwich.

The problem is that our semantics cannot generate the many-sandwiches reading and also allow for exceptions. On the one hand, we need universal quantification, but on the other hand, it is too strong. We need some kind of not-quite-universal quantification.

The groups approach to nonmaximality endorsed by Landman and Lasnik does not fare any better. It predicts that the nonmaximal reading of the sentence in (16) should be possible only on its collective reading; on a distributive reading, the groups approach predicts that we should have true universal quantification.

The fact that the quantification associated with definite plurals is somehow weaker than universal quantification has already been widely observed in the literature on reciprocals (Fiengo and Lasnik 1973; Langendoen 1978; Dalrymple et al. 1998), and more recently in the literature on donkey sentences.

For example, the sentences in (17)–(18) are taken from Fiengo and Lasnik.

- (17) The men are hitting each other.
- (18) Each of the men is hitting the other(s).

Fiengo and Lasnik observe that in a situation where the *men* refers to a group of two men, (17) and (18) are synonymous. However, if the group of men is a bit larger, say 7 or 8 men, then (17) can be judged true even if (18) is false. It seems that (17) can be judged true even if there are some among *the men* who are not hitting anyone at all. This indicates that the sentence in (17), at least in certain circumstances, allows a “weak” reading of the reciprocal. On this weak reading, the sentence can be true even if it is not the case that every man is hitting every other man in the group.

Williams (1991), in a reply to Heim et al. (1991), expands on this point to argue for a distinction between “strong distributivity” and “weak distributivity”, even in the absence of reciprocals. He cites Fiengo and May’s examples and adds the sentence in (19):

- (19) The men were hitting Bill.

Just as (17) can describe a “general melee” which allows there to be non-hitters among those referred to by *the men*, the same is true for (19), Williams says: “It is compatible with a situation in which there were some nonhitters” (p. 162). Williams’ “weak distributivity” is what I am calling nonmaximality.

In a more recent paper, Yoon (1996) proposes that taking into account the phenomenon of nonmaximality of definite plurals offers a solution to the proportion problem in donkey sentences. The proportion problem is illustrated by the sentences in (20)–(21).

- (20) Most boys who had a baseball card in their pockets soiled it while playing in the mud.
- (21) 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 (20) 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, (21) 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.

Like Williams with reciprocals, Yoon argues that there is a parallel between the weak reading of a donkey sentence and the weak readings of sentences with definite plurals. She gives the examples in (22)–(23).

(22) The glasses are dirty.

(23) The glasses are clean.

In a situation where someone is setting a table for a formal dinner, she argues, (22) would count as true even if only a few of the glasses are dirty; but for (23) to be true it seems we would say that all the glasses have to be clean.

Yoon proposes that nonmaximality is a lexically-specified property of a relatively small class of predicates, such as *dirty* and *open*, which are part of an antonym pair (e.g., *dirty/clean*, *open/closed*). However, we have already seen that phrasal predicates without antonyms (*ate a sandwich*) can also show nonmaximality. Since Yoon’s analysis of nonmaximality is too narrow, I will not adopt it here. What her paper adds to the discussion is to show that the proportion problem is another phenomenon that is related to the nonmaximality associated with definite plurals.

Even more recent is a proposal by Lasersohn (1999) in which the non-maximality of definite plurals is proposed to be part of a more general phenomenon that he calls “pragmatic slack”. The idea is that speakers do not require one another to be completely truthful and accurate all the time, and that nonmaximality can be traced to a willingness by speakers to grant each other a certain amount of pragmatic slack.

He develops a formal mechanism called a “pragmatic halo” that gives speakers a set of objects that are “sufficiently close” to the denotation of a natural language expression like *the girls*. The halo of *the girls*, for example, is a set of sets of individuals which are very similar to the set denoted by *the girls*, but which differ from that set in minor ways, such as leaving out one or two individuals who are irrelevant or otherwise pragmatically ignorable. If a sentence is not literally true but true when we replace an expression with something from its pragmatic halo, a hearer may be willing to accept the sentence as true. Lasersohn’s proposal, unlike Yoon’s, treats nonmaximality as a very general phenomenon and so it is closer to what we need here. However, while it is even more general than the proposal I will make here (Lasersohn does not limit himself to the issue of definite plurals and *all*), it does not account for some of the subtle phenomena related to collective predicates that we will discuss in the second half of this paper.

3. CAPTURING NONMAXIMALITY

To capture nonmaximality I will propose an adaptation to the theory of distributivity of Schwarzschild (1996). The crucial property of this theory, which makes it different from every other theory of distributivity that I know of, is the presence of a domain variable in the restriction of the quantifier. I will not review here the evidence for this approach; for that see Schwarzschild (1996).

Schwarzschild proposes that the D operator (which he calls *Part*, for partition; I refrain from adopting this name for consistency with other literature) is always accompanied by a **context-dependent domain selection variable**. He calls this variable *Cov*, because the value assigned to the variable always takes the form of a cover of the universe of discourse. (For an earlier use of covers to account for properties of plurals, see Gillon (1987)). The definition of a cover is given in (24).

- (24) X covers Y iff:
- a. X is a set of nonempty subsets of Y
 - b. $\forall y \in Y \exists x \in X[y \in x]$

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 Westerstähl (1985) and von Stechow (1994) on quantifiers and their domain restrictions).

When a D operator is present on the predicate *jumped in the lake*, as in (25), the sentence will be interpreted as in (26). (Because we always interpret expressions relative to a model *M* and an assignment function *g*, I will leave off these superscripts.)

- (25) The girls ^{D_i}jumped in the lake.

- (26) $\forall x[x \in \llbracket \text{Cov}_i \rrbracket \ \& \ x \subseteq \llbracket \text{the.girls}' \rrbracket \rightarrow x \in \llbracket \text{jumped.in.the.lake}' \rrbracket]$

Note that the universal quantifier contributed to (26) by the D operator has two conditions in its restriction: *x* must be a subset of *the girls* (standard D operators require that *x* be an element of $\llbracket \text{the.girls}' \rrbracket$) and *x* must be an element of the cover assigned to *Cov_i*.

In order to evaluate the truth conditions of (26), we must assign a value to *Cov_i*, 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 (27).

- (27) $U = \{a, b, c, s, t, \{a,b\}, \{a,c\}, \{a,s\}, \{a,t\}, \{a,s,t\} \dots\}$
 $\llbracket \text{the.girls}' \rrbracket = \{a,b,c\}$
 $J = \{\{a\}, \{b\}, \{c\}, \{s,t\}\}$
 $K = \{\{a\}, \{c\}, \{b,s,t\}\}$

Suppose the context assigns the value J to Cov_i in (26). Then the sentence will be interpreted just as if we had used a “standard” D operator, because since each girl occupies a singleton set of the cover assigned to Cov_i , each girl is asserted to be in the extension of *jumped.in.the.lake*. 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 J to Cov_i in (26) leads to the same interpretation that any ordinary D operator, without Cov_i in its restriction, would yield.³

There is another possibility, one which is very important to us here. Because the value of Cov is a cover of the entire domain of discourse, that means it is possible to assign a value to Cov that is oddly-shaped, or “ill-fitting” with respect to the subject DP. This can be illustrated using the same sentence from above, with a different value for Cov_i .

- (28) The girls jumped in the lake.
- (29) $\forall x[x \in \llbracket \text{Cov}_i \rrbracket \ \& \ x \subseteq \llbracket \text{the.girls}' \rrbracket \rightarrow x \in \llbracket \text{jumped.in.the.lake}' \rrbracket]$
 $U = \{a, b, c, s, t, \{a,b\}, \{a,c\}, \{a,s\}, \{a,t\}, \{a,s,t\} \dots\}$
 $\llbracket \text{the.girls}' \rrbracket = \{a,b,c\}$
 $J = \{\{a\}, \{b\}, \{c\}, \{s,t\}\}$
 $K = \{\{a\}, \{c\}, \{b,s,t\}\}$

Let’s consider the possibility of assigning the cover K to Cov_i in (29). The difference between K and J (the value we assigned earlier) is that in K Betty does not occupy a singleton cell: she is in a cell with the two non-girls, Stan and Tim. Call this an ill-fitting cover, because it is ill-fitting with

³ 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 *jumped.in.the.lake* 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 girls in the extension of *jumped in the lake*, only individual girls).

respect to the set of girls – there is no set of cells whose union is equivalent to the set of girls.

A consequence of assigning this type of cover to Cov_i in (29) is that the semantics in some sense ‘doesn’t care’ whether Betty jumped in the lake or not. Since the set $\{b, s, t\}$ is not a subset of the set $\{a, b, c\}$, there is no cell containing Betty that satisfies the restriction of the quantifier. The sentence can come out true whether Betty jumped in the lake or not.

This possibility was pointed out by Lasersohn (1995), who saw it as a problem with Schwarzschild’s proposal. His objection was that if the semantics and pragmatics work in the way just described, it means a sentence like *John and Mary went to school* can come out true in a situation where Mary didn’t go to school. Schwarzschild’s response was that such a choice of cover would be so uncooperative of a speaker as to be “pathological”. He shows that it is possible to formulate a rule that prevents the choice of such a cover, but argues that it is more plausible to simply assume that pathological covers are ruled out by pragmatic principles (1996: 77, referring to Schwarzschild 1994: 228–233).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 the phenomenon of nonmaximality shows.

In fact, it is quite clear that there are circumstances in which both speaker and hearer share the assumption that one (or more) individuals who are part of the denotation of the definite plural is (are) excluded from the domain of the D operator. Imagine we are counselors at a girls’ summer camp, and suppose that we are talking about the girls of cabin number six. At lunchtime, each girl was presented with a tray including a sandwich, milk, an apple, and a cookie. We can say *the girls ate a sandwich* even if we both know that Betty never eats her sandwich, and she just picked at the cookie and the apple.

Suppose there is a cell of the cover where we keep all the things that may be salient in the discourse as a whole, but are excluded from the present consideration. Let us call this cell the “junkpile”. Speakers can agree to put Betty in the junkpile if they both know there is good reason to ignore her.

However, I do not assume that nonmaximality arises only in cases where both speaker and hearer agree that individuals may be excluded

⁴ To the extent that he discusses nonmaximality, Schwarzschild adopts the general approach that I discussed earlier as the groups approach.

from the domain of the D operator. Rather, I take it that it is a very general phenomenon because speakers and hearers are constantly leaving room to accommodate the others' potentially ill-fitting cover. This is true not only for Cov, but for contextually-determined domain variables in general: speakers and hearers are always engaged in a process of guessing about what each other have in mind. (For more discussion on this point, see von Fintel (1994).) In fact, I assume that in the case of plurals and distributivity, this guessing is largely responsible for the phenomenon of nonmaximality. It's not that speakers and hearers always have in mind ill-fitting covers; it's that speakers and hearers must always make room for the *possibility* of ill-fitting covers.

A theory of distributivity like the one I have outlined gives us a simple account of nonmaximality, in which nonmaximality is just a domain-of-quantification effect. With this proposal, we capture the intuition that distributive quantification is not-quite-universal, as Landman, Williams, and others have observed.⁵

However, in addition to Lasersohn's objection to Schwarzschild, there is another objection to the idea that Cov ranges over the entire domain of discourse (due to Manfred Krifka, p.c.). This argument says, if we are talking about the girls jumping in the lake, why should Cov also include the boys further down who stayed on shore, and the geese that might be flying overhead, and so on? Doesn't this force us to make odd assumptions about pragmatics?

However, I know of no theory of pragmatics that prevents these assumptions. In addition, I take it that a good argument for including Cov as the domain variable for the D operator comes from the body of recent work on natural language quantification, including the work by von Fintel previously cited. Von Fintel argues that natural language quantifiers always come with resource domain variables whose values are fixed by the context of utterance. If this is true, then the null hypothesis for the D operator would be that it, too, should have a context-sensitive variable in its domain.

This line of reasoning, however, still does not require us to adopt Cov as that variable (although to my knowledge there is no other proposal for what the domain variable of the D operator should look like). And it also does not require us to suppose that Cov always ranges over the entire domain of discourse (although, again, resource domain variables generally *are* taken to range over the entire domain of discourse).

⁵ This also means that D operators (with their context-sensitive domain restrictions) appear on any predicate that is interpreted distributively, not just on predicates that are ambiguous between a distributive and a collective reading, as some authors (e.g. Lasersohn 1995) have proposed.

The upshot is that we do not have good empirical evidence either for or against the idea that Cov is a cover of the entire domain of discourse.⁶ The evidence for this idea will have to come from the use I make of it in this proposal. If the reader deems the arguments presented here for a theory that accounts for nonmaximality to be convincing, then we will have at least one good reason to suppose that Cov ranges over the whole domain.

The theory of nonmaximality I have proposed here offers a simple, straightforward solution to the “many sandwiches” problem. It gives us a quantifier with not-quite-universal force that can take scope over an existentially-quantified object. Recall our example sentence, repeated below, with its interpretation in (31), assuming Cov_i has the value in (32).

(30) The girls ate a sandwich.

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

(32) $\{\{a\}, \{c\}, \{b,s,t\}\}$

This sentence is 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 need: a universal quantifier ensures that we have a one-sandwich-per-girl reading, but its quantification over the set

⁶ It is worth pointing out that Schwarzschild did give an argument for this from VP ellipsis:

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:

(i) The computers were paid for in two installments and the diskettes were too. (1996, p. 76)

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. This means that the Cov associated with the predicate *be-paid-for-in-two-installments* must contain singleton cells corresponding to each computer, and a nonsingleton cell corresponding to the set of diskettes. The most straightforward way to capture this is to suppose that in fact Cov is a cover of the entire universe of discourse.

However, Manfred Krifka points out (p.c.) that since D adjoins, modifier-like, to VP, perhaps we could just copy the VP without the D operator, and adjoin a new D operator with a new index (and hence a new Cov variable) to the new copy of the VP. Therefore, it’s not clear that the D operator must be copied with the VP, as would be required for Schwarzschild’s argument to go through.

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.

It has been suggested to me by several people, including an anonymous reviewer, that perhaps nonmaximality is really just a problem of fixing the precise reference of *the girls* in a sentence like (30). Perhaps when the speaker uttered the phrase *the girls* she didn’t really mean Alice, Betty, and Carmen, but just Alice and Carmen. But this hypothesis won’t explain more complex examples with VP or sentence conjunction, as shown below.

(33) The girls built a raft together, and then took a nap.

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

If we tried to capture nonmaximality by attributing it to a problem in fixing the reference of the DP, we would predict that the actual participants in the raft building should be the same as the actual participants in the nap-taking, because both VP’s are predicated of the same DP. But the crucial observation here is that nonmaximality doesn’t have to apply the same way to both conjuncts: the actual participants in the raft-building do not have to be the same as in the nap-taking; likewise for the actual participants in basketball and in swimming.

This suggests that the phenomenon of nonmaximality is independent of the process of identifying the referent of a plural NP. It supports the hypothesis I’ve proposed here, which locates the source of nonmaximality in the VP – more specifically, in the D operator, with its context-sensitive domain variable. Finally, this example also shows that each D operator in a discourse or a sentence has its own Cov, which may have a different value than other Covs associated with other D operators. This is consistent with Schwarzschild’s original formulation, in which each D operator is indexed to a particular Cov variable.

4. THE SEMANTICS OF *all*

This analysis of nonmaximality makes possible a new account of the meaning of *all* which will explain why *all* is compatible with collective predicates, unlike the universal quantifiers *each* and *every*. The first observation in this direction is that *all*, unlike *every* or *each*, always combines with a definite plural, rather than a common noun, in episodic sentences. This is shown in (35)–(36).

(35) All the girls went to the gym yesterday.

- (36) *All girls went to the gym yesterday.

Since distributive quantification is associated with definite plurals, the fact that *all* occurs with definite plurals rather than common nouns⁷ is one piece of evidence that the meaning of *all* is related to distributivity.

Consistent with this fact, I propose that *all* is not a determiner-quantifier but rather interacts with the quantification introduced by the D operator to rule out the nonmaximality that a D operator normally allows. Recall our observation about the difference between sentences like (37) and (38).

- (37) The girls jumped in the lake.

- (38) The girls all jumped in the lake.

The authors who have written on the meaning of *all* have commented that *all* seems to have a “maximizing effect” (Dowty 1987) or a “totality effect” (Link 1983).

Let us take this intuitive characterization seriously. I propose that *all*’s contribution to (38) is to rule out the possibility of nonmaximality.⁸ Since the cause of nonmaximality 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 (39).

- (39) **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 $\forall y[y \in X \rightarrow \exists Z[Z \in \text{Cov} \ \& \ y \in Z \ \& \ Z \subseteq X]]$

⁷ Note that when *all* combines with a common noun, if the sentence is grammatical it is interpreted generically, as observed by Partee (1995).

- (i) All doctors wear white coats.

This suggests that *all* can only combine with referring expressions, including definites and kinds.

⁸ This is very similar in spirit, though quite different in execution, to Lasnik’s (1999) approach, discussed briefly in section 1.

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. In order to keep this paper simpler I will stipulate that floated *all* is always construed with the subject DP, and to mark this in our examples I will mark DP's associated with *all* with superscript ^{gf}. (For a more explicit account of the syntax, see Brisson (1998).) Of course, a prenominal occurrence of *all* is associated with its sister DP.

All has its maximizing effect by requiring that the value assigned to Cov be a good fit with the denotation of the DP it is associated with. Because covers are a feature of the context of utterance, this means that *all*'s contribution is not a component of truth-conditional meaning, but something more like presupposition, or a focus-sensitive operator. I'll call *all*'s contribution to meaning the "domain-adjusting meaning".⁹ The idea is that the presence of domain-adjusting meaning is triggered by the presence of *all* and is derived alongside the ordinary meaning of a sentence, similarly to the way the focus denotation and the ordinary denotation are derived in the structured meaning approach to focus (see Rooth 1998).

Here is a sketch of how this might work. First we have a translation rule for *all*:

(40) ***translation rule for all:***

all has no ordinary translation, and a domain-adjusting meaning of $[\lambda xgf(\text{Cov})(x)]$

The domain-adjusting meaning of *all* is written inside the characters [] to orthographically mark that good fit is not evaluable as part of the

⁹ I am sidestepping the question of just how closely related the notion of domain-adjusting meaning is to presupposition. 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 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 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 from calling the good fit requirement a presupposition because it would raise numerous new questions. For one thing, as 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, and I don't see any evidence that it does, although this might be due to the fact that its effect is extremely local.

Note also that the idea of "domain adjusting" is similar to the proposal by Kadmon and Landman (1993) for *any*, in which *any* lacks quantificational force of its own but has a "domain widening" effect on the domain of another quantifier present in the sentence.

truth conditions of the sentence, but interacts with the context to limit the possible choices of Cov. It is then possible to write a series of rules for combining regular meanings with domain adjusting meanings. For example, the translation of *all the girls* would be :

$$(41) \quad \text{the.girls}', [\text{gf}(\text{Cov})(\text{the.girls}')]]$$

and the translation of *all the girls left* would be:

$$(42) \quad \text{left}'(\text{the.girls}'), [\text{gf}(\text{Cov}_i)(\text{the.girls}')]]$$

For reasons of space I will not give all of the rules to compose domain-adjusting meanings alongside ordinary meanings. The mechanical details are worked out in Brisson (1998). For the rest of the discussion I will simply use the superscript notation or the $[\]$ to indicate when the domain-adjusting meaning of *all* is imposing the good fit requirement.

Let's return to our sample universe and sample values of Cov_i to see how this works. Consider the sentence in (43), with its translation in (44) and interpretation in (45).

$$(43) \quad \text{All the girls jumped in the lake.}$$

$$(44) \quad \text{D}_{\text{j.i.t.l}}'(\text{the.girls}'), [\text{gf}(\text{Cov}_i)(\text{the.girls}')]]$$

$$(45) \quad \forall x[x \in [\text{Cov}_i] \ \& \ x \subseteq [\text{the.girls}']] \rightarrow x \in \llbracket \text{jumped.in.the.lake}' \rrbracket, [\text{gf}(\text{Cov}_i)(\text{the.girls}')]]$$

$$U = \{a, b, c, s, t, \{a, b\}, \{a, c\}, \{a, s\}, \{a, t\}, \{a, s, t\}, \dots\}$$

$$\llbracket \text{the.girls}' \rrbracket = \{a, b, c\}$$

$$J = \{\{a\}, \{b\}, \{c\}, \{s, t\}\}$$

$$K = \{\{a\}, \{c\}, \{b, s, t\}\}$$

$$L = \{\{a\}, \{b\}, \{c, s, t\}\}$$

If *all* interacts with the context to limit the set of possible values for Cov_i in the way I have described, then K and L (and other ill-fitting covers) are not possible values for Cov_i in (45). They have been eliminated as possibilities by the presence of *all*. This leaves J as the only possibility. (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 J will be assigned as the value of Cov_i in (45). Hence the D operator will have true universal quantificational force over the set denoted by *the girls*.

This explains the maximizing effect of *all*, and hence the difference between (37) and (38). (37) can be true, in some contexts, if a speaker is willing to allow assignment of an ill-fitting value to Cov_i . But (38) will not be true unless every girl jumped in the lake because the presence of *all* ensures that an ill-fitting cover cannot be assigned to Cov_i .

At this point, the theory of what *all* does works nicely with the theory of nonmaximality that I have proposed, but it isn't a crucial component of it. So far, it doesn't improve on the empirical coverage of a theory that would treat *all* like a determiner-quantifier, and one could adopt the theory of nonmaximality that I have proposed without adopting the theory about *all*. However, in the next sections we will see that the theories of nonmaximality and *all* proposed here contribute to a theory of collective predication in which we will predict a range of facts about the interaction of *all* and collective predicates that is unexpected if we treat *all* simply as a universal quantifier.

5. *All* AND COLLECTIVE PREDICATES

I begin this discussion by giving a definition of "collective predicate".

- (46) A *collective predicate* is one that requires a plural argument; or a predicate that with a plural argument has an interpretation that is different from the interpretation that comes from universally quantifying over the parts of the plural argument.

This definition includes many subclasses of predicates, including a class I'll call collective action, which we have already seen instances of:

- (47) The girls gathered in the hallway.
- (48) *Jane gathered in the hallway.
- (49) The girls built a raft.
- (50) Every girl built a raft. (*only has 1 raft per girl reading*)

The definition also includes the class of reciprocal predicates, including "lexcially reciprocal" predicates like *be alike*, and predicates that are made reciprocal by the presence of the anaphor *each other* in their object position.

- (51) The dogs are alike.

(52) *Rex is alike.

(53) The boys recognized each other.

(54) *John recognized each other.

It also includes predicates that are made reciprocal by the addition of collectivizing adverbials such as *together*.

(55) The planes landed together/as a group/in formation.

(56) *The DC-10 landed together/as a group/in formation.

As we saw in the introduction, *all*, unlike *every*, is generally compatible with collective predicates, as in (57)–(58).

(57) The girls all gathered in the hallway.

(58) The campers all built a raft.

However, Dowty (1987) observed that collective predicates are not uniform in this respect. Some collective predicates do not allow *all*, as in (59)–(60).

(59) *The teachers are all a group of four.

(60) *The children are all a big group.

In his discussion of collectives and *all*, Dowty did not look at reciprocal predicates and predicates that have been “collectivized” by adverbial phrases. We will see later that whether or not this was an intentional omission, it is a useful way to proceed and will be justified by the facts. For now, though, to keep our discussion of the problem simpler, we will restrict our attention (as did Dowty) to collective predicates that don’t have collectivizing adverbials and are not interpreted reciprocally.

Looking only at the predicates that I have called collective action predicates, and predicates like *be a group of four* or *be a big group*, Dowty offered an explanation for the difference between collective predicates that allow *all* and those that don’t. He suggested that the former have “distributive subentailments” and the latter don’t. He doesn’t provide a definition of distributive subentailments, 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 is that *all* is a universal quantifier that distributes this subentailment down to every individual in the denotation of the subject. This is in part an elaboration of Link’s (1983) earlier idea that *all* introduces a ‘partakes in’ operator.

He says of the predicates that do not allow *all*, *like be a good team* and *be a big group*, that they are “pure cardinality predicates”. As such, they do not have any subentailments for *all* to “operate on”, which explains the ungrammaticality of sentences like (59)–(60).

However, 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? And are these part of the lexical specification of the predicate?

Lacking a clear definition of what subentailments are, or independent evidence for their existence, the account offered by Dowty suffers from circularity. The only evidence for the existence of subentailments in the lexical specification of any given predicate is the distribution of *all*, yet subentailments are called upon to explain the distribution of *all*.

This problem with Dowty’s theory was pointed out by Taub (1989). She also added an important observation about the types of collective predicates that do and do not allow *all*. I call it Taub’s generalization, and it is given in (61).

- (61) **Taub’s generalization:** the collective predicates that disallow *all* are the collective predicates denoting states and achievements.

The evidence for Taub’s generalization comes from the following examples:

Collective states:

- (62) *All the boys are a big group.
- (63) *All the trees are dense in the middle of the forest. (*does not allow collective reading*)

Collective activities:

- (64) All the boys carried the piano around for an hour.

Collective accomplishments:

(65) All the students gathered in the hallway.

(66) All the girls built a raft.

Collective achievements:

(67) *All the senators passed the pay raise.

(68) *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.¹⁰ But it is not the case that only cardinality predicates disallow *all*. The example in (63) is from Taub's paper (she attributes the example to Barbara Partee) and I would add (69). On a collective reading, both are infelicitous with *all*.

(69) *The bottles are all too heavy to carry.

Also, Manfred Krifka (p.c.) points out that the following "minimal pair" follows Taub's generalization.

(70) *All the girls are a big group. (state)

(71) All the girls formed a big group. (accomplishment)

In addition, it turns out that the same generalization that Taub offers about collective predicates with *all* also holds true of collective predicates with *except* phrases, as shown in (72)–(75).¹¹

(72) *The girls are a big group, except for Kim and Hannah.

¹⁰ Note that reciprocals and "collectivized" predicates do not follow Taub's Generalization, as shown in (i)–(iii). (Thanks to an anonymous reviewer for pointing this out.) *Be alike* is a state, *recognize* and *arrive* are achievements.

(i) All the dogs look alike.

(ii) All the boys recognized each other.

(iii) All the planes arrived together.

This is why we are excluding them from consideration for the time being. We will understand why these predicates behave the way they do after sections 5.5 and 6.

¹¹ An anonymous reviewer finds that some English speakers do not accept the contrast in (72)–(75). In my tests I have not found any speakers who reject it.

- (73) The campers built rafts every summer, except for the youngest ones.
- (74) The students built a raft, except for Maggie and Josh.
- (75) *The students elected Mike, except for the sophomores.

Finally, the collective states and achievements seem to resist nonmaximality in a way that collective activities and accomplishments do not. For example, sentences with definite plurals allow a phrase starting with *but* ... that can reinforce or make explicit that a nonmaximal reading is intended, as shown in (76)–(77).

- (76) The girls jumped in the lake, but Julia didn't.
- (77) The girls ate a sandwich, but Elizabeth didn't.

Collective activities and accomplishments also allow this reinforcement of nonmaximality.¹²

- (78) The girls gathered in the auditorium, but Mary and Jean didn't.
- (79) The girls built a raft, but Helen and Vivian didn't.

In contrast, however, the collective states and achievements sound quite odd with this kind of continuation.

- (80) *The girls are a group of four, but not Kim/but Kim isn't.
- (81) *The girls elected a president, but Jane didn't.

The idea of allowing a predicate like 'be a group of four' or 'elect a president' to apply nonmaximally seems almost like a category mistake.

¹² Lasersohn (1999) cites (i) (from Kroch 1974) to make the opposite point.

- (i) Although the townspeople are asleep, some of them are awake.

He says that this sentence "sounds contradictory". I agree that this sentence sounds worse than (76)–(79) in the text. It appears that reinforcing nonmaximality is limited to certain kinds of phrases, similar to the kinds of phrases which can be used to cancel or reinforce conversational implicatures.

Why should this be, since other collective predicates are compatible with nonmaximality?

The explanation for all of these properties of collective predicates will be that some collective predicates actually can have a D operator, hence a quantificational component of their meaning. The collective predicates that do have a D operator are then predicted to participate in all of the phenomena that we have shown are associated with distributive quantification: nonmaximality (a domain-of-quantification effect); compatibility with *except* phrases, which according to von Stechow (1995) are ‘domain subtractors’; and compatibility with *all*, which crucially interacts with the D operator. Those collectives that do not have any quantificational component of their meaning are shown not to exhibit any of these qualities. In the next section I’ll sketch an analysis of collectives that will do the work we need to explain these phenomena.

5.1. *Collectives and DO*

Using Taub’s generalization as a guide, I propose that a bleached-out activity predicate called DO, which is a subcomponent of the meaning of activity and accomplishment verbs (Mittwoch (1982), and see section 5.4), is projected as a separate and phonetically empty verbal head in the syntax (see Larson 1988). This predicate can host a D operator and when it does, the result is a “collective” reading for the sentence. On the other hand, we will see that collective states and achievements cannot host a D operator, and so they will not show any of the properties associated with distributive quantification (except in certain circumstances, which we will see are actually predicted by the proposal). The account will be spelled out in the framework of neo-Davidsonian event semantics.

5.2. *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 (82) is interpreted as (84) (as opposed to (83)).

(82) Alvin kissed Roxy.

(83) kiss’(Roxy)(Alvin)

(84) $\lambda e[\text{kiss}'(e) \ \& \ \text{Ag}(e, \text{Alvin}) \ \& \ \text{Th}(e, \text{Roxy})]$

(85) $\exists e[\text{kiss}'(e) \ \& \ \text{Ag}(e, \text{Alvin}) \ \& \ \text{Th}(e, \text{Roxy})]$

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 of existential closure takes the predicate of events in (84) and changes it to a proposition as in (85) by binding the variable over events e with an existential quantifier.

I use the term *event* to include states, because the distinction between events and states does not do any work for us here. Bach (1981) proposed the term *eventualities* to include both states and events, and many contemporary authors do not see a useful distinction between states and events (e.g., Kratzer 1998). So I will use the term *event* in its more general sense.

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. So we will redefine the distributivity operator so that it works in event semantics. The new D operator is defined as in (86); note that we introduce existential quantification over subparts of an event.

$$(86) \quad \lambda P \lambda x \lambda e \forall y \exists e' [y \subseteq x \ \& \ y \in \llbracket \text{Cov}_i \rrbracket \rightarrow P(e')(y) \ \& \ e' \leq e]$$

This will apply to a VP like (87) to yield (88). (I treat P here as a variable that ranges over functions from individuals to sets of events (type $\langle i, \langle v, t \rangle \rangle$, see below), because this is what a VP denotes in event semantics.)

$$(87) \quad {}^D \lambda x \lambda e [ate'(e) \ \& \ Ag(e, x) \ \& \ Th(e, a.sandwich')] =$$

$$(88) \quad \lambda x \lambda e \forall y \exists e' [y \subseteq x \ \& \ y \in \llbracket \text{Cov}_i \rrbracket \rightarrow ate'(e') \ \& \ Ag(e', y) \ \& \ Th(e', a.sandwich') \ \& \ e' \leq e]$$

When we have fed in *the girls* as the subject argument and existentially bound the variable e , we get (89).

$$(89) \quad \exists e \forall y \exists e' [y \subseteq \llbracket the.girls' \rrbracket \ \& \ y \in \llbracket \text{Cov}_i \rrbracket \rightarrow ate'(e') \ \& \ Ag(e', y) \ \& \ Th(e', a.sandwich') \ \& \ e' \leq e]$$

Thus we take the distributive reading of *the girls ate a sandwich* to assert the existence of an event which contains several subevents: these subevents are individual ate-a-sandwich events for each girl.

Before we can go any further in discussing the parts of an event, we must give more specific details about the assumptions about event structure that make this possible. I assume that the domain of events is cumulative; that is, that events can form sums which are themselves events. This

is similar, though not identical, to ideas in Landman (1996) and others (e.g., Krifka 1992.) Landman assumes that there is a difference between singular and plural events, a distinction that doesn't do any work for us here.

Every account of event semantics has to address the following problem. Because we existentially quantify over events, there are many events that could potentially make the sentence in (90), translated as (91), true.

(90) John ate beans.

(91) $\exists e[\text{ate}'(e) \ \& \ \text{Ag}(e,J) \ \& \ \text{Th}(e,\text{beans}')]$

For example, obviously an event in which John ate beans makes (91) 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 can cause problems in any event semantics.

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 *Cecelia sang*, and suppose we are in an opera house where Cecelia just sang an aria. If we use minimal events to evaluate the sentence *Cecelia 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 Cecelia 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 notion of an event for a proposition is defined so that we are allowed to choose sums of 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 Cecelia singing the whole aria as the event for the proposition *Cecelia sang*, as long as it doesn't include any parts that are not events of Cecelia singing.

The definition of an event for a proposition is given in (92) (Kratzer 1998). (Recall that propositions denote sets of events before existential closure.)

- (92) An event e is an event for a proposition p iff:

$$\forall e'[e' \leq e \ \& \ e' \notin p \rightarrow \exists e''[e' \leq e'' \leq e \ \& \ \min(e'')[e'' \in p]]]$$

The expression $\min(e'')[e'' \in p]$ 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 $p(e)$, we always check whether e is an event for that proposition p , that is, whether it is an event e that does not contain extraneous stuff.

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 Lasnik (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 $\langle e, t \rangle$, but $\langle i, \langle v, t \rangle \rangle$. It can be represented equivalently as either (93) or (94).

- (93) sing'

- (94) $\lambda x \lambda e [\text{sing}'(e) \ \& \ \text{Ag}(e, x)]$

Even though we have some new types in our system I will retain the conventions of variable use as closely as possible. So, for example, 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.

Finally, I will adopt two stipulations introduced by Landman (1996). The first is the “unique role requirement”, which says 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.¹³ The second is the “scope domain principle”, (SDP)

¹³ The URR is not uncontroversial. One problem is that it cannot straightforwardly account for the fact that if I touch John's hand, then I have also touched his arm, and I have also touched John. I thank Manfred Krifka (p.c.) for bringing this to my attention, but I will not address this issue here.

which says that quantified DPs must take scope outside of the existential quantifier introduced by existential closure (this is the scope domain). This is one way of capturing in event semantics the idea that quantified expressions must undergo QR.¹⁴ (See Landman (1996) for discussion and arguments in favor of these rules.)

For concreteness, we will suppose that existential closure is introduced at the IP level, and that QR is adjunction to IP or a higher verbal projection.¹⁵ However, unlike Landman, I will assume that plural definites are never quantificational, and hence never subject to the SDP. (Landman assumes that definites may optionally take scope outside of the scope domain.)

5.3. Generalized Distributivity

Lasersohn (1998) 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”.

Based on the definition of generalized conjunction in (95)–(96), Lasersohn gives a definition of a distributable type and a rule for a D operator, shown in (97)–(98).

(95)a. t is a conjoinable type

b. If $\langle \mathbf{a}, \mathbf{b} \rangle$ is a type and b is a conjoinable type, then $\langle \mathbf{a}, \mathbf{b} \rangle$ is a conjoinable type.

(96)a. If $X \subseteq \mathbf{D}_t$, then $\sqcap X = 1$ if $X = \{1\}$; $\sqcap X = 0$ otherwise.

b. If $X \subseteq \mathbf{D}_{\langle \mathbf{a}, \mathbf{b} \rangle}$ (where $\langle \mathbf{a}, \mathbf{b} \rangle$ is a conjoinable type), then $\sqcap X$ is that function $f \in \mathbf{D}_{\langle \mathbf{a}, \mathbf{b} \rangle}$ such that for all a , $f(a) = \sqcap \{f'(a) \mid f' \in X\}$

(97) *Distributable types:*

If \mathbf{a} is a conjoinable type, then $\langle \mathbf{e}, \mathbf{a} \rangle$ is a distributable type.

(98) Where α is an expression of some distributable type $\langle \mathbf{e}, \mathbf{a} \rangle$ and x is any individual (i.e., $x \in \mathbf{D}_e$): $\llbracket^D \alpha \rrbracket^{\mathbf{M}, \mathbf{g}}(x) = \sqcap \{ \llbracket \alpha \rrbracket^{\mathbf{M}, \mathbf{g}}(y) \mid y \leq x \}$

¹⁴ 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.

¹⁵ 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.

However, we need a generalized D operator that is suitable for event semantics, and includes the variable Cov. Lasersohn's definition is compatible with that. He also shows how to revise the definition of a conjoinable type and the D operator in to take the type of events into account. These revisions are given in (99) and (100) (adapted from Lasersohn 1998: 87).

(99)a. $\langle v, t \rangle$ is a conjoinable type.

b. If $\langle \mathbf{a}, \mathbf{b} \rangle$ is a type and \mathbf{b} is a conjoinable type, then $\langle \mathbf{a}, \mathbf{b} \rangle$ is a conjoinable type.

(100)a. If $X \subseteq \mathbf{D}_{\langle v, t \rangle}$ then $\sqcap X$ is that function $f: \mathbf{D}_v \rightarrow \mathbf{D}_t$ such that for all $e \in \mathbf{D}_v$:

$$f(e) = 1 \text{ iff } \forall f' \in X \exists e' \leq e f'(e') = 1.$$

b. If $X \subseteq \mathbf{D}_{\langle \mathbf{a}, \mathbf{b} \rangle}$ (where $\langle \mathbf{a}, \mathbf{b} \rangle$ is a type and \mathbf{b} is a conjoinable type), then $\sqcap X$ is that function $f \in \mathbf{D}_{\langle \mathbf{a}, \mathbf{b} \rangle}$ such that for all a , $f(a) = \sqcap \{f'(a) \mid f' \in X\}$.

Our final definition of a D operator, given in (102), also takes into account the presence of Cov as a domain restrictor.

(101) *Distributable types:*

If \mathbf{a} is a conjoinable type, then $\langle i, \mathbf{a} \rangle$ is a distributable type.

(102) Where α is an expression of some distributable type $\langle i, \mathbf{a} \rangle$ and x is any individual (i.e., $x \in \mathbf{D}_i$): $\llbracket \mathbf{D}\alpha \rrbracket^{\mathbf{M}, \mathbf{g}}(x) = \sqcap \{ \llbracket \alpha \rrbracket^{\mathbf{M}, \mathbf{g}}(y) \mid y \subseteq x \ \& \ y \in \llbracket \text{Cov}_i \rrbracket^{\mathbf{M}, \mathbf{g}} \}$

To make our derivations simpler, we'll adapt Lasersohn's definition of a D operator into something that can more transparently apply to objects of type $\langle i, \mathbf{a} \rangle$. I give a revised definition of a generalized D operator below.

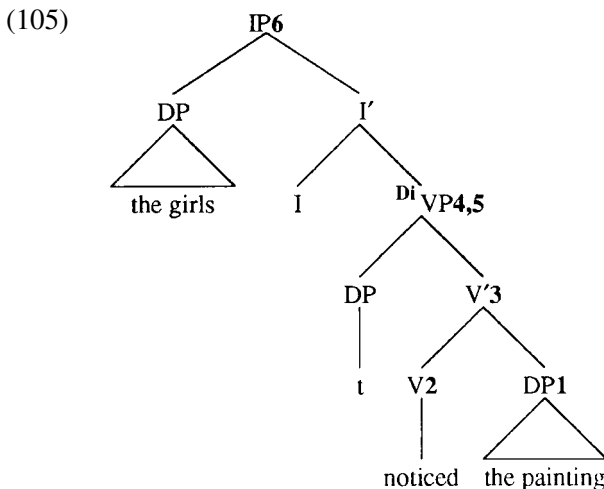
(103) Where Z is a variable of type $\langle i, \mathbf{a} \rangle$ (i.e., a distributable type),
 $\mathbf{D} =_{\text{df}} \lambda Z \lambda x \sqcap \{Z(z) \mid z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$

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 applies to a thing that takes a plural DP as its (first) argument.^{16,17}

Let us apply this to an example. We will do some “unpacking” at the end of the derivation.

(104) The girls noticed the painting.¹⁸



1. the.painting'
2. $\lambda x \lambda y \lambda e [\text{noticed}'(e) \ \& \ \text{Ag}(e,y) \ \& \ \text{Th}(e,x)]$
3. $\lambda y \lambda e [\text{noticed}'(e) \ \& \ \text{Ag}(e,y) \ \& \ \text{Th}(e,\text{the.painting}')]$
4. $\lambda P \lambda x \{P(z) | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$ (*the meaning of^D*)
5. $\lambda x \cap \{\lambda e [\text{noticed}'(e) \ \& \ \text{Ag}(e,z) \ \& \ \text{Th}(e,\text{the.painting}')]| z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$ (*the result of 4(3)*)
6. $\cap \{\lambda e [\text{noticed}'(e) \ \& \ \text{Ag}(e,z) \ \& \ \text{Th}(e,\text{the.painting}')]| z \subseteq \llbracket \text{the.girls}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
 $= \lambda e' [e' \in \cap \{\lambda e [\text{noticed}'(e) \ \& \ \text{Ag}(e,z) \ \& \ \text{Th}(e,\text{the.painting}')]| z \subseteq$

¹⁶ I will later argue that we need a slightly more restrictive rule that is sensitive both to context and to considerations of economy (see section 5.3).

¹⁷ An anonymous reviewer asks why the D operator is defined in such a way that it does not apply to objects of type $\langle v,t \rangle$. In event semantics, sentences denote not truth conditions, but predicates of events (before existential closure). Therefore Lasnik (1998) defines conjoinable types at those ending in $\langle v,t \rangle$, which is the type that corresponds to $\langle t \rangle$ in semantics without events.

¹⁸ I have called the subject of *notice* an agent here, although it's arguable that the thematic role is something like experiencer instead. This is not crucial to the operating of the D operator.

$\llbracket \text{the.girls}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \}$

after \exists -closure: $= \exists e'[e' \in \sqcap \{ \lambda e[\text{noticed}'(e) \ \& \ \text{Ag}(e,z) \ \& \ \text{Th}(e,\text{the.painting}')] | z \subseteq \llbracket \text{the.girls}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \}]$

The part-of structure on events that the D operator introduces is, in a sense, “buried” in the meaning of generalized conjunction. However, an unpacking of the function that generalized conjunction yields will make it more apparent.

The \sqcap operator takes as its argument the set of functions F that we get by plugging each individual part of the denotation of *the girls* as the agent argument of *noticed*. If $\llbracket \text{the.girls}' \rrbracket = \{a,b,c\}$, then F could alternately be written as in (106).

- (106) $\{ \lambda e[\text{noticed}'(e) \ \& \ \text{Ag}(e,a) \ \& \ \text{Th}(e,\text{the.painting}')], \lambda e[\text{noticed}'(e) \ \& \ \text{Ag}(e,b) \ \& \ \text{Th}(e,\text{the.painting}')], \lambda e[\text{noticed}'(e) \ \& \ \text{Ag}(e,c) \ \& \ \text{Th}(e,\text{the.painting}')] \}$

When the generalized conjunction operator takes this set of functions as its argument, it yields a function f' of type $\langle v,t \rangle$. This function will yield the value true for some event e only if for every function f in the set of functions in (106), 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 girls*. It looks like this:

- (107) $\lambda e \forall x \exists e' [x \subseteq \llbracket \text{the.girls}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{noticed}'(e') \ \& \ \text{Ag}(e', x) \ \& \ \text{Th}(e', \text{the.ptg}') \ \& \ e' \leq e]$

- (108) $\exists e \forall x \exists e' [x \subseteq \llbracket \text{the.girls}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{noticed}'(e') \ \& \ \text{Ag}(e', x) \ \& \ \text{Th}(e', \text{the.ptg}') \ \& \ e' \leq e]$

In other words, (108) means the same thing as the last line of (105). As we move on to look at more examples in this paper, we will continue to “unpack” the meaning of a sentence with a \sqcap operator where it will make the discussion easier to follow.

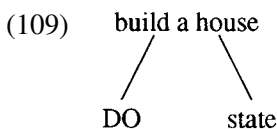
At this point, we are finally ready to look at the theory of collective predication that will explain the phenomena discussed at the beginning of section 4. It is worth pointing out that all of the machinery related to events introduced in this section has been independently proposed and argued for; the innovation proposed in this paper is contained in the section to come, in which we will introduce syntactic structure in collective activities and accomplishments that will “make room” for a low-scope D operator.

5.4. *The Structure of Collective Activities and Accomplishments*

By wide consensus in the literature, what activities and accomplishments have in common is a subcomponent of their meaning called "DO." This component has been given various names by various authors, including "activity", "process" and "DO". (See e.g., Ross (1972), Dowty (1979), Mittwoch (1982), Pustejovsky (1991), Grimshaw and Vikner (1993) and others.) The basic idea is that "DO" is, in a sense that is made precise differently in different theories, an "activity". 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. 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 (109).¹⁹



The fact that accomplishments also have a stative predicate in their lexical structure does not affect the proposal that I will make here. The crucial idea is that the activity component of the predicate is separate from the lexical head of the verb.

In contrast to activities and accomplishments, states and achievements lack DO. In fact, more importantly for our purposes, they lack the complex internal structure that I assume is present in activities and accomplishments. The idea that a stative eventuality does not have internal structure is relatively uncontroversial, so I will assume it without further argument.

¹⁹ 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)). See below for more discussion on this point.

Achievements, however, are more controversial. Many authors (e.g., Krifka (1989), Parsons (1990)) deny that they constitute a separate aspectual class at all. However, Mittwoch (1991), Pustejovsky (1991), and Piñon (1997) defend the necessity for treating achievements separately from accomplishments on several grounds. Here are just two differences between achievements and accomplishments.

Achievements allow modification by a point adverbial but accomplishments do not.

(110) Jane arrived at noon.

(111) *Sally built a house at noon.

Accomplishments and achievements show different behavior with *in* adverbials. (112) and (113) are roughly synonymous, while (114) and (115) mean very different things.

(112) Jane recognized Alex in twenty minutes.

(113) Jane recognized Alex after twenty minutes.

(114) Sally baked a cake in twenty minutes.

(115) Sally baked a cake after twenty minutes.

While (115) refers, roughly, to what happened at the end of a twenty-minute period, (114) refers to what happened *during* a twenty minute period.

Both of these differences (for more discussion and evidence, see the works cited above) suggest that the difference between achievements and accomplishments is that accomplishments include reference to the event leading up to the endpoint, while achievements refer just to the endpoint itself.

Some authors (e.g., Pustejovsky) assume that achievements are nevertheless complex, in that they include the preceding state; that is, the event of Jane recognizing Alex includes the state in which she had not yet recognized him. But I take the position that the preceding state is not part of the denotation of the predicate, but something more like a presupposition. I also assume that if “become” or “cause” are present, then they are operators, or perhaps relations between this presupposed event and the event denoted by the predicate (see Davidson (1967) on relations between events, and Piñon (1997) on achievements).

Thus I am proposing that the crucial difference between achievements and states on the one hand, and activities and accomplishments on the other, is that the former lack the substructure that is contained in the latter. The complex substructure of activities and accomplishments will give us “room” to introduce distributivity on just part of the (complex) eventuality denoted by the predicate.

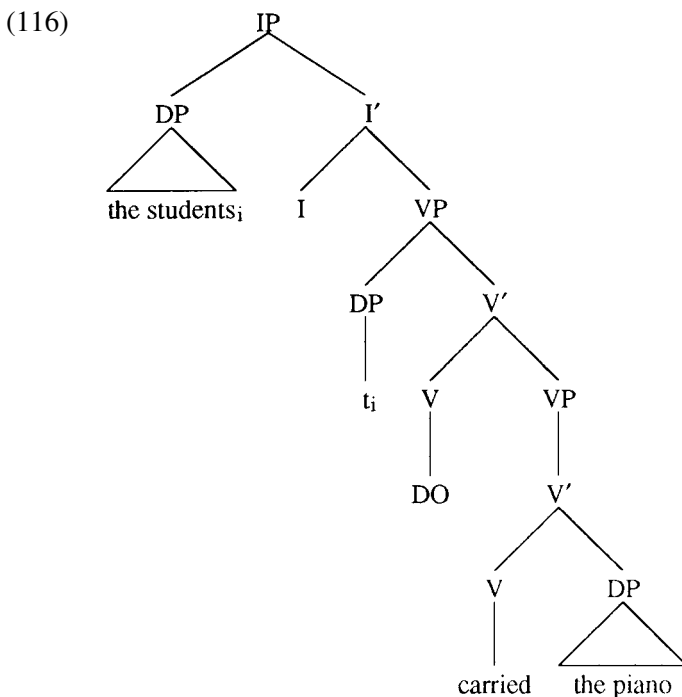
The mechanism for instantiating this idea works as follows. I propose that the DO component of activities and accomplishments is projected into the syntax as a kind of aspectual head that takes the phrase projected by the lexical head of the verb (for instance, *build*) as its sister. This proposal is 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 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) 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 (i.e., single-object) verbs.

The proposal I am making here is probably closest to the syntactic proposals in McClure (1994). McClure proposes two aspectual projections above VP that are licensed (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 the agent) as its argument. The structure of a sentence like *the students carried the piano*, then, looks like this:



Two syntactic issues should immediately be addressed. First, in my examples I show that the only functional head above DO is I. 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 (i.e., 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 (e.g., 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. What is crucial here is that there are two separate heads within the VP, and that the lexical semantics of the verb is divided between the two heads in the manner described below.

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).

$$(117) \quad \text{DO} = \lambda x \lambda e' [\text{DO}(e) \ \& \ \text{Ag}(e, x)]$$

$$(118) \quad \text{carry}' = \lambda x \lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, x)]$$

Given the structure proposed in (116), 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".

5.5. *Event Composition*

Based on the evidence from many theories of argument structure (see, e.g., 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 proposes that the external argument is introduced by an independent functional head, 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 be reminiscent of Higginbotham's (1985) "theta identification".) Kratzer's definition of event identification is given in (119).

(119) Event Identification:

f	g	\rightarrow	h
$\langle i \langle v, t \rangle \rangle$	$\langle v, t \rangle$		$\langle i \langle v, t \rangle \rangle$
			$\lambda x \lambda e [f(x)(e) \ \& \ g(e)]$
$\lambda x \lambda e [\text{Ag}(e, x)]$	$\lambda e [\text{wash}'(\text{clothes}')(e)]$		$\lambda x \lambda e [\text{Ag}(e, x) \ \& \ \text{wash}'(\text{clothes}')(e)]$

Event identification in the form Kratzer proposes won't work for the structure I am adopting for activities and accomplishments. To see why, I

refer to the structure given in (116). Whereas Kratzer's "f" is a function of the form $\lambda x \lambda e [Ag(e)(x)]$, our "f", which is the DO part of the predicate, also includes the predicate of events DO, as in $\lambda x \lambda e [DO(e) \ \& \ Ag(e,x)]$. If we combine this expression with an expression like $\lambda e [carry'(e) \ \& \ Th(e,piano')]$, then we would end up claiming that the DOing and the carrying are the same thing: event identification would give us $\lambda x \lambda e [carry'(e) \ \& \ Th(e,piano) \ \& \ DO(e) \ \& \ Ag(e,x)]$ as the meaning of the higher VP in (116).

The claim I am making here is that the DOing event is a separate component of the meaning of activity and accomplishment predicates. Therefore, for these predicates, we cannot use event "identification". I propose instead an operation I call event composition, in which the DO part of the event is composed with the lexical meaning of the verb.

Event composition:

$$\begin{array}{lll}
 f & g & \rightarrow h \\
 \langle v, t \rangle & \langle i, \langle v, t \rangle \rangle & \langle i, \langle v, t \rangle \rangle \\
 \lambda e [carry'(e) \ \& \ Th(e,piano)] & \lambda x \lambda e [DO(e) \ \& \ Ag(e,x)] & \rightarrow \\
 & \lambda x \lambda e [carry'(e) \ \& \ Th(e,piano') \ \& \ \exists e' [DO(e') \ \& \ Ag(e',x) \ \& \ e' \leq e]]
 \end{array}$$

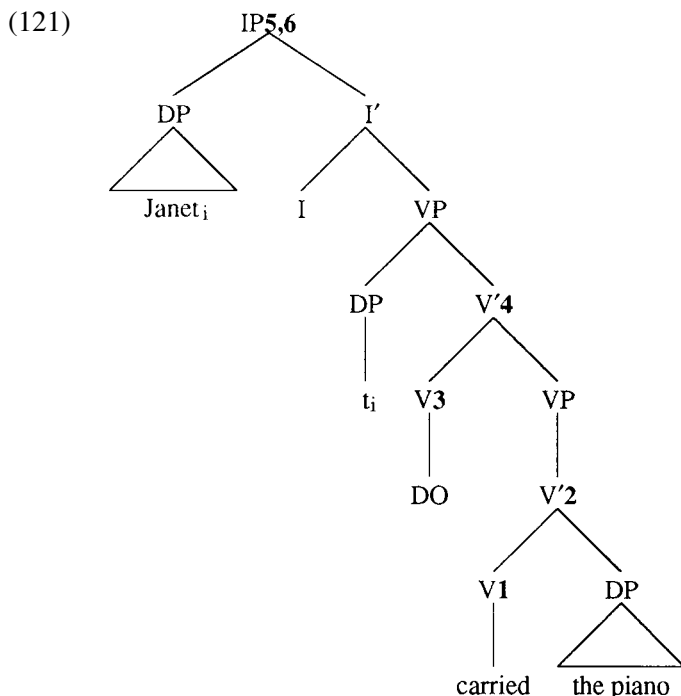
In fact since we are not "identifying" the events, we could also define event composition as an operator that can be used whenever we want to combine functions of the right type. The operator would be defined as follows, where Q is a variable over objects of type $\langle v, t \rangle$.

$$(120) \quad \lambda Q \lambda P \lambda x \lambda e [Q(e) \ \& \ \exists e' [P(x)(e')] \ \& \ e' \leq e]$$

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. Kratzer stipulates that event composition is available whenever we want to combine two expressions of the types corresponding to "f" and "g".

Consistent with the hypothesis about the difference between predicates that contain DO and the predicates that don't, I assume that states and achievements combine with their external arguments by means of Kratzer's original process of event identification. The idea is that those predicates with complex event substructure compose by event composition, and those with simple event structure compose by event identification. If we use event identification where we need event composition, or vice versa, the resulting meaning is ill-formed and hence the derivation using the incorrect process is ruled out.

At this point, we have enough tools in hand to spell out the details of the structure of a sentence with an activity or accomplishment predicate. At first I will use an example with a singular DP subject, and then we will see how these predicates, when collective, contain hidden distributivity. As shown in (121), the lexical head projects a lower VP, and DO heads a projection that dominates that VP. The semantic derivation proceeds as shown below.



1. $\lambda x \lambda e[\text{carry}'(e) \ \& \ \text{Th}(e,x)]$
2. $\lambda e[\text{carry}'(e) \ \& \ \text{Th}(e,\text{the.piano}')]]$
3. $\lambda x \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e,x)]$ (2 and 3 combine via event composition)
4. $\lambda x \lambda e[\text{carry}'(e) \ \& \ \text{Th}(e,\text{the.piano}') \ \& \ \exists e'[\text{DO}(e') \ \& \ \text{Ag}(e',x) \ \& \ e' \leq e]]$
5. $\lambda e[\text{carry}'(e) \ \& \ \text{Th}(e,\text{the.piano}') \ \& \ \exists e'[\text{DO}(e') \ \& \ \text{Ag}(e',J) \ \& \ e' \leq e]]$
6. $\exists e[\text{carry}'(e) \ \& \ \text{Th}(e,\text{the.piano}') \ \& \ \exists e'[\text{DO}(e') \ \& \ \text{Ag}(e',J) \ \& \ e' \leq e]]$

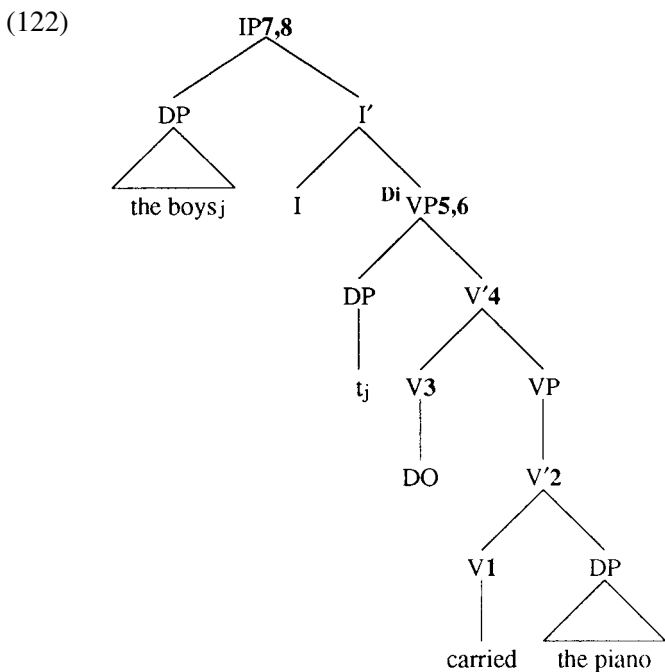
The expression in line 6 of (121) has the following truth conditions. It says that there is an event of carrying the piano in which Janet is the agent of its DO subpart. Notice that the notion of an event for a proposition does some work for us here. Without it, this sentence might count as true by virtue 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 the piano being carried with Janet as 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.

6. COLLECTIVE PREDICATES WITH HIDDEN QUANTIFICATION

This analysis of activity and accomplishment predicates makes possible a new analysis of the collective readings of these predicates. The D operator interacts with the syntax and semantics we have just proposed. The distinction between distributive and collective readings is captured by the two possible insertion sites for a D operator: on DO, which yields a collective reading, and on the VP dominating DO, which yields a distributive reading. (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.) Thus, the reading that we have been calling “collective” turns out to actually contain a sort of hidden distributivity.

When the D operator is on the highest VP, the result is an ordinary distributive reading – in other words, the same reading that other theories of distributivity predict when a D operator is present.

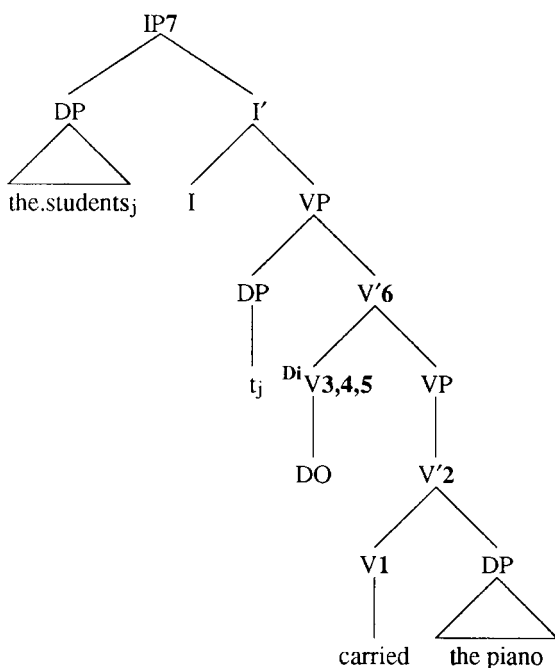


1. $\lambda x \lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, x)]$
2. $\lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}')]]$
3. $\lambda x \lambda e' [\text{DO}(e') \ \& \ \text{Ag}(e', x)]$ (2 and 3 combine via event composition)
4. $\lambda x \lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e]]$
5. $\lambda P \lambda x \sqcap \{P(z) | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$ (the meaning of ^D)
6. $\lambda x \sqcap \{\lambda x \lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e]] \ (z) | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$ (the result of 5(4))
7. $\sqcap \{\lambda x \lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e]] \ (z) | z \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
- 7'. $= \sqcap \{\lambda e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \leq e]] | z \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
8. $= \lambda e \forall z \exists e'' [z \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{carry}'(e'') \ \& \ \text{Th}(e'', \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \leq e'' \ \& \ e'' \leq e]]$
- 8'. $= \exists e \forall z \exists e'' [z \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{carry}'(e'') \ \& \ \text{Th}(e'', \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \leq e'' \ \& \ e'' \leq e]]$

This sentence asserts the existence of a separate carry-the-piano event for each one of the boys (modulo nonmaximality). 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 yields what we have been calling a collective reading.

(123)



1. $\lambda x \lambda e[\text{carry}'(e) \ \& \ \text{Th}(e, x)]$
2. $\lambda e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}')]]$
3. $\lambda x \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, x)]$
4. $\lambda P \lambda y \sqcap \{ P(z) | z \leq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \}$ (*the meaning of D*)
5. $\lambda y \sqcap \{ \lambda x \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, x)](z) | z \subseteq y \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \}$ (*the result of 4(3)*)
- 5'. $= \lambda y \sqcap \{ \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, z)] | z \subseteq y \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \}$ *to make event composition easier, we will use the following notational variant of 5'*
- 5''. $= \lambda y \lambda e' [e' \in \sqcap \{ \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, z)] | z \subseteq y \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \}]$ *2 and 5' combine via event composition*
6. $\lambda y \lambda e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \in \sqcap \{ \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, z)] | z \subseteq y \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \} \ \& \ e' \leq e]$
7. $\exists e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \in \sqcap \{ \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, z)] | z \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ Z \in \llbracket \text{Cov}_i \rrbracket \} \ \& \ e' \leq e]$

To “unpack” the function given by the generalized conjunction operator into something that is more transparent, let us suppose that the students are Harry, Bill, and Tom. Then the set F that the \sqcap operator takes as its argument is the set in (124), namely, the set of functions we get by substituting each one of the students for the Agent argument of DO.

$$(124) \quad \{ \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, h)], \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, b)], \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, t)] \}$$

Because the generalized conjunction operator introduces existential quantification over events when it applies to sets of functions of type $\langle v, t \rangle$, when it applies to the set in (124) it will yield the following function.

$$(125) \quad \lambda e''' \forall x \exists e'' [x \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow [\text{DO}(e'') \ \& \ \text{Ag}(e'', x)] \ \& \ e'' \leq e''']$$

Now we can plug this back in to the formula in line 7 of the derivation in (123).

$$\begin{aligned}
 (126) \quad \text{line 7:} \\
 & \exists e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \in \sqcap \{ \lambda e[\text{DO}(e) \ \& \ \text{Ag}(e, z)] | z \subseteq \llbracket \text{Cov}_i \rrbracket \} \ \& \ e' \leq e] \\
 & = \exists e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \in \lambda e''' \forall x \exists e'' [x \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow [\text{DO}(e'') \ \& \ \text{Ag}(e'', x)] \ \& \ e'' \leq e''' \ \& \ e' \leq e]] \\
 & \exists e[\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}') \ \& \ \exists e' \forall x \exists e'' [x \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow [\text{DO}(e'') \ \& \ \text{Ag}(e'', x)] \ \& \ e'' \leq e' \ \& \ e' \leq e]
 \end{aligned}$$

The last expression in (126) 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 nonmaximality).

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. An important difference is that in this account 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 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: what is entailed of each member of the group is agenthood in an event of DOing that is part of the collective activity.

Now we have an account of the collectivity of activities and accomplishments that predicts that they should have the same characteristics that we find in other phenomena that can be traced to the presence of distributive quantification. We will see that this explains why these collectives are compatible with *all*, why they are compatible with *except*-phrases, and why they show the effect of nonmaximality.

6.1. *Consequences: all*

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 (127), the derivation will yield a distributive reading, represented by (128).

(127) The students carried the piano.

(128) $\exists e \forall z \exists e'' [z \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{carry}'(e'') \ \& \ \text{Th}(e'', \text{the.piano}') \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \leq e'' \ \& \ e'' \leq e]]$

When D is inserted on the aspectual head DO of a sentence like (127), the derivation will yield a collective reading, represented by (129).

(129) $\exists e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}')] \ \& \ \exists e' \forall x \exists e'' [x \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{DO}(e'') \ \& \ \text{Ag}(e'', x) \ \& \ e'' \leq e' \ \& \ e' \leq e]]$

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 (130), with *all*, the derivation will yield a distributive and “maximized” reading, represented by (131).

(130) The students all carried the piano.

(131) $\exists e \forall z \exists e'' [z \subseteq \llbracket \text{the.boys}^{\text{gf}} \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{carry}'(e'') \ \& \ \text{Th}(e'', \text{the.piano}) \ \& \ \exists e' [\text{DO}(e') \ \& \ \text{Ag}(e', z) \ \& \ e' \leq e'' \ \& \ e'' \leq e]]$

Of course (130) can be interpreted collectively and “maximally”, if D is inserted on DO.

(132) $\exists e [\text{carry}'(e) \ \& \ \text{Th}(e, \text{the.piano}')] \ \& \ \exists e' \forall x \exists e'' [x \subseteq \llbracket \text{the.students}^{\text{gf}} \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{DO}(e'') \ \& \ \text{Ag}(e'', x) \ \& \ e'' \leq e' \ \& \ e' \leq e]]$

Thus the “maximizing effect” of *all* with the collective reading comes about because each individual student is asserted to be an agent of a DO-ing event that is part of the carrying event.

Given this analysis of lexically collective predicates 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 (134).

(133) Every girl built a raft.

(134) $\forall x[x \in \llbracket \text{girl}' \rrbracket \rightarrow \exists e[\text{built}'(e) \ \& \ \exists e'[\text{DO}(e') \ \& \ \text{Ag}(e',x) \ \& \ e' \leq e] \ \& \ \text{Th}(e,a.\text{raft}')]]]$

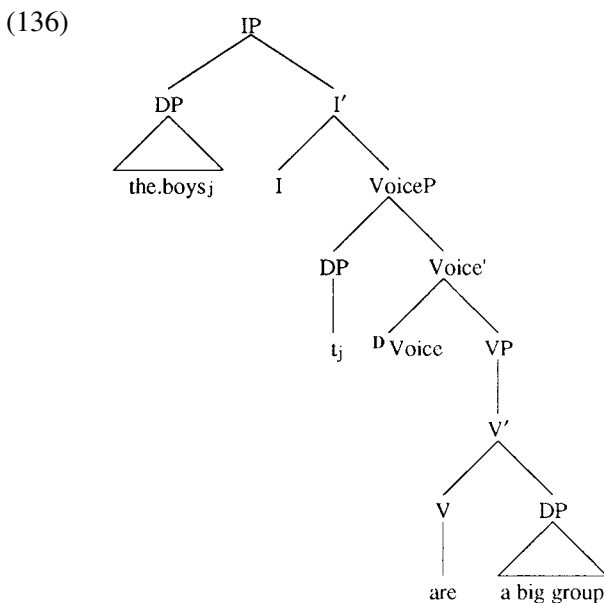
6.2. *The Other Half of Taub's Generalization: Distributivity and Economy*

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 the derivation of the sentence in (135).

(135) The boys are a big group.

Adopting Kratzer's hypothesis about the external argument, we assume that the external argument and its thematic role (let us simply call it "role" in this case) are introduced by a separate head, as shown below.



For Kratzer, the voice head is external to VP, unlike the DO head which I have proposed is internal to VP (this difference is not crucial). In this sentence there are two places where we can insert a D operator: voice head, which is a distributable type because it takes a DP as its first argument (its translation is $\lambda x \lambda e [\text{role}(x)(e)]$); and VoiceP, which is equivalent to putting D on VP in the examples we have already been considering. In either case, the derivation is ill-formed.

If we insert D on voice, we get a result that is semantically ill-formed. Distributivity applies only to the thematic role, resulting in the odd proposition shown in (137).

- (137) $\exists e \forall z \exists e' [z \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{role}(z)(e') \ \& \ e' \leq e$
 $\ \& \ \text{be.a.big.group}'(e)]$

This proposition asserts the existence of “role” events that are not directly linked to any verbal (or adjectival, or nominal) predicate. But thematic roles are not events in and of themselves.

A better candidate emerges if we insert D on Voice P. The result of that derivation is shown in (138).

- (138) $\exists e \forall x \exists e' [x \subseteq \llbracket \text{the.boys}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{big.group}'(e') \ \& \ \text{role}(e', x) \ \& \ e' \leq e]$

But there are still problems here. (138) 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.

One way to make it possible for a D operator to appear here is to allow it to quantify over a singleton domain: if the value assigned to Cov contains a set that is equal to the denotation of *the boys*, then the truth conditions of (138) are not any different from the truth conditions of this same sentence without a D operator. But this would also mean that the D operator does not change the truth conditions in any way. This violates principles of economy, which would rule out adding an operator to a derivation when it doesn't have any effect on the truth conditions.

If this is the case, then we expect that *all*, which depends on distributivity, should not be possible with (135). This is of course the case, as we have already seen, and as Taub's generalization predicts.

- (139) *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. Again, adopting Kratzer's hypothesis about the external argument means that there are two places where a D operator in principle could be inserted. The result of inserting D on the voice head is shown in (141), and the result of inserting D on the VoiceP is shown in (142).

(140) The students elected a president.

(141) $\exists e \forall z \exists e' [z \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{role}(z)(e') \ \& \ e' \leq e \ \& \ \text{elect}'(e) \ \& \ \text{pt}(e, \text{a.pres}')]$

(142) $\exists e \forall x \exists e' [x \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{elect}'(e' \ \& \ \text{pt}(e, \text{a.pres}')) \ \& \ \text{role}(e', x) \ \& \ e' \leq e]$

(141) is ill-formed, for the same reason as we discussed earlier. (142) 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 (140) 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 nonmaximality with these predicates is also explained, since nonmaximality is due to the presence of a D operator.

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, (143) is a stative predicate that is ambiguous between a distributive and a collective reading.

(143) 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 nonmaximality, 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.

- (144) The bottles are all too heavy to carry.

As expected, (144) 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 (145)–(146).

- (145) The trees are (all) denser in the middle of the forest. (*att. to B. Partee*)

- (146) The students (all) voted in favor of the proposal. (*att. to B. Ladusaw*)

The sentences in (145)–(146), without *all*, are ambiguous. For example, (145) 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, (146) 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, which was not observed by Dowty, is that *be dense* is a stative predicate, and *vote for* is an achievement predicate. The present analysis predicts that these predicates are not compatible with *all* on their collective readings, so we are not surprised that *all* apparently makes their collective readings disappear; we expect it.

6.3. *Some Exceptions (That Prove the Rule)*

There are some exceptions to Taub’s generalization that, on close inspection, actually provide evidence in favor of the proposal I have made here.²⁰ These are cases where the predicates of Taub’s generalization do accept

²⁰ There are some other 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

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 (147).

- (147) All the students elected a president.

This sentence 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, a D operator inserted on the VP will not be ruled out by economy, because in this context the D operator actually does some work for us. The sentence in (147) would be interpreted as in (148).

- (148) $\exists e \forall x \exists e' [x \subseteq \llbracket \text{the.students}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{elect}(e') \ \& \ \text{theme}(e', x) \ \& \ e' \leq e]$

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 nonmaximality is also due to the presence of a D operator, we expect that (147) should license nonmaximality as well. It does, as (149) shows.

- (149) The students elected a president, but not the fourth graders because nobody would run.

This kind of example provides important evidence about the nature of the condition that licenses insertion of a D operator: it shows that contextual conditions play a role in licensing a D operator. A D operator will only readings, and unfortunately for me, the collective readings with *all* are not as bad as I would predict.

For example, (i) allows a collective reading; this reading is even more salient in (ii).

- (i) The students all own that house.
(ii) The grandchildren all inherited that house.

I don't have an explanation for these exceptions. However, the fact that they 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* 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*.

be useful to us if it helps to distinguish among several possibilities. At its heart, this is clearly an economy condition, but it needs to be fleshed out a little more, as in (150).

(150) **Economy-based condition on insertion of a D operator:**

A D operator is licensed for a predicate P taking a plural argument Y if Y has at least two and as many as n contextually relevant distinct subparts $x_1 \dots x_n$, and $P(x_1)$ or $P(x_2)$ or $\dots P(x_n)$ are live possibilities in the discourse.²¹

What should be clear by this point is that Taub's generalization works, as far as it goes, because collective activities and accomplishments contain hidden distributivity. Collective states and achievements do not normally involve distributivity, but they can in a rich enough context; and when they do, we expect to see the same family of properties that we find with lexically distributive predicates. So the real generalization is about distributivity itself: wherever distributivity is licensed, we expect to find *all*.

6.4. *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 Stechow (1994) *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 (151).

(151) Every girl went to the gym, except for Jackie.

Here *except* subtracts Jackie from the set of girls, which is the restriction of *every*. *All* of course has a similar function – the good fit requirement affects the domain of quantification of the D operator that *all* is associated with. 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

²¹ I assume that atomic individuals are always available as contextually relevant distinct subparts.

argument. Since there is no distributivity, and hence no quantification, the ill-formedness of *except* clauses with these predicates is predicted:

(152) *The boys are a big group, except for Jason.

(153) *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 (154). Recall again our context where the students are holding elections for class president by grade.

(154) 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.

6.5. *Reciprocals*

Recall that we pointed out earlier that reciprocal predicates seem to violate Taub's generalization. Both lexically reciprocal predicates as in (155), and predicates that have been made reciprocal by the addition of the anaphor *each other*, as in (156), allow *all*.

(155) The children all look alike.

(156) The students all recognized each other.

These predicates also allow *except* phrases, whose distribution we have seen parallels Taub's generalization.

(157) The children look alike, except for Paul.

(158) The students recognized each other, except for Jenny.

Since *look alike* is a collective state, and *recognize each other* is a collective achievement, why do they allow *all*? Doesn't this violate Taub's generalization?

It should be clear by now why we have deferred discussion of these examples. We have discovered that the underlying reason for Taub's generalization is that collective activities and accomplishments do contain a hidden distributivity. Since reciprocals also always contain hidden distributivity, regardless of their aktionsart category, we expect that they should be compatible with *all*.

On every analysis of reciprocity that I am aware of (cf. works already cited in section 1), quantification over the parts of a plural is a component. The source of the distributive quantification is not universally agreed upon. For overt reciprocals such as *each other*, Heim et al. (1991) propose that the quantification comes from a decomposition of the anaphor, specifically from the component *each*. But in section 1 we also saw a very brief discussion of whether *each other* has the same "strong distributivity" that *each* has, where Williams (1991) disagreed with HLM.

For lexical reciprocals, the literature takes it to be a given that quantification is part of the lexical meaning, although again there is disagreement about the precise shape of the quantification. For discussion see especially Langedoen (1978), Moltmann (1997), and Dalrymple et al. (1998).

Although there is disagreement about the particulars, what is uncontroversial is that reciprocal sentences always contain quantification, regardless of their aktionsart class. Given the independent evidence for this, we expect that *all* should be compatible with reciprocal sentences. In fact, the analysis given here could be taken as evidence that the quantifier present in reciprocals is a D operator with Cov in its restriction. For a proposal along these lines, see Beck (2000).

7. OTHER CONSEQUENCES

The account I've just elaborated also explains another fact about the distribution of *all* – it explains why *all* can cooccur with collectivizing adverbial phrases such as *together* and *at once*, and why *every* cannot. The data is shown in (159)–(160).

(159) All the planes landed together/in formation/as a group/at once.

(160) *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 turns it into a predicate that will apply only to pluralities: a single plane can land, but it cannot land together.

The difference between (159) and (160) is predicted. We have already seen that a D operator can take lower scope than a quantified DP; since quantified DP's must take scope outside of the scope domain, we expect that *every plane* will not be able to combine with the VP *landed together*. But the D operator can take scope inside the scope domain, and since the "scope" of *all* is determined by the scope of the D operator, we expect that it can cooccur with collectivizing adverbials on the VP.

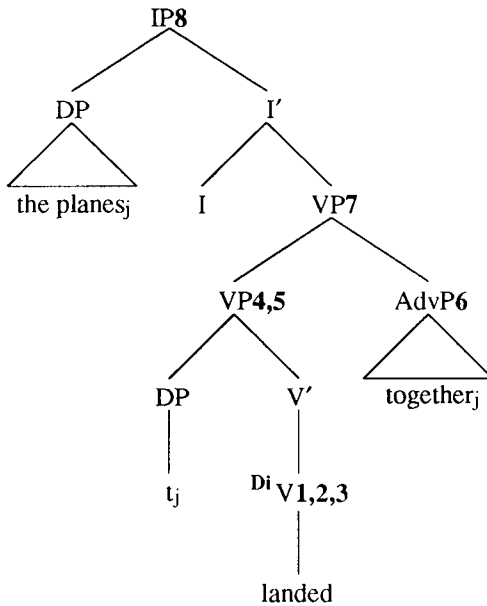
To see how this works, we first need an analysis of collectivising adverbials. I will treat *together* as involving a kind of quantification over individual parts of a plural DP; this is contrary to a proposal by Lasnik (1995) in which the quantification introduced by *together* is over events. However, I borrow from Lasnik 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. Lasnik 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 . τ is a function that maps events to the space-time they occupy. Then we can give the following translation for VP-attached *together*.

$$(161) \quad \lambda P \lambda x \lambda e [P(x)(e) \ \& \ \forall y, y', e', e'' [y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ P(y)(e') \ \& \ P(y')(e'') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \cap \tau(e'')]]$$

This predicate has the following truth-conditional effect. It looks at the subevents of an event e , and requires all of the subevents of e that are P 'ing events by a part of x to be in a relation of spatiotemporal overlap. 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 (159). We have already said that distributivity can apply to a VP. If we take the structure in (162), and apply distributivity at the level of V, then the derivation can proceed as given in (163). (Since *land* is intransitive, we could equivalently apply D to the lowest VP.) (Note that I assume that the trace t_j contributes a variable x_j 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.)

(162)



- (163)
1. $\lambda x \lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, x)]$
 2. $\lambda P \lambda x \lambda \Pi \{P(z) | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
 3. $\lambda x \Pi \{\lambda x \lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, x)](z) | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
 $= \lambda x \Pi \{\lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, z)] | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
 4. $\Pi \{\lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, z)] | z \subseteq x_j \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}$
 5. $\lambda x_j [\Pi \{\lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, z)] | z \subseteq x_j \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}]$
 6. $\lambda P \lambda x \lambda e [P(x)(e) \ \& \ \forall y, y', e', e'' [y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ P(y)(e') \ \& \ P(y')(e'') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]$
 7. $\lambda x \lambda e [\Pi \{\lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, z)] | z \subseteq x \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}](e) \ \& \ \forall y, y', e', e'' [y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]$
 8. $\lambda e [\Pi \{\lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, z)] | z \subseteq \llbracket \text{the.planes}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket\}](e) \ \& \ \forall y, y', e', e'' [y \in \llbracket \text{the.planes}' \rrbracket \ \& \ y' \in \llbracket \text{the.planes}' \rrbracket \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]$

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).

The function yielded by the generalized distributivity operator applied to the set in line 8 could be equivalently written as (164) (recall that \sqcap of sets of functions of type $\langle v, t \rangle$ introduces existential quantification over parts of an event).

$$(164) \quad \sqcap \{ \lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, z)] \mid z \subseteq \llbracket \text{the.planes}' \rrbracket \ \& \ z \in \llbracket \text{Cov}_i \rrbracket \} \\ = \lambda e''' \forall x \exists e' [x \subseteq \llbracket \text{the.planes}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{land}'(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e''']$$

We can use this equivalence to change line 8 into the expression in (165).

$$(165) \quad \lambda e [\lambda e''' [\forall x \exists e' [x \subseteq \llbracket \text{the.planes}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{land}'(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e'''] (e) \ \& \ \forall y, y', e', e'' [y \in \llbracket \text{the.planes}' \rrbracket \ \& \ y' \in \llbracket \text{the.planes}' \rrbracket \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$$

Lambda conversion turns (165) into (166).

$$(166) \quad \lambda e [\forall x \exists e' [x \subseteq \llbracket \text{the.planes}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{land}'(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e \ \& \ \forall y, y', e', e'' [y \in \llbracket \text{the.planes}' \rrbracket \ \& \ y' \in \llbracket \text{the.planes}' \rrbracket \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$$

We apply existential closure to arrive at the proposition in (167).

$$(167) \quad \exists e \forall x \exists e' [x \subseteq \llbracket \text{the.planes}' \rrbracket \ \& \ x \in \llbracket \text{Cov}_i \rrbracket \rightarrow \text{land}'(e') \ \& \ \text{Ag}(e', x) \ \& \ e' \leq e \ \& \ \forall y, y', e', e'' [y \in \llbracket \text{the.planes}' \rrbracket \ \& \ y' \in \llbracket \text{the.planes}' \rrbracket \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$$

The expression in (167) says that there is an event which has as its subevents a landing event for each plane (modulo nonmaximality), 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”.

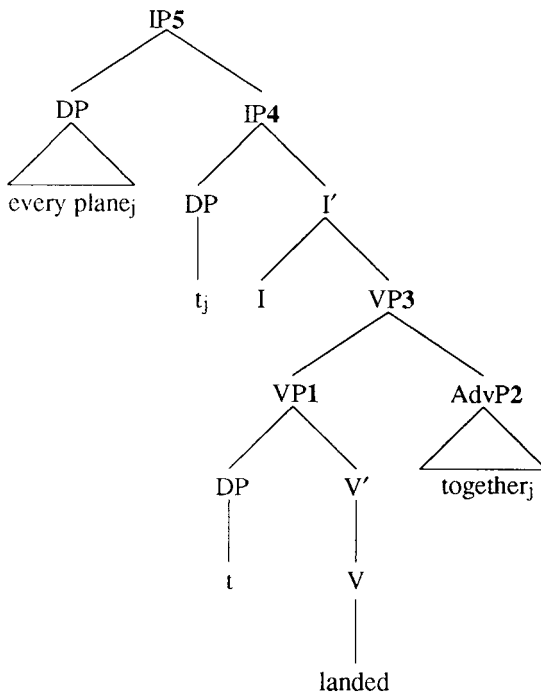
The derivation of (167) has two instances of universal quantification over the parts of $\llbracket \text{the.planes}' \rrbracket$: one introduced by the D operator, and one introduced by *together*. Crucially, though, neither of them has scope over the other, so neither of them interferes with the other.

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 (159) to be (168), 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.

(168)



- (169)
1. $\lambda x \lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, x)]$
 2. $\lambda P \lambda x \lambda e [P(x)(e) \ \& \ \forall y, y', e', e'' [y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$
 3. $\lambda x \lambda e [\text{land}'(e) \ \& \ \text{Ag}(e, x) \ \& \ \forall y, y', e', e'' [y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$
 4. $\exists e [\text{land}'(e) \ \& \ \text{Ag}(e, x) \ \& \ \forall y, y', e', e'' [y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e', y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'', y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$

5. $\forall x[\text{plane}'(x) \rightarrow \exists e[\text{land}'(e) \ \& \ \text{Ag}(e,x) \ \& \ \forall y,y'e''[y \in x \ \& \ y' \in x \ \& \ y \neq y' \ \& \ \text{land}'(e') \ \& \ \text{Ag}(e',y) \ \& \ \text{land}'(e'') \ \& \ \text{Ag}(e'',y') \ \& \ e' \leq e \ \& \ e'' \leq e \rightarrow \tau(e') \circ \tau(e'')]]]$

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*.²²

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.

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 made here, in which *all* has the same low “scope” as the D operator it is associated with.

8. SUMMARY

In this paper I proposed a new account of the nonmaximal nature of distributive quantification. The account proposed is very general, predicts that nonmaximality is possible wherever we find distributivity, and, building on

²² *Together* also has a use that has been called “antidistributive” (Schwarzschild 1992); 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 (i).

- (i) The students built a house together.

It is not clear to me why *together* should have this effect. Yael Sharvit (p.c.) has tried to convince me that the spatiotemporal overlap reading is possible, if one works hard enough to find a context. But I think even if it is possible, it is clear that at least it is 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.

work by Schwarzschild (1996), further aligns the theory of distributivity with other current theories of quantification, in which we find that pragmatics has an effect on the domain of quantification. I then further propose a meaning for the prenominal and floated quantifier *all* in which it does not have any quantificational force of its own, but interacts with the D operator and has the function of ruling out a nonmaximal interpretation for the plural DP that the operator and *all* are both associated with.

Then I introduced an analysis of collective predication in which collectivity is not a uniform phenomenon. Collectivity is widely treated as being the absence of quantification, but I showed this to be true for only a subset of collective predicates. Collective activity and accomplishment verbs actually do involve distributive quantification, and this explains several phenomena, including their ability to cooccur with *all*, the fact that they show nonmaximality, and their ability to cooccur with *except* phrases. The analysis also explains why even the collective and stative predicates, which are normally not interpreted with a D operator, can have a D operator when we interpret these predicates being applied to subsets of a larger group – and correctly predicts that *all* should be compatible with these readings. The long-standing puzzle of “meaning shift”, observed by Dowty (1987), in which *all* appears to take away the collective reading of an ambiguous stative or achievement predicate, also finds a solution here.

Finally, we looked at two other subclasses of collective predicates which apparently ignore the generalization about aktionsart class: reciprocals and verbs with collectivizing adverbials. We found that the quantification present in reciprocals neatly explains their compatibility with *all*. We also found that the account of low-scope distributivity and the interaction of *all* with the D operator explains the differing behavior between *all* DP’s and *every* DP’s with collectivizing adverbials.

REFERENCES

- Bach, Emmon: 1981, ‘On Time, Tense, and Aspect: An Essay in English Metaphysics’, in P. Cole (ed.), *Radical Pragmatics*, Academic Press, New York.
- Beaver, David: 1996, ‘An Infinite Number of Monkeys’, in *Acta Linguistica Hungarica (Special edition devoted to Proceedings of the 5th Hungarian Symposium on Logic and Language)*.
- Beck, Sigrid: 2000, *Exceptions in Relational Plurals*, Talk presented at SALT 10, Cornell.
- Brisson, Christine: 1998, *Distributivity, Maximality, and Floating Quantifiers*, Doctoral dissertation, Rutgers University, New Brunswick.
- Dalrymple, M., M. Kanazawa, Y. Kim, S. Mchombo, and S. Peters: 1998, ‘Reciprocal Expressions and the Concept of Reciprocity’, *Linguistics and Philosophy* **21**, 159–210.

- Davidson, Donald: 1967, 'Causal Relations', *Journal of Philosophy* **64**. Reprinted in *Donald Davidson: Essays on Actions and Events*, Clarendon Press, Oxford, 1980, pp. 149–162.
- Dowty, David: 1979, *Word Meaning and Montague Grammar*, Reidel, Dordrecht.
- Dowty, David: 1987, 'Collective Predicates, Distributive Predicates, and All', in F. Marshall (ed.), *Proceedings of the 3rd ESCOL*, Ohio State University, Ohio.
- Fiengo, Robert and Howard Lasnik: 1973, 'The Logical Structure of Reciprocal Sentences in English', *Foundations of Language* **9**, 447–468.
- Gillon, Brendan: 1987, 'The Readings of Plural Noun Phrases in English', *Linguistics and Philosophy* **13**, 477–485.
- Grimshaw, Jane: 1990, *Argument Structure*, MIT Press, Cambridge.
- Grimshaw, Jane and Sten Vikner: 1993, 'Obligatory Adjuncts and the Structure of Events', in E. Reuland and W. Abraham (eds.), *Knowledge and Language*, Vol. II, Kluwer, Dordrecht.
- Hale, Ken and Samuel Jay Keyser: 1987, 'Explaining and Constraining the English Middle', in Carol Tenny (ed.), *Studies in Generative Approaches to Aspect*, Lexicon Project Working Papers 10, Center for Cognitive Science, MIT.
- 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.
- Johnson, Kyle: 1991, 'Object Positions', *Natural Language and Linguistic Theory* **9**, 577–636.
- Kadmon, Nirit and Fred Landman: 1993, 'Any', *Linguistics and Philosophy* **16**, 353–422.
- Keenan, Edward and Leonard Faltz: 1985, *Boolean Semantics for Natural Language*, Kluwer, Dordrecht.
- Kratzer, Angelika: 1994, 'On External Arguments', in E. Benedicio and J. Runner (eds.), *UMOP Occasional Papers 17 - Functional Projections*, GLSA, Amherst.
- Krifka, Manfred: 1992, 'Thematic Relations as Links between Nominal Reference and Temporal Constitution', in Ivan Sag and Anna Szabolsci (eds.), *Lexical Matters*, Stanford University, Palo Alto.
- Kroch, Anthony: 1975, *The Semantics of Scope of English*, Doctoral dissertation, MIT, Cambridge, MA.
- Landman, Fred: 1989, 'Groups, I', *Linguistics and Philosophy* **12**, 559–605.
- Landman, Fred: 1996, 'Plurality', in S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, Blackwell, Cambridge.
- 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.
- Lasnik, Peter: 1995, *Plurality, Conjunction, and Events*, Kluwer, Dordrecht.
- Lasnik, Peter: 1998, 'Generalized Distributivity Operators', *Linguistics and Philosophy* **21**, 83–93.
- Lasnik, Peter: 1999, 'Pragmatic Halos' *Language* **75**, 522–551.
- Lewis, David: 1975, 'Adverbs of Quantification', in E. Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge University Press, Cambridge.

- Link, Godehard: 1983, 'The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach', in Bauerle et al. (eds.), *Meaning, Use, and Interpretation of Language*, DeGruyter, Berlin.
- McClure, William: 1994, *Syntactic Projections of the Semantics of Aspect*, Doctoral dissertation, Cornell, Ithaca, NY.
- Mittwoch, Anita: 1982, 'On the Difference between *Eating* and *Eating Something*: Activities versus Accomplishments', *Linguistic Inquiry* **12**, 113–122.
- Mittwoch, Anita: 1991, 'Defence of Vendler's Achievements', in C. Vetters and W. Vandeweghe (eds.), *Perspectives on Aspect and Aktionsart (Belgian Journal of Linguistics 6)*, Editions de l'Université de Bruxelles, Bruxelles, pp. 71–85.
- Moltmann, Friederike: 1997, *Parts and Wholes in Semantics*, Oxford University Press, Oxford.
- Parsons, Terence: 1990, *Events in the Semantics of English*, MIT Press, Cambridge.
- Partee, Barbara: 1995, 'Quantificational Structures and Compositionality', in Emmon Bach et al. (eds.), *Quantification in Natural Language*, Kluwer, Dordrecht.
- Partee, Barbara and Mats Rooth: 1983, 'Generalized Conjunction and Type Ambiguity', in Bauerle et al. (eds.), *Meaning, Use, and Interpretation of Language*, DeGruyter, Berlin.
- Pustejovsky, 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.
- Ross, J. R.: 1972, 'Act', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, Reidel, Dordrecht.
- Schwarzschild, Roger: 1992, 'Together as a Non-Distributivity Marker', in P. Dekker and M. Stokhof (eds.), *Proceedings of the 8th Amsterdam Colloquium*, Institute for Logic, Language, and Computation, University of Amsterdam.
- Schwarzschild, Roger: 1994, 'Plurals, Presuppositions, and the Sources of Distributivity', *Natural Language Semantics* **2**, 201–248.
- Schwarzschild, Roger: 1996, *Pluralities*, Kluwer, Dordrecht.
- Scha, Remko: 1981, 'Distributive, Collective, and Cumulative Quantification', in J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Truth, Interpretation, and Information (GRASS 2)*, Foris, Dordrecht.
- Speas, Margaret: 1990, *Phrase Structure in Natural Language*, Kluwer, Dordrecht.
- Stalnaker, Robert: 1973, 'Presuppositions', *Journal of Philosophical Logic* **2**, 447–457.
- Taub, Alison: 1989, 'Collective Predicates, Aktionsarten and All', in Emmon Bach, Angelika Kratzer, and Barbara Partee (eds.), *Papers on Quantification*, University of Massachusetts at Amherst.
- von Fintel, Kai: 1994, *Restrictions on Quantifier Domains*, Doctoral dissertation, University of Massachusetts at Amherst.
- Westerståhl, Dag: 1985, 'Determiners and Context Sets', in Johan van Bentham and Alice ter Meulen (eds.), *Generalized Quantifiers in Natural Language*, Foris, Dordrecht.
- Williams, Edwin: 1991, 'Reciprocal Scope', *Linguistic Inquiry* **22**, 159–173.
- Yoon, Y.: 1996, 'Total and Partial Predicates and the Weak and Strong Interpretations', *Natural Language Semantics* **4**, 217–236.

University of Pennsylvania

Philadelphia, PA

USA

E-mail: brisson@babel.ling.upenn.edu