Plurals, Derived Predicates, and Reciprocals*

Uli Sauerland
MIT

1 Introduction

This paper addresses the syntax and semantics plurals, and then applies it to reciprocal expressions. In the course of this investigation, I address two problems for the conventional view that a reciprocal makes essentially the same semantic contribution to the sentence as other noun phrases, but has an interesting internal structure. I will show that both problems are properties of plurality in general, and can be successfully explained along these lines. As a result, the paper is more about plurality in general than reciprocals though the goal of the paper is to account for the two problems relating to reciprocals.

Let me start with the conventional account of reciprocals, and then formulate the two problems for this account. To see that reciprocals have a more interesting semantics that reflexives and other pronominals, look at the examples in (1): The sentence with the reflexive *themselves* can be paraphrased by replacing the reflexive with its antecedent as shown in (1a). For the reciprocal in (1b), however, this paraphrase would be quite inaccurate, and a correct paraphrase that doesn't use any pronominal expressions can only be given by using two conjoined sentences, as shown in (1b).

(1) a. John and Mary photographed themselves.
John and Mary photographed John and Mary.
b. John and Mary photographed each other.
John photographed Mary and Mary photographed John.

One quite successful and attractive line to explain the semantics of reciprocal expressions, reduces the problem posed by (1b) to that of the semantics for (2). Theoretically, this approach attempts to account for reciprocals by means of distribution of the antecedent, as marked by the *each* in (2), and a complex lexical entry for *each other*. This is attractive, if the semantics for (2), in turn, can be ultimately stated by just using 'ordinary' variable binding, as it is invoked in the explanation of reflexives and other pronominals.

*This paper is based on an unpublished paper of mine from 1995, which I partially rewrote in early 1998. The main change in the theoretical position taken is that I have adopted the position of Schwarzschild (1992) with respect to the ⋆-operator in section 2. During the long gestation period of this research quite a number of people have given me useful comments. I would like to thank Danny Fox, Irene Heim, Roger Schwarzschild, Wolfgang Sternefeld, Yoad Winter, Kazuko Yatsushiro, and all others for their help. I am also grateful to Danny Fox, Irene Heim, and Lisa Matthewson for reading through a draft of this version of the paper. Sections of this paper was presented at the MIT LingLunch Series, at ESCOL '94 at the University of South Carolina, WECOL 1994 at UCLA, and at CONSOLE III at the University of Venice. All remaining errors are of course my own.

The Interpretive Tract. MIT Working Papers in Linguistics 25, 177–204
U. Sauerland and O. Percus (eds.)

(2) John and Mary each photographed (each of) the other(s).

The most well known analysis of this type is the proposal of Heim *et al.* (1991a) that the *each* morpheme undergoes covert syntactic movement from its surface position to the antecedent of the reciprocal. A slightly different line, first pursued in Roberts (1991), proposes that reciprocal sentences force the presence of a phonologically null distributivity operator adjoined to the antecedent of the reciprocal, which has a semantics similar to the overt *each* in (2).

Neither of these two options postulates any syntactic operations specific to reciprocals, since both *each*-movement (Safir and Stowell 1987) and a null distributivity operator (Link 1983, Roberts 1987, Lasersohn 1995) have found independent support. Moreover, the semantics of a sentence like (2) is quite easily stated precisely as we will see below and, in a way, reduces the complexities of reciprocal expressions to the interaction of known principles. Therefore, I consider it worthwhile defending the approach sketched above, which is the goal of this paper.

The two problems for the above view I know of are exemplified in (3) and (4). The first one, which was first brought up by Sternefeld (1993), who attributes this type of example to Heim (p.c.), are sentences containing a third plural noun phrase such as (3a). (3a) clearly is not accurately paraphrased by (3b); at best (3b) paraphrases one reading of (3a). For example, (3a) is true in a situation where three of the letters were sent one way and the other three letters were sent the other way between two correspondents. But, (3b) is false in the same situation.

(3) a. They wrote these six letters to each other.
b. They each wrote these six letters to (each of) the other(s).

The second type of problematic example is illustrated by (4) from Dalrymple *et al.* (1994a). Again, the reciprocal sentence (4a) is not accurately paraphrased by (4b), where *each* is added to the antecedent of the reciprocal. Since in other cases, namely in (5), the same transformation yields an accurate paraphrase, it is often claimed that the meaning of reciprocity depends on the predicate. (Fiengo and Lasnik 1973, Langendoen 1978, Moltmann 1992, Dalrymple *et al.* 1994a).

(4) a. The children followed each other into the room.
b. The children each followed (each of) the other(s) into the room.

(5) a. The children know each other.
b. The children each know (each of) the other(s).

In this paper, I try to provide a solution to both problems. In particular, I hope to show that both problems are related to similar problems in the area of plurality and that the solutions to these problems also explain the problems with reciprocals. Let me now lay out the strategy for the rest of the paper.

For the understanding of plurality I have to assume as a background, I rely mostly on the insights of Roger Schwarzschild's work on the topic (Schwarzschild 1991, 1992, 1994). Especially the $\star$-operator of Schwarzschild (1994), which is similar to the independently developed proposal in Sternefeld (1993, 1998), is going to be important.

One use of the $\star$-operator is in an account of the general 'vagueness' observed with sentences that contain more than one plural noun phrase. This ‘vagueness’ is illustrated by (6a), which is true if, for each of the children, he or she sent only a few of the letters to only some of the adults (cf. Kroch 1974 and Scha 1984). This is a surprising fact, because intuitively plural subjects often behave like universal quantifiers. But, (6b) clearly doesn’t allow the ‘vague’ interpretation possible with (6a). In section 2, I’ll summarize Schwarzschild’s approach to plurality and to examples like (6) in a form that uses a slightly different ontology, and also repeat some of his arguments.

(6) a. The children sent these letters to the adults.
 b. Every child sent these letters to the adults.

Section 3 makes a new proposal concerning cumulative readings, namely, the claim that in many cases the ‘cumulative’ readings require the formation of derived binary predicates. For this purpose, I propose that movement always creates a function-argument structure that is visible to the syntax, with the moved phrase usually corresponding to the argument (cf. Cooper 1979, Heim and Kratzer 1998). Then subsequent movements can target a position between the function and the argument and thereby create a structure with two arguments.

In section 4, I address the Sternefeld’s problem for the semantics of reciprocals illustrated by (3) above. The intuition of Sternefeld (1993, 1998) is that the ‘vagueness’ observed in (6) is also at the heart of first problem. This intuition I believe is correct, but I disagree with Sternefeld about the details. What I show, is that a solution to the first problem follows quite straightforwardly from the theory of plurals laid out in section 2 and traditional assumptions about the reciprocal. One adjustment of the view presented above, however, is argued for, namely, that the reciprocal must be a definite expression, despite the overt appearance of it being universally quantified (cf. Heim *et al.* (1991b)).

Finally in section 5, I present a pragmatic account of the effect of different predicates that Dalrymple *et al.* (1994a) observed. Again, I hope to show that only the traditional assumptions about reciprocals are needed. Added to this picture is a new pragmatic principle, which I call *benevolence*. This will allow a selective weakening of propositions that otherwise contradict common world knowledge.

2 Plurality and Generalized Distributivity

At least since the debate among Link (1983, 1991), Landman (1989), and Schwarzschild (1991), any discussion of plurals must begin with clarifying the ontology and thereby the reference of plural noun phrases. I adopt the position that the referents of plural noun phrases are made up of singular entities by combining them with the mereological union operation $\oplus$. So e.g. *John and Mary* corresponds to the plural entity John$\oplus$Mary.[1] I assume that all plural individuals formed by means of $\oplus$ also belong to the domain of type *e*. Calling the type *e* that of individuals is now somewhat misleading because both singular and plural entities are contained in this type-domain, but I will continue with this usage.

[1]The use of mereological concepts rather than set-theoretic ones, might superficially be seen as a disagreement with Schwarzschild’s (1991) position. In fact, the use of mereological lattices incorporates the key postulate of Schwarzschild’s union theory, that the semantics of NP-conjunction is associative, into the ontology. In addition, I don’t need to introduce *Quine’s invention* to avoid the difference between an individual and the singleton set containing that individual, that set-theory would allow us to make.

2.1 Distributivity

The example in (7a) illustrates what is known as the ambiguity between a distributive and a collective interpretation. On the collective interpretation, (7a) is judged to be true when *the men* is used to refer to two men who each weigh 150 lbs. Therefore this understanding of (7a) roughly corresponds to (7b). On the distributive interpretation, (7a) is made true by a situation where two men that each weigh 300 lbs. are the only salient men around.

(7) a. The men weighed 300 lbs.
b. The men together weighed 300 lbs.
c. The men each weighed 300 lbs.

There is a straightforward way of expressing the perceived difference offered by the ontology laid out above. Let's look at this under the assumption that *the men* are John and Bill. Then, in the collective situation, the predicate *weigh 300 lbs.* is true of the plural individual John⊕Bill, and in the distributive situation, it is true of John and is also true of Bill. Looking at this way of expressing the distinction, what seems to be different in the distributive situation is the contribution the subject makes to the meaning. Indeed, an ambiguity of the subject has been postulated to assign to distinct representations to (8a) (Bennett 1974). However, because of sentences such as (8) where the subject allows a distributive construal in one conjunct and a collective one in the other, this view was given up (see Roberts 1987, Schwarzschild 1994, and Lasersohn (1995) for discussion). Instead, the reference of *the men* in the example is nowadays assumed to be always the same plural individual, on our assumptions John⊕Bill.

(8) The men weighed 300 lbs. (each) and lifted the piano (together).

How can the predicate *weigh 300 lbs.* be true of John⊕Bill? What we need to say, is that predicates are also true of plural individuals if these individuals are the mereological sum of smaller individuals that the predicate is true of. Then, since *weigh 300 lbs.* is true of John and is also true of Bill it will also be true of John⊕Bill. Expressing this intuition, Link (1983), Schwarzschild (1994), Sternefeld (1993) and others all define operators similar to the $\star$-operator defined in (9).[2]

(9) $\star F$ is the F' such that
a. For all x: if $F(x)$, then $F'(x)$
b. For all x, y: if $F'(x)$ and $F'(y)$, then $F'(x \oplus y)$
c. For any function F'' that satisfies a. and b.:
$$\forall x\text{: if } \star F'(x) = 1, \text{ then } F''(x) = 1$$

An easy way to understand the idea of $\star$-operator is to look at predicates as the sets they are the characteristic functions of. I will call this set the *set-extension* of a predicate. So for example, if the predicate F is *weigh 300 lbs.* and John and Bill each weigh 300 lbs., its set-extension is {John, Bill}. Now, the predicate $\star F$ will be true of John⊕Bill as well as of John and Bill. Then, its set-extension is {John,

[2] As far as I can see, the differences between my definition in (9) and the definitions of others are entirely due the different assumptions about plurals adopted, but without empirical import.

Bill, John$\oplus$Bill}. In general, looking at the set-extensions, the $\star$-operator assigns to any set S the smallest superset closed under the mereological sum operation $\oplus$.

Notice that the 'starred' predicate is still true of the individuals the original predicate was true of. This raises the possibility that the 'starred' predicate is for all semantic purposes the meaning of the predicate. So on this view, which is the one advocated by Schwarzschild (1994), the distinction between a distributive and a collective construal is not relevant in the syntax. The alternative view is to propose that insertion of the $\star$-operator is optional, and therefore all one-place predicates are ambiguous, with respect to whether the $\star$-operator is applying or not. This view postulates an ambiguity in the LF-representation between the distributive and collective construal; therefore Schwarzschild (1994) calls it the ambiguity view.

Schwarzschild (1994) argues against the ambiguity view. One argument he discusses is that, as the number of men under discussion goes up, the number of such readings goes up as well. E.g. with three men, (9a) can also be judged true, if two men together weigh 300 lbs. and the third man on his own weighs 300 lbs. Admittedly, (9a) seems quite weird as description of such a situation under normal circumstances, but imagine the following circumstances: A long line of men is waiting in front of two elevators. One elevator has only 300 lbs. capacity, the other one 400 lbs. Your job is to arrange the men in groups such that the two elevators are used as efficiently as possible. So, you have them tell you their exact weights and group them accordingly. Then, the men weighing 300 lbs. stand on the left, the men weighing 400 lbs. stand on the right. (See Schwarzschild 1991, 1992 for discussion of similar examples)

A second argument for Schwarzschild's view is the based on his example (72), given in (10a). Consider (10) to be a command given by a head mobster to Beasly. According to Schwarzschild (1994), the order cannot be fulfilled but only ensuring that either those guys don't win with a group ticket or by ensuring that they just don't win individually. Example (10b), due to Danny Fox (p.c.), makes the point, but with an assertion: (10b) is only judged true if there was a celebration after each of their individual lottery wins as well as the wins where two or more of them shared a lottery ticket. Since in other cases we judge a sentence true if any reading of it makes it true (cf. Abusch 1994), the ambiguity view incorrectly predicts (10a) to be true in the described situation without any restriction on the context it appears in. (11a) and (11b) attest that indeed the sentences have the weaker truth conditions when *together* is inserted to enforce the collective 'reading'.

(10) a. Beasly, better make sure those guys don't win the lottery this week!
b. Whenever those three guys win the lottery, there's a celebration.

(11) a. Beasly, better make sure those guys together don't win the lottery this week!
b. Whenever those three guys together win the lottery, there's a celebration.

On Schwarzschild's (1994) theory, the problem with (10) is of a different nature; it's correctly predicted to be false in the situation described above; but to capture the fact that (10b) is true in the context where it's clear that only the guys common wins are of interest, and that (11) are true in the situations from above is not predicted yet. To account for the former fact, Schwarzschild (1994) proposes that the $\star$-operator is sensitive to context. To this purpose he adds a restriction to salient individuals to (9), which on the formalization I use results in (12). In a context where only the plurality John⊕Bill⊕Jack, assuming these are the three guys names, is salient, the set-extension of ⋆*win the lottery* will be the empty set, even if the set-extension of *buy three apples* includes the singular individuals John and Bill.

(12) $\star F$ is the F' such that
  a. For all **salient** x: if $F(x)$, then $F'(x)$
  b. For all x, y: if $F'(x)$ and $F'(y)$, then $F'(x \oplus y)$.
  c. For any function F'' that satisfies a. and b.: if $\forall x : \star F'(x) = 1$, then $F''(x) = 1$

2.2 Collectivity

The problem addressed in this section is illustrated by the example in (13) , and was noted with (11) as well. The fact is that (13) is false in a situation where John weighs 250 lbs. and Bill does as well. (13) seems to pose a problem for Schwarzschild's (1994) proposal that the ⋆-operator applies to all predicates, because that means giving up the distinction between a distributive and a collective interpretation of the VP. But, adding *together* as in (13) forces a collective interpretation, and (13) isn't true in a distributive situation, where John weighs 250 lbs. and Bill does, too. The modified definition of the ⋆-operator could provide a way of dealing with (13). However, because (13) is even false in a discourse where the singular individuals John and Bill are salient, such an approach seems to be on the wrong track.

(13) John and Bill together weigh 250 lbs.

Schwarzschild's account of collectivity intuitively says that *together* requires the subject of a predicate to be a plurality and that the predicate be 'exactly' (without a ⋆-operator) true of this plurality. We need the concept that a predicate is true of a plural individual only by means of the ⋆-operators applying. Since Schwarzschild's (1994) way of providing this concept is quite technical, I am cutting some corners here. Let me define a second meaning for constituents, the 'picky' meaning, denotated as $[\![]\!]^{picky}$. The value of $[\![]\!]^{picky}$ is defined exactly like the ordinary meaning $[\![]\!]$ except for the case of the ⋆-operator. The new meaning function, $[\![]\!]^{picky}$, treats the ⋆-operator as semantically vacuous, as defined in (15).

(14) $[\![\star\text{XP}]\!]^{picky}$ is equal to $[\![\text{XP}]\!]^{picky}$

(15) $[\![\text{together}]\!](\text{VP})(x)$ is defined only if x is a salient plurality
  $[\![\text{together}]\!](\text{VP})(x) = 1$ if and only if $[\![\text{VP}]\!]^{picky}(x) = 1$

Let us see how this works by applying the definitions in a situation where John and Bill each weigh 250 lbs. Then the $[\![]\!]$-extension of ⋆(weigh 250 lbs.) is {John, Bill, John⊕Bill}, but the $[\![]\!]^{picky}$-extension of the same predicate is just {John, Bill}, because here the ⋆-operator isn't taken into account. Therefore, for the whole VP *together weigh 250 lbs.*, we get the following meaning, because John⊕Bill is not in the $[\![]\!]^{picky}$-extension of *weigh 250 lbs.*:

(16) $[\![$together weigh 250 lbs.$]\!]$ is defined for the plurality John⊕Bill
  $[\![$together weigh 250 lbs.$]\!]$(John⊕Bill) $= 0$

Notice, that even applying the $\star$-operator to the VP after *together* applied would not affect its meaning, because the singular individuals John and Bill aren't in the domain of the VP after *together* applied. This predicts that if the $[\![]\!]^{picky}$-extension of the predicate includes salient pluralities, *together* is not incompatible with a distributive construal that distributes over such salient pluralities. Example (17) confirms this prediction, where the pluralities John$\oplus$Bill and Mary$\oplus$Sue are made salient by the extra uses of *and*.

(17) John and Bill and Mary and Sue $\star$[together $\star$[weigh 250 lbs.]]

2.3 Codistributivity (or Cumulativity)

As mentioned above, sentences that contain more than one plural noun phrase seem to show even more flexibility or vagueness in their truth conditions than the combinations of distributive vs. collective construals of the individual noun phrases would predict: Kroch (1974) and Scha (1984) noticed that a sentence like (18a) can be true in a situation where each woman faces only one of the men. A paraphrase like (18b) seems to capture this construal of (18a), which Scha (1984) refers to as the *cumulative* reading. Because I adopt the view that Scha's reading is really a form of a distributive construal, as explained below, I call the relevant interpretation the *codistributive* construal.

(18) a. The women face the men.
b. For each of the women there is a man who she faces, and for every man there is a woman who faces him.

Before I present Schwarzschild's account of codistributivity, let me address factor interfering with the judgements: Sauerland (1994) and Winter (1997) point out that in examples where the second NP is a definite description as in (18a) the relevant interpretation can be achieved by binding of an implicit variable. More precisely, if the first NP in (18a) is construed distributively and the second NP contains an implicit variable bound by the first, (18a) should be interpreted like (19). In fact, Winter (1997) claims that this is the only source of codistributive readings.

(19) The women (each) face their man.

Disagreeing with Winter (1997), I believe that there are codistributive readings independently of variable binding. For one, it is possible to get codistributive readings in examples like (20a), without needing a discourse context that establishes a functional dependency between the mafiosi and the policemen. A second argument is based on a test Winter (1997) proposes. Look at (20b). The second noun phrase in (20b) is a conjunction of names and therefore doesn't lend itself to a binding analysis. Nevertheless, (20b) allows a codistributive reading.[3]

[3]Winter (1997) points out that a codistributive reading is impossible for (ia) and thereby concludes that codistributivity arises only via implicit variables. In my judgement, (ia) can receive a codistributive reading in a special context, e.g. if it is the goal of a game to separate John from two women using a wall and simultaneously separate Bill from the two other women using the wall. The fact that the codistributive reading isn't readily available in (i), in my opinion, has to do with the discussion of pragmatics in section 5, especially at the end of the section: The reasoning there is that a predicate like *marry* easily gets a codistributive reading because world knowledge rules out stronger readings. But, with a predicate like *separate*, a codistributive reading needs to motivated by the context.

(20) a. The mafiosi shot the policemen.

b. I know the Smith brothers married Sue, Jill, and Sarah.
But, I don't exactly know who married whom.

So, what is Schwarzschild's account of codistributivity? Sternefeld (1993) and Schwarzschild (1994) point out that Scha's cumulative reading can be subsumed under the concept of a distributive construal. Intuitively, these readings seem to involve distribution over two arguments 'in parallel'. More formally, Sternefeld (1993) and Schwarzschild (1994) propose to deal with these examples using a general distributivity operator $\star$ that applies to predicates of n-arguments. I define this operator here for functions that take n-arguments of type e.[4] Notice that the $\star$-operator defined in (9) is the special case $n = 1$ of the definition in (21).[5]

(21) For F of type $\langle \underbrace{e, \langle e, \ldots, \langle e}_{n\text{-times}}, t\rangle \ldots\rangle$, $\star F$ is the function such that:

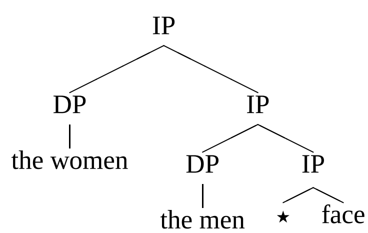
a. $\forall x_1, \ldots, x_n$: if $F(x_1)\cdots(x_n) = 1$, then $\star F(x_1)\ldots(x_n) = 1$

b. $\forall x_1, \ldots, x_n, y_1, \ldots, y_n$: if $\star F(x_1)\cdots(x_n) = 1$ and $\star F(y_1)\cdots(y_n) = 1$, then $\star F(x_1 \oplus y_1)\ldots(x_n \oplus y_n) = 1$

c. For any function F' that satisfies a. and b.:
$\forall x_1, \ldots, x_n$: if $\star F(x_1)\cdots(x_n) = 1$, then $F'(x_1)\cdots(x_n) = 1$

Using the generalized distributivity operator we can represent the codistributive reading of the sentence in (18-a) as follows:

(22)



(i) a. John and Bill are separated from Mary, Sue, Ann, and Ruth by a wall.

b. John and Bill are separated from Mary and Sue, and Ann and Ruth by a wall.

Winter (1997) also points out that for a sentence like (ib) a codistributive reading is possible, which he refers to a *respectively* reading and suggest to relate it to wide scope conjunction. I would like to argue for Schwarzschild's (1996) analysis to treat *respectively* as codistributivity with additional restrictions on the context, but at moment I don't understand Winter's (1997) suggestion enough to do so. But, whatever the analysis of *respectively* is, this cannot be the explanation of (20-b): Adding *respectively* to (20-b) makes the continuation *I don't know who married whom* non-sensical. But, without *respectively* this continuation is possible. Therefore, I believe (20-b) has a true codistributive reading.

[4]Since I use a functional type theory the definition given here looks more complicated than the underlying idea. If we again look at the sets of n-tuples that the functions we're defining $\star$ for are the characteristic functions of, and extent the definition of $\oplus$ to tuples such that $(a_1, \ldots, a_n) \oplus (b_1, \ldots b_n)$ is defined as $(a_1 \oplus b_1, \ldots, a_n \oplus b_n)$, the $\star$-operator again closes a set of n-tuples under the sum operation $\oplus$.

(i) For a set M of n-tuples let $\star M$ be the smallest set M' with $M \subset M'$ and $\forall a, b \in M'$: $a \oplus b \in M'$.

[5]The definition here doesn't yet incorporate the context-sensitivity of (12). We will come back to this issue in section 5.

Let us check that it is indeed true in a situation where Mary faces John, Carol faces Martin, and Lucy faces Tim, nobody faces anybody else and these are all the men and women present. The crucial step of the calculation is the application of the $\star$-operator given in (23). This adds to the denotation of the two-place predicate *face*, amongst others, the pair where the first component is the group of the women and the second the group of the men. Hence the sentence (18-a) is true in the described situation.

(23) $\star[\![face]\!] = \star\{$(Mary,John),(Carol,Martin),(Lucy,Tim)$\}$
$= \{$(Mary$\oplus$Carol$\oplus$Lucy, John$\oplus$Martin$\oplus$Tim),
(Mary$\oplus$Carol, John$\oplus$Martin$\oplus$Tim), (Mary$\oplus$Lucy, John$\oplus$Martin$\oplus$Tim),
(Carol$\oplus$Lucy, John$\oplus$Martin$\oplus$Tim), (Mary$\oplus$Carol$\oplus$Lucy, John$\oplus$Martin),
..., (Mary,John),(Carol,Martin),(Lucy,Tim)$\}$

2.4 Quantification over Plurals

A second place where according to Schwarzschild (1994) the ‘picky’ meaning, that was introduced for *together*, plays a role are downward entailing quantifiers like *less than two*. A sentence like (24a) can be true if John and Bill ate one apple each, even though in the same situation its negation could also be judged true. The latter is expected because for the plurality John$\oplus$Bill there is a plurality of apples $a_1 \oplus a_2$ such that $\star$*eat* is true of them. For (24a) to be true, a lexical entry for *less than two* that makes reference to the ‘picky’ interpretation, such as the one in (24b), is helpful.

(24) a. John and Bill ate less than two apples.
b.[6] $[\![$less than two$]\!]$(R)(N) $= 1$ if and only if
there is no plurality x with $\#x \geq 2$ and $R(x) = 1$ and $[\![N]\!]^{picky}(x) = 1$

Once we look at quantifiers like *exactly two*, it becomes apparent that a quantifier can be true either because of the ‘picky’-meaning or the $\star$-meaning of its restrictor. We can assume therefore that all quantifiers obey a definition schema like the one in (25). Obviously with upward entailing quantifiers the interpretation with a ‘picky’ restrictor will always entail the truth of the starred restrictor interpretation, whereas with downward entailing quantifiers the implication will be the reverse.

(25) $[\![\text{Q(NP)(VP)}]\!] = 1$ if and only if $[\![\text{Q}]\!]([\![\text{NP}]\!])([\![\text{VP}]\!])$ or $[\![\text{Q}]\!]([\![\text{NP}]\!])([\![\text{VP}]\!]^{picky})$

Notice here that (24a) is in fact very similar to (10-a), repeated in (26). However, Schwarzschild (1994) proposes a different account for the two examples; one applying domain restriction the other using ‘picky’ meanings. The difference between the two examples is that (26) is only judged true in a special context, where the individuals aren’t salient, whereas (24a) is always judged true. Whether this really is sufficient motivation for two distinct ways to deal with sentences

[6] I ignore here the formal problem that $[\![N]\!]^{picky}$ is not really defined. What is meant here is the result of applying $[\![]\!]^{picky}$ to the syntactic constituent that corresponds to the nuclear scope of the quantifier *less than two.*

where a distributivity construal might occur in a downward entailing environment is not clear to me. For now, I'll stick with Schwarzschild's proposal.[7]

(26) John and Bill didn't buy three apples.

3 Binary Predicate Formation via Movement

In this section, I show that some instances of codistributivity require a change in the scopal order of the arguments of a predicate via the application of syntactic movement. The argument is based on the prediction of the ⋆-operator approach to codistributivity. Namely, the approach predicts that whenever two noun phrases in a sentence exhibit a codistributive reading, they must be arguments of the same predicate at some level. This is a prediction, because the ⋆-operator can only apply to predicates.

What I will show here is that it is not always sufficient to apply the ⋆-operator only to the lexical predicates. I such cases, I claim the relevant predicates are created by movement. However, unlike in other instances of movement the predicates that need to be formed for codistributivity are binary predicates. This, I argue requires a particular view of movement; namely, movement must create a predicate-argument structure visible to the syntax. Then, a second movement can target a position between the predicate and the argument created by the first instance of movement, thereby creating a two-place predicate.

3.1 The Need for Movement

The example we're concerned with in this section is given in (27).[8] (27) is true in a situation where each child gave exactly one coin to exactly one street artist.[9]

(27) The children gave exactly one coin to the street artists.

The relevant interpretation of (28) seems to involve distribution over both the children and the artists, but not over the coins. Therefore, we would want to apply the ⋆-operator to the two-place predicate *give exactly one coin to*, in the same fashion as we did with *buy* in the previous section. However, this predicate is not readily available, because it has been argued by e.g. Larson (1988) that the predicate *give* applies to the goal object *to the street artists* first, then to the theme object, and then to the subject.[10] Therefore, the only predicates that the ⋆-operator could apply to are the ternary predicate *give*, the binary predicate *give to the street*

[7]Schwarzschild (1996) doesn't discuss this this issue very much, but contains one passage on pages 88–89 that seems to suggests a departure from the 'picky' meanings as pointed out to me by Irene Heim (p.c.).

[8]Winter (1997) discusses a similar example, which he attributes to Dorit Ben-Shalom.

[9]I'm being sloppy here, in using definite plurals despite the argument, mentioned above, that with definites a bound construal of the second NP is an additional source of such an interpretation. However, the relevant interpretation is also observed in (i), which doesn't allow a binding construal.

(i) John, Bill and Joe will give exactly one flower to Sue, Mary and Sarah.

[10]The arguments for this base order, which I take to reflect order of predication, are based on scope and binding evidence. Even if the base order among the two objects was the opposite, the point made in the text could be made with (i):

(i) The children gave the small coins to exactly one juggler.

artists, and the unary predicate *give exactly one coin to the street artists*. This is shown in (29). In (29), I assume that all arguments of the verb moved once to a clausal position, so that quantifiers can be interpreted without type shifting (Cresti 1995 refers to this as θ-saturation).

(28) $\lambda x\star[\lambda y\star[\lambda z\star[\star[\star[\![\text{give}]\!](z)](y)](x)([\![\text{the artists}]\!])]([\![\text{exactly one coin}]\!])]([\![\text{the children}]\!])$

The representation in (28) is not true in the relevant situation because the predicate that applies to *exactly one coin* is ⋆*give to the artists*, but in the relevant situation there's no single thing such that it was given to the plurality of the artists. Only the plurality of all the coins involved in the transaction will fulfill this predicate, but this entity doesn't have exactly one element. Therefore, *the artists* needs to take scope over *exactly one coin*. Since such scope shifting is generally accomplished by syntactic movement, the representation in (29), where *the artist* has undergone further movement, is the candidate to look at next.

(29) $\lambda x\star[\overbrace{\lambda z\star[\lambda y\star[\star[\star[\![\text{give}]\!](z)](y)](x)([\![\text{exactly one coin}]\!])}^{=F}]([\![\text{the artists}]\!])]([\![\text{the children}]\!])$

However, (29) is also not true in the relevant situation. Look at the predicate marked as F. F is true of all z which were given exactly one coin by x. In the situation we're looking at, for any given x, there'll be only one such z. Therefore, the predicate will not be true of the *the artists* for a particular choice of x, even after the ⋆-operator has applied.

Actually, we want to form a binary predicate *give exactly one coin to*. What went wrong in (29) is that we assumed that movement only creates unary predicates. However, this is by no means a necessary assumption, and in fact it would need to be stipulated on a view where movement creates a predicate which then applies to the moved constituent, as Cooper (1979) and Heim and Kratzer (1998) suggest for considerations of semantic simplicity.[11] On this view, as stated in (30), movement that targets YP—at least XP-movement—creates two new syntactic constituents above YP. The lower one of these, YP′, corresponds to a functional abstract, over the variable left in the position XP originated from. The higher one, YP″, is interpreted as the result of the function YP′ applying to the meaning of the moved constituent XP.

(30) **Movement:** When XP moves from a position inside YP to the sister-position of YP, the following operations take place:

a. XP is replaced with a variable x which doesn't occur yet in YP.

b. A new node YP′ is formed with the two daughter nodes YP and λx.

c. A new node YP″ is formed with the two daughter nodes YP′ and XP.

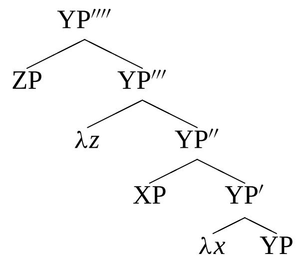
[11] Binary predicate formation via movement between an abstractor and its argument as proposed here might be independently necessary for e.g. multiple exceptives (cf. Moltmann 1995):

(i) Every child gave a coin to every artist, except Julia to the magician.

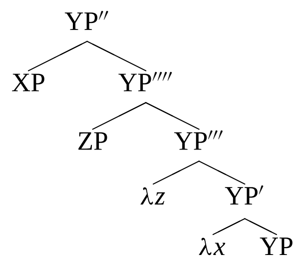
Nissenbaum (1998) presents an additional empirical argument supporting the view of movement presented here from a study of parasitic gap licensing.

Since there are these two constituents YP′ and YP″ created by one instance of movement, we expect that subsequent movement can targeted either of the two. If the higher one is target, only unary predicates are created as shown in (31a), but, if we target the lower one, a binary predicate is created as shown in (31b).[12] To be more precise, YP‴ in (31b) corresponds to a unary function that assigns to an individual another unary function, but this is what semanticists standardly regard as a ‘binary’ function, and there’s an isomorphism between functions of this type and true binary functions. This isomorphism is known as Schönfinkelization or Currying (cf. Heim and Kratzer 1998).

(31) a.



b.



In (29), what we always did is target the higher one of the constituents created by movement. But, as we see now, this was a mistake because we require a binary predicate. In (32), we see the representation we get by moving *the street artists* between the moved phrase *the children* and its abstractor.

(32) $\star\,\overbrace{\lambda z \lambda x \star[\lambda y \star[\star[\star[\![\text{give}]\!](z)](y)](x)([\![\text{exactly one coin}]\!])]}^{=F}([\![\text{the artists}]\!])([\![\text{the children}]\!])$

Indeed (32) is true in the situation we’re looking at: The binary predicate marked as F is true of an x and z if x gave exactly one coin to z. Therefore, it’s true for any one child x and any one artist z in the situation we’re concerned with. Then, applying the $\star$-operator to this binary predicate yields a binary predicate that is true for all the children and all the artists.

3.2 Evidence for Movement

The claim that codistributivity must involve movement if the codistributed predicate is not a basic predicate predicts that codistributive readings are restricted to cases

[12] In Richards (1997) syntactic evidence for movement to a position below a phrase that has moved previously—*tucking in* is the term Richards uses—is presented.

where such movement is possible. This section corroborates this prediction.[13]

One obvious question to ask here is about the locality conditions of codistributivity. The prediction is that the availability of the codistributive interpretation obeys the same locality restrictions that quantifier raising in other cases obeys. For quantifier raising the consensus in the literature is that it is largely clause-bound, although not all the judgements are unproblematic. We would hence expect to find the same clause-boundedness with respect to codistributive interpretation. As the contrast in (33) shows this prediction is in principle borne out, although the data is not always so clear (see also Winter 1997). In (33), imagine a situation where Sue and Linda are two of the participants of a game which can only have one winner. Then, John and Bill must have not understood the game for (33b) or (33c) to be true, but (33a) can still be true.

(33) a. John and Bill expected Sue and Linda to win.
b. #John and Bill expected that Sue and Linda would win.
c. #John and Bill had the expectation that Sue and Linda would win.

The second test in (34a) shows that a codistributive reading can force wide scope. (34a) cannot be true in a situation where John had the expectation that the winner would be male without having a specific boy in mind, and Bill had the expectation the winner would be female. Rather, a codistributive construal of (34a) requires that *a boy* and *a girl* take scope above *expect*. This is predicted if the codistributive reading can only arise from a representation like (34b), where the the ECM-subject and the matrix subject are coarguments of a derived binary predicate. Crucially, the derived predicate and therefore the ECM-subject must have scope over *expect*, because otherwise binding of the the subject trace of *expect* is impossible. Notice that (34c) allows a reading where the ECM-subject takes scope below *expect*, as does (34a) on a reading where John and Bill have the same expectation.

(34) a. John and Bill expected a boy and a girl to win.
b. $\star[\lambda y \lambda x$ expect(win(y))(x)](a boy and a girl)(John and Bill)
c. John expected a boy and a girl to win.

A third test can be drawn from the fact that a bound variable must be in the scope of its binder. In (35), this should block the movement of the plural *the teachers who liked her*$_i$ to form the predicate necessary for a codistributive interpretation with the subject. Indeed it seems that a codistributive construal is blocked: (35) cannot be true in a situation where John introduced exactly one girl to the math teacher, and Bill introduced exactly one girl to the physics teacher.

(35) The men introduced exactly one girl$_i$ to the teachers who liked her$_i$.

4 Reciprocals and Distributivity

This section shows how Sternefeld's (1993) puzzle mentioned in the introduction can be explained by combining the view of distributivity from the previous section with a conventional view that the reciprocal itself has a complex structure and its

[13] I thank Danny Fox for providing the examples in this section to me.

antecedent is interpreted distributively (Heim *et al.* 1991a, Roberts 1991). The problem Sternefeld (1993) raises, repeated here briefly from section 1, is that on the conventional view of reciprocals the logical form representation of (36a) is similar to (36b). But, in (37) (repeated from (3)), a similar representation doesn't seem to give the right result.

(36) a. The students know each other.
  b. The students each know (each of) the other(s).

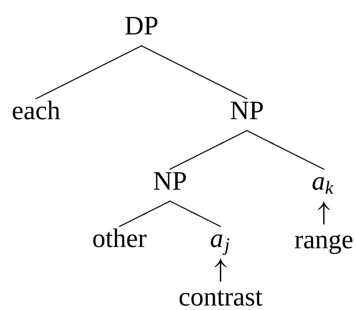
(37) a. They wrote these six letters to each other.
  b. They each wrote these six letters to (each of) the other(s).

In this section, I will first give the semantics of a simple reciprocal sentence, such as (36a) in the first subsection. There, I adopt from Roberts (1991) that reciprocals have a complex structure involving two bound variables, but differ from her in how the binding of the second argument is accomplished. In the next subsection, I will discuss some evidence for the two pronominal elements, as well as a restriction on their binders. Finally, in the third subsection, I address Sternefeld's puzzle.

4.1 Basic reciprocal sentences

The internal structure of the reciprocal that I assume (until the revision in (52)) following Roberts (1991) is shown in (38). It can be paraphrased as: *each one other than himself$_j$ amongst them$_k$*. The two arguments of *other* in (38) are called the *contrast* argument a_j and the *range* argument a_k. Even for reciprocal elements with a different morphological shape, like *one other* in English, I assume that they have exactly the same complex structure. [14] However, I show below that reflexives that can also be true in a reciprocal situation differ from true reciprocals structurally, as discussed below.

(38)



The semantic interpretation of *each* and *other* in this structure does not differ from that of *each* or *other* when they are occurring independently. Their lexical entries are given in (39a) and (39b). In (39a), x^e is the contrast argument of *other*, and y^e is the range argument. In addition, *other* comes with the presupposition that the

[14] Underlying this belief is my hope that general constraints on potential lexical items in human language force any expression that has the complex referential properties of a reciprocal, to have the complex structure Roberts (1991) proposes for *each other*. This assumption seems to be confirmed by the fact that a reciprocal-anaphor with a quite different morphological analysis like Chichewâ *an* shows exactly the same behavior as English *each other* (Dalrymple *et al.* 1994b).

contrast argument x^e is a part of the range argument y^e, as witnessed by the fact that *a boy other than Mary* is odd.

(39) a. $[\![\text{other}]\!](x^e)(y^e)(z^e) = 1$ if and only if z^e is part of y^e and z^e is not equal to nor a part[15] of x^e

b. $[\![\text{each}]\!](X^{et})(Y^{et}) = 1$ if and only if $\forall z$(z is a singular individual and $X^{et}(z^e) = 1 \Rightarrow Y^{et}(z^e)$)

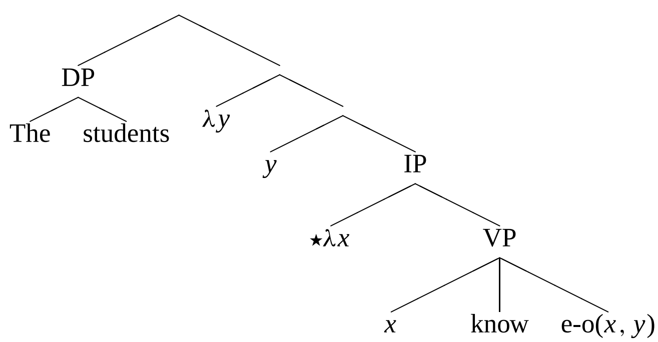
In the following I will abbreviate the structured representation in (38) with *e-o*(a_c, a_r), where a_c is the contrast argument and a_r is the range argument.

One straightforward syntactic argument in favor of such a complex lexical entry is presented by Yatsushiro (1997). She observes that reflexives and reciprocal in Japanese behave differently with respect to the Chain Condition of Rizzi (1986). The relevant corollary of this condition rules out a configuration where the trace of A-movement is c-commanded by a cobound pronominal expression, which doesn't c-command the moved phrase. Yatsushiro (1997) observes that in Japanese reflexives are sensitive to this condition, but not reciprocals. This she explains by claiming that the reciprocal has a complex representation, such that the pronominal parts of it don't c-command the relevant A-trace. The reflexives, in contrast, have a simple structure and therefore violate the corollary of Rizzi's condition.

Look now at the first example of a reciprocal sentence in (40a). I assume that the LF-representation of (40a) is the one given in (40b), modulo the following two simplifications: One, given the lexical entry for *each*, the reciprocal itself is a quantifier phrase, and therefore might have to undergo quantifier raising for type resolution when it occurs in the object position, but this is not represented in (40b).[16] Two, I assume following Schwarzschild (1994) that in all instances of predication the $\star$-operator is applied to the predicate. However, in (40b) only that $\star$-operator is shown, that really plays a role in describing the truth conditions of (40a).

(40) a. The students know each other

b.



[15]The *nor a part of*-part of the lexical entry for *other* is new here compared to Heim *et al.* (1991a). It is needed in examples like (i), where the contrast argument, taken from the context, is a group of two people, not just a single individual.

(i) Two of the three students live in Cambridge. The other student lives in Somerville.

[16]Actually, I argue below that the *each*-morpheme of the reciprocal is not interpreted as *each*, but rather as a definite determiner.

The most important element of the LF in (40b) is how the contrast and range argument of *other* are bound. Following most work in formal semantics, I assume that binding of a variable by its antecedent requires that the antecedent undergo movement, which will create an abstractor, which then can be coindexed with the variable to be bound. What is new about (40b) however, is that two variables are bound by the same antecedent, namely *the students* binds both the contrast and the range argument of *other*. Therefore, I assume that *the students* moves twice, creating two λ-predicates. Because the lexical entry of *other* presupposes that the contrast argument is a true part of the range argument, the two arguments cannot be cobound. Rather, the contrast argument must be bound distributively, whereas the range argument must be bound collectively for the presupposition to be satisfied. Then, the contrast argument can refer to one of the students, while the range argument refers to the total group of the students, such that the complete expression $other(a_c, a_r)$ is true for a z if z is a subgroup of all the students a_r, but a_c is not part of z.

Heim *et al.* (1991a) present two kinds of evidence for the claim that the lexical entry of *each other* contains two variables which are usually bound from the same position; one distributively, and one collectively. One kind of evidence argues that generally an anaphoric element can be related to its antecedent distributively or collectively. (41a) from Higginbotham (1985), which Heim *et al.* (1991a) call the puzzle of grain, makes this point. A second kind of evidence argues that the two variables within *each other* can sometimes be bound from different positions. Examples showing this are discussed in the next subsection. As for (41a) we predict that *they* can have three different antecedents in the sentence (41a): In (41b), *they* is cobound with the contrast argument of the reciprocal, which is distributively related to its antecedent. Therefore, (41b) is true if John told Mary that John should leave, and Mary told John that Mary should leave. In (41c), *they* is cobound with the range argument of the reciprocal, and therefore the sentence is predicted to be true if John told Mary that John and Mary should leave, and Mary told John that John and Mary should leave. Finally in (41d), *they* is bound by the reciprocal, and therefore true if John told Mary that Mary should leave, and Mary told John that John should leave. Since all three readings are attested, (41a) argues that in fact, the prediction that three binders are available for *they* is borne out.[17]

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

b. $[\lambda y[[\star\lambda x[[\lambda z[x \text{ told } z \text{ that } x \text{ should leave}]](\text{e-o}(x, y))]](y)]](\text{John}\oplus\text{Mary})$

c. $[\lambda y[[\star\lambda x[[\lambda z[x \text{ told } z \text{ that } y \text{ should leave}]](\text{e-o}(x, y))]](y)]](\text{John}\oplus\text{Mary})$

d. $[\lambda y[[\star\lambda x[[\lambda z[x \text{ told } z \text{ that } z \text{ should leave}]](\text{e-o}(x, y))]](y)]](\text{John}\oplus\text{Mary})$

[17] Heim *et al.* (1991a) could have made the same point—that a distinction between distributively binding and collectively binding is independently needed—by means of the example (i), which doesn't involve a reciprocal.

(i) a. John and Mary send a letter to their parents.

b. $[\lambda y[[\star\lambda x[\text{send a letter to } x\text{'s parents}]](y)]](\text{John}\oplus\text{Mary})$

c. $[\lambda y[[\star\lambda x[\text{send a letter to } y\text{'s parents}]](y)]](\text{John}\oplus\text{Mary})$

4.2 Contrast and Range Binding

The second kind of evidence Heim *et al.* (1991a) provide for the two variables in the lexical entry of *each other* shows that the two variables can sometimes be bound from different positions. The relevant examples are (42a) from Higginbotham (1980) and (42b) from Heim *et al.* (1991a). (42a) can be true if John and Mary both think John and Mary like each other, but it can also be true John thinks that he likes Mary, and Mary thinks that she likes John. The latter interpretation is more clearly evidenced by (42b), where it's the only sensible one.

(42) a. John and Mary think that they like each other.
b. John and Mary think that they are taller than each other.

Heim *et al.* (1991a) claim that (42a) and (42b) are ambiguous with respect to the binder of the range argument of *each other*. If the range argument is bound by *they*, as in (43a), the sentence will only be true if John and Mary both think that John and Mary like each other. But, if the range argument is bound by the matrix subject, as in (43b), the sentence can be true if John thinks he likes Mary, and Mary thinks she likes John.

(43) a. $[\lambda z[z$ think that $[\lambda y\ [[\star\lambda x\ [x$ like e-o$(x, y)]](y)]](z)]]$(John$\oplus$Mary)
b. $[\lambda y\ [[\ \star\lambda z\ [z$ think that $[\lambda x\ [x$ like e-o$(x, y)]](z)]](y)]]$(John$\oplus$Mary)

This argument in favor of the two variables in the reciprocal is strengthened by the observation (Rizzi p.c. to Heim *et al.* 1991a), that in languages where some reflexives allow a reciprocal-like interpretation, the interpretation with wide scope binding of the range argument, (43b), is absent. Such languages are Italian with the clitic *si* in (44a), and German with the reflexive *sich* in (45a).[18] Both Italian and German also have 'real' reflexives that show the same ambiguity in (44b) and (45b) that English *each other* does in (42). The contrast between *true reciprocals* and reflexives remains even when a reciprocal interpretation of the reflexives is forced via adding an adverb translating into *mutually*.[19]

(44) a. #I due pensano di esser-si battuti (reciprocamente).
the two think be-self beaten (mutually)
b. I due pensano di avere prevalso l'uno sull'altro.
the two thought have prevailed the one over the other

(45) a. #Kai und Toni glauben, daß sie sich überragen.
Kai and Toni think that they self be taller than
b. Kai und Toni glauben, daß sie einander überragen.
Kai and Toni think that they each other be taller than

[18] Heim *et al.* (1991a) claim that the clitic-hood of *si* in Italian is responsible for the absence of a non-contradictory reading. But, Dalrymple *et al.* (1994b) point out that Chicheŵa has a reflexive clitic *an* which allows for non-local binding of the range argument of *other*.

[19] Heim *et al.* (1991a) attribute (44) to Luigi Rizzi (p.c.). I thank Alessandro Zucchi (p.c.) for confirming the judgement and providing the example with *reciprocamente*. As for the German example, adding the German translation of *mutually*, *gegenseitig*, does improve the example a little bit, but it still seems worse that the sentence with the 'real' reciprocal.

Notice that these *fake* reciprocals differ in meaning from the *real* ones only in these very complex sentences. Hence it seems very unlikely that these differences could be acquired just as two different non-complex lexical entries with similar but not identical semantics.

On the proposal here the appearance of a reciprocal reading with a reflexive as in (46a) is explained as a codistributive reading of the reflexive anaphor and its antecedent.[20] The representation in (46b) will be true in a reciprocal situation where Kai saw Toni and Toni saw Kai, as is easy to verify. The unavailability of such a 'pseudo-reciprocal' reading in (45a) and (44a) is then predicted, because, as argued above, codistributive readings are restricted by the clause boundedness of quantifier raising.

(46) a. [Kai und Toni]$_1$ sehen sich$_1$ (gegenseitig).
Kai and Toni see self (mutually)

b. [λz [$\star$see $(z)(z)$](Kai$\oplus$Toni)

Notice that in the explanation of (42) the range variable of the reciprocal is bound from outside of its governing category. Given the fact that *each other* must always have a local antecedent, is seems natural to assume that the contrast variable always has to be bound like on ordinary reflexive. This assumption accounts for the anaphoric behavior displayed by the complex *each other*.

However, it is necessary to stipulate an additional restriction on the binding of the range variable as Rooth (p.c. to Heim *et al.* 1991a) observes. If the range variable could be bound from anywhere, we predict a reading for (47) that isn't in fact observed. Namely (47) isn't true in a situation where the women told each of the youngest three of them to give lectures to all the other women. A clearer example of the same type is (47b). Here, the women were definitely wrong in their denial if the youngest two of them each knew the other one of the youngest two. But, if the matrix subject was the antecedent of the range argument of the reciprocal, the women are predicted to be right in a situation where the youngest two of them each know the other, but the youngest one doesn't know one of the older women.

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

b. The women denied that the youngest two of them knew each other.

Descriptively, the generalization is that the range variable must be bound by an NP that also binds the contrast variable, either directly or indirectly. Here, 'binding indirectly' can be defined as the transitive closure of the binding relation (cf. Higginbotham 1983). At this point, this remains as a stipulation.[21]

[20]The absence of such readings with 'self'-anaphora such as the English reflexives, I hope can be explained by making reference to the reflexive marking of Reinhart and Reuland (1993).

[21]Roberts (1991) attempts to get rid of this stipulation by modifying the theory in such a way that the range variable must always be bound from the closest NP in the argument position of a predicate where a distributing $\star$ is inserted. But, even if we grant that a $\star$ may only be inserted in positions where it is needed for the truth of the sentence, her proposal makes the wrong prediction for examples like (i), where the object *each of these books* is the closest NP to *each other* receiving a distributive interpretation.

(i) They gave (each of) these books to each other.

4.3 Sternefeld's Problem

Putting together codistributivity and the account of reciprocals gives us an account of Sternefeld's example (3), which is repeated in (48).

(48) They wrote these six letters to each other.

Consider first the logical form representation in (49), which is similar to that of a simple reciprocal sentence, but with *six letters* being interpreted codistributively with the antecedent of the contrast argument.

(49) $[\lambda y[\star\overbrace{[\lambda z\lambda x[\underbrace{[\lambda v[x \text{ send } z \text{ to } v]](\text{e-o}(x, y))}_{=G}]]}^{=F}(\text{six letters})(y)]](\text{they})$

Assume first that John and Tom are the persons referred to by *they*. Then the representation (49) will be true in a situation where some of the six letters were sent from John to Tom, and the others were sent by Tom to John. Namely, the predicate marked as F is true of one of John or Tom if the other argument of F is something that he sent to the other one of the two. Therefore, $\star F$ is true of John$\oplus$Tom and the mereological sum of any number of letters that were sent from one of them to the other. Therefore, at least the subcase of Sternefeld's (1993) puzzle where the antecedent of *each other* consists of two individuals is solved.

Now consider the the case where *they* refers to a group of three, namely John, Tom, and Bill. As Sternefeld (p.c.) pointed out to me, the representation (49) is not true in a situation where there are six different letters, l_1 to l_6, and the set-extension of *write* is as given in (50): Consider the subformula $G = [\lambda v[x \text{ send } z \text{ to } v]](\text{e-o}(x, y))$ of (49) and assume that v refers to John; for any choice z, G will be false, because there is no letter such that John sent *them* to each of Bill and Tom.

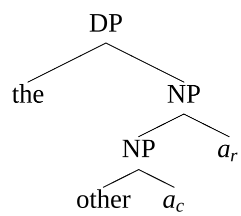
(50) {(John,l_1,Bill), (John,l_2,Tom), (Bill,l_3,John), (Bill,l_4,Tom), (Tom,l_5,John), (Tom,l_6,Bill)}

However, in this interpretation of G we didn't take fully into account that according to Schwarzschild's proposal the $\star$-operator also applies to the predicate *send*. Whether this helps or not depends on whether there is a letter or collection of letters such that John sent them to each of Bill and Tom. What is true, is that John sent $l_1 \oplus l_2$ to Tom$\oplus$Bill. So, the question is whether this makes G true in the relevant case. The answer is no, if the *each* that occurs in *each other* is interpreted as the quantifier *each*. As (51) shows, the quantifier *each* doesn't allow a codistributive construal with a coargument.

(51) John sent these letters to each of Tom and Bill.

Is it really clear that the *each* occurring in the reciprocal forces us to treat reciprocals as distributive quantifiers of themselves? I believe that the occurrence of *each* in the lexical realization of the reciprocal could have other explanations. It might be that as Lebeaux (1983) and Heim *et al.* (1991a) propose, *each* moves away from the *other*-part of the reciprocal. Or, it could be that the occurrence of *each* as part of the reciprocal is just a morphological accident, and the semantic representation of the reciprocal contains just a definite article, as in (52). Then, we are really looking at $G' = [\lambda v[x \star\text{send } z \text{ to } v]](\text{e-o}(x, y))$. If G' is correct, then the question we asked above can be answered with 'yes'. Then, also for values of x other than John, there is a way to make G true, namely with z referring to $l_3 \oplus l_4$ and $l_5 \oplus l_6$ respectively. But, then the predicate F in (49), has the set-extension {(John, l1⊕l2), (Bill, $l_3 \oplus l_4$), (Tom, $l_5 \oplus l_6$)}. Now it's easy to see, after changing the lexical entry for *each other* to that in (52), (49) is predicted to be true in the problematic situation. Therefore, I consider Sternefeld's puzzle solved.

(52)



The revision of the lexical entry for reciprocals in (52), can be understood as expressing the observation that the reciprocal itself is not a universal quantifier. In particular, we observed in (51) that universal quantifiers don't allow codistributive readings, but force a singly distributive interpretation. The reciprocal on the other hand seems to allow a codistributive construal, and thereby patterns with definite noun phrases. It might be possible to gather additional support for this conclusion by looking at examples like (53): The prediction is that (53a) can be true if the children painted one picture each showing all the children except for the child painting it. (53b), on the other hand, requires that for each teacher one picture of him must have been painted. Indeed the judgements seem to go this way, but are not completely clear. In particular, (53a) also could be marginally true in a situation where all children together collaborate on one picture of themselves.

(53) a. The children painted a picture of each other.
b. The children painted a picture of each of the teachers.

Let me briefly compare the proposal above with that of Sternefeld (1998) as I see it. The main difference between the two accounts is that Sternefeld allows the reciprocal to be codistributive with its antecedent. On my proposal the reciprocal could only be codistributive with another coargument, but not with its antecedent. This restriction follows from the assumption that the antecedent binds distributively into the reciprocal, because then the $\star$-operator that distributes the antecedent must c-command the reciprocal. This makes it impossible to achieve coargumenthood for the 'undistributed' antecedent and the reciprocal. Assume for an illustration we wanted to achieve codistributivity between the subject antecedent and the object in (54a). The representation in (54b) shows that in such a representation the contrast variable x of the reciprocal is unbound.

(54) a. The children like each other.
b. $\lambda y[\star[\lambda z \lambda x[x \text{ like } z]](\text{e-o}(x, y))(y)](\text{the children})$

Sternefeld gets around the binding restriction by decomposing the reciprocal into an inequality statement and a reflexive part. The semantic representation he gives for (54a) is given in (55). One way suggested by Irene Heim (p.c.) to derive (55), though not Sternefeld's way, is incorporation of $\neq$ from the reciprocal into the verb, which then is interpreted by means of conjunction. However, the incorporation analysis would fail to explain Yatsushiro's (1997) evidence that reciprocals behave differently form reflexives with respect to the chain condition, because after incorporation this account predicts no difference between reciprocals and reflexives. Furthermore, the incorporation analysis would in a case like (53a) have to postulate incorporation into derived predicates, which contradicts the common syntactic assumption that incorporation targets only heads.

(55) $\lambda y[\star[\lambda z \lambda x[x \text{ like } z \wedge x \neq z]](y, y)]$(the children)

Sternefeld's (1998) actual derivation for (55) is quite different: He proposes that the inequality $x \neq z$ occupies the argument position of *like*. On this view, z is the trace of y, and therefore the underlying form of the reciprocal must of the form $x \neq y$ with x the contrast argument and y the range argument, where the restrictions on their binding properties are needed as on my account: the contrast argument must be bound by a local antecedent, the range argument must be bound by the same DP that in effect binds the contrast argument. In addition, however, Sternefeld needs to account for the fact that the internal argument of *like* must be coindexed with the trace of movement of the range argument, a stipulation that the account above can do without. But, further work is needed to draw the distinction between the two analyses.

5 Capturing Pragmatic Effects

This section addresses the reported variability of truth conditions in basic reciprocal sentences depending on the predicate (Fiengo and Lasnik 1973, Langendoen 1978, Moltmann 1992, Dalrymple *et al.* 1994a). The problem is epitomized by the contrast between the two examples in (56): (56b) is true if the children entered the room in a sequence where each child except for the first followed the child preceeding him or her. (56a), however, is false if each child knew only one other child and there was even one child not knowing any of the children, at least in a null context.

(56) a. The children knew each other. (= (5))
b. The children followed each other into the room. (= (4))

I claim that despite appearance the interpretation of the reciprocal doesn't depend directly on the predicate it appears with. Rather, I hope to show that the meaning of the reciprocal is primarily affected by contextual restrictors and it is the values of the contextual restrictors that are affected by the predicates and other material in the sentence. In particular, I propose that the default preference for null restrictors is overruled by the desire to give an utterance a chance of being true. Therefore, if the lexical properties of the predicate are such that the sentence cannot be true with null restrictors, stronger restrictors are accommodated.

In this section, I introduce two kinds of mechanisms that are sensitive to the context in the first subsection. The first of this mechanism fills a gap left by the definition of the n-ary $\star$-operator in (21) above, where we lost the contextual restriction of the one-place $\star$-operator of (12). In the second subsection, I go on to show how the determination of these two restrictors can account for the facts of Dalrymple *et al.* (1994a) and related facts without reciprocals.

5.1 Contextual Restrictors

Schwarzschild (1991, 1994) showed the great influence of the context on the interpretation of plural noun phrases in general. To give an example, imagine a context where a dance instructor says (57) (cf. (18-a)) to his students. What the instructor expects is that the woman of each couple faces her partner, not just some other man in the room.

(57) For the next dance, the women face the men, please.

While Schwarzschild captures the contextual influence by incorporating it into the definition of the $\star$-operator as we saw in (12), I introduce here contextual restrictors that are separate from the $\star$, as suggested by Heim (p.c.). I assume that contextual restrictors are functions from individuals or tuples of individuals onto truth values, which are true of contextually relevant individuals or tuples of individuals. I represent the contextual restrictors in the semantic representation as free functional variables κ_n. These are adjoined to predicates of the same type and combine with the predicates via predicate modification, which is annotated as $\cap$.

Using this idea we can account for the contextual influence on the interpretation of (57) using the logical form in (58a). If we assume the contextual relevance expressed by the function in (58b), we achieve the desired interpretation for (57).

(58) a. $\star[\kappa_1 \cap [\![\text{face}]\!]]\ ([\![\text{the men}]\!])([\![\text{the women}]\!])$
b. $\kappa_1(x, y) = 1$ if and only if x and y are a couple.

I assume that the κ-restrictors can optionally be inserted above any predicate. Note that these restrictors make a proposition logically stronger or weaker, depending whether they appear in an upward or downward entailing position. Nevertheless, they alone cannot solve the puzzle exhibited by (56) above. Look at (59a), repeated from (56), and its semantic representation in (59b), where two κ-restrictors, κ_1 and κ_{other}, have been inserted.

(59) a. The children followed each other into the room.
b. $[\lambda y[\star\lambda x[\star[\kappa_1 \cap [\![\text{follow}]\!]]([\![\text{the}]\!](\kappa_{other} \cap [\![\text{other}]\!](x)(y)))]](y)]([\![\text{the children}]\!])$

Let's first consider the effect of κ_1. Because it appears in an upward entailing context, it makes the proposition it occurs in logically stronger. Therefore, it clearly won't help us with (59a). κ_{other}, on the other hand, since it occurs in a downward entailing context might help. Since it intersects with *other*, κ_{other} must be a three place predicate of type $\langle e, \langle e, \langle e, t\rangle\rangle\rangle$. A natural restriction, however, would be to tuples of children who entered the room one immediately following the other, without restricting the range argument position. One definition for κ_{other} is given in (60). With this restriction, the subformula $[\![\text{the}]\!](\kappa_{other} \cap [\![\text{other}]\!](x)(y))$ of (59b) has as its value the child immediately preceeding x into the room, if x is not the first child. But, if x is the first child that entered the room, the argument of *the* has the empty set as set-extension. Therefore the existence presupposition of *the* is violated.

(60) $\kappa_{other}(x)(y)(z) = 1$ if and only if x immediately followed z into the room.

One might attempt to remedy the presupposition violation by accommodating it into the restriction of *the children*. This amounts to saying that the extension of *the children* doesn't include the first child. But, it's easy to see that now we get a presupposition violation in the case of the second child.

Therefore, I propose to introduce a second mechanism to 'soften' the meaning of plural predication. Basically, it should say that a predicate is true of a group if it is true of a substantial part of that group. A slightly more formal version of such a functor, I call it ENOUGH, is given in (61). The examples in (62) might also make use of such an operation: (62a) can be true even if a small portion of the animals had a sore throat, and didn't make any noise. (62b) can be true even if John himself is a student but didn't request that he himself leave.

(61) ENOUGH$(P)(y) = 1$ if and only if there is an x such that x is a substantial part of some y and $P(x) = 1$

(62) a. The dogs barked and the cats miaued.
b. John asked the students to leave.

I leave it to the reader to verify that the semantic representation in (63) is indeed true in the situation sketched above.

(63) ENOUGH$(\lambda y[\star\lambda x[[\star[\![\text{follow}]\!]]([\![\text{the}]\!](\kappa_{other} \cap [\![\text{other}]\!](x)(y)))]](y))([\![\text{the children}]\!])$

5.2 Benevolence

The open questions at this point are where these restrictors are inserted and how their values are determined. We can reduce the first question to the latter, if we assume that restrictors whose values haven't been determined yet are true of any value in their domain. I will assume that there are two possibilities for how their values are determined. One is that, as was illustrated by (57), the restrictor reflects what is relevant or prominent in the extralinguistic context. The second possibility is to assume that an appropriate value of the restrictor exists, even though the actual value is not known. This is similar to the mechanism of presupposition accommodation, as it is described in Lewis (1979): In order to keep the conversation going, a participant, even though he does not know the relevant contextual restriction, just assumes the existence of an appropriate restriction.[22] Therefore, I call this process *restrictor accommodation*.

Restrictor accommodation should be restricted in its application like presupposition accommodation to circumstances where it is needed to keep the conversation going. If it could always apply we wouldn't predict the contrast in (56). I propose therefore that restrictor accommodation can only apply if the lexical properties of the sentence are such that there's no reasonable hope for it to be true without restrictor accommodation. I assume that this is a general principle of pragmatics, and state it as such in (64).

[22] As Irene Heim (p.c.), points out the presupposition of *too* cannot be satisfied by accommodating the existence of a state of affairs accommodating it. In the text, I have the existence presupposition of *the* in mind.

(64) **Benevolence:**[23] Assume that an assertion is true in at least one situation compatible with common world knowledge.

Benevolence offers a way to give a pragmatic explanation for the *strongest meaning hypothesis* of Dalrymple *et al.* (1994a). Their generalization is that for a simple reciprocal sentence of the form "Subject Verb *each other*" the reciprocal can be interpreted using one reading out of certain finite set of possible interpretations. The possible readings are ordered according to their logical strength – the number of pairs that are required to stand in the relation denoted by the verb to make the sentence true. However, the speaker also knows that some verbs have logical properties like being *asymmetric* that make them incompatible with the stronger readings. The strongest meaning hypothesis now states that from the possible readings that strongest one is chosen that could be true given the independently known logical properties of the verb.[24] An example of how this works is the problem of (56) this section is addressing. In (65), I give the version of this problem that is discussed by Dalrymple *et al.* (1994a): The contradictory feeling that example (65a) has in contrast to (65b), is explained as the fact that *know* expresses a relation that is not necessarily asymmetric, whereas *follow* expresses an asymmetric relation. Hence for the interpretation of (5) the strongest possible interpretation for the simple reciprocal sentence is chosen; i.e. the one where all pairs of non-identical willow-school-fifth-graders have to stand in the relation *know*. For the interpretation of (65b) however a weaker interpretation of the sentence is chosen because the verb *follow* expresses an asymmetric relation. Hence the claim Harry didn't follow any of his classmates does not contradict the preceeding claim.

(65) a. #The willow school fifth graders know each other, but the oldest doesn't know the youngest.

b. The willow school fifth graders followed each other into the class room, and Harry went first.

Since this statement of the generalization involves real world knowledge, a pragmatic account of it is desirable. The interaction of contextual restrictors and benevolence offers such an account. Benevolence allows restrictor accommodation in (65b), but not in (65a), because there's no possible situation compatible with our world knowledge about *following somebody into a room* where it is true for a group of children that each child followed every other child. The reason for this is simply that *follow into a room* is not a reflexive relation according to our world knowledge. For (66a), on the other hand, there exist possible scenarios compatible with world knowledge that make it true. Therefore, benevolence doesn't apply in the interpretation of (65a).

[23] In earlier versions of this paper, I called this principle *Charity*. Since then, I was informed that this term is already in use in the literature; in fact, for a concept very similar or maybe the same to what I propose here. However, I was not able to find any use of the term *Charity* in print. Therefore, I use a different term for the moment.

[24] The actual formulation of Dalrymple *et al.* (1994a:p. 73) is given in (i):

(i) **The Strongest Meaning Hypothesis:** A reciprocal sentence is interpreted as expressing the logically strongest candidate truth conditions which are not contradicted by known properties of the relation expressed by the reciprocal scope when restricted to the group argument.

One argument in favor of a pragmatic account of this observation is that the effect of the *strongest meaning hypothesis* is absent in a ‘loaded’ context as in (66): The second sentence of (66) can be true, even if only the every resident on the eastern side of Mass. Ave. only knows his neighbors. On my proposal, it doesn’t matter whether benevolence applies in (66), because the actual context provides a restriction that makes the second sentence of (66) true.

(66) Walking down Mass. Ave. from Arlington to Boston the sociologist found out: The residents on the eastern side of Mass. Ave. know each other.

Examples where the antecedent of the reciprocal denotes a group of two individuals pose a problem for the Dalrymple *et al.*’s (1994a) proposal, but seemingly do for my proposal as well: Both proposals apparently predict (67) to be true, because *procreate* is an antisymmetric relation, which for Dalrymple *et al.*’s (1994a) would force a weak reciprocity construal in the interpretation of (67). Also, on the benevolence account pursued here, the antisymmetry of procreation that is part of our world knowledge should allow restrictor accommodation. But, notice that restrictor accommodation alone doesn’t help; in addition, the ENOUGH operator must apply, because only my mother procreated. Therefore, what (67) shows is that my mother alone is not a sufficiently substantial part of my mother$\oplus$me, for the implication from *My mother procreated.* to *My mother and I procreated.* to hold.

(67) #My mother and I procreated each other.

A second argument for Benevolence was brought to my attention by Roger Schwarzschild (p.c.). He observes that an effect similar to the Strongest Meaning Hypothesis is observed with respect to the availability of codistributivity. For example, (68a) is only true out of context if every student knows every professor, but (68b) can be true if one student talked to one professor each. (see footnote 1 and Winter (1996))

(68) a. The students know the professors.
b. The students talked to the professors.

On the approach, we account for this observation under the assumption that a codistributive interpretation is dispreferred, just like restrictor accommodation was dispreferred. For example it might be that a codistributive interpretation always requires covert movement.[25] Then, these movement operations would only be licensed if either common world knowledge coupled with Benevolence or a loaded context requires them, for the sentence to have a chance of being true. Similarly, wide is more easily possible in (69b) than (69a) for most people.

(69) a. The teacher assigned a book to every girl.
b. The teacher gave a book to every girl.

[25]This was not what I assumed above, but it would e.g. follow from a decomposition of predicates in monadic elementary predicates. E.g. Kratzer (1993) proposes such a decomposition of predicates.

6 Conclusion

The analysis of plurality and reciprocity proposed in this paper had the following five ingredients:

- the n-place $\star$-operator that applies to all predicates (cf. Schwarzschild 1992, Sternefeld 1993),
- a concept of meaning 'without the $\star$-operator', here defined as $[\![\,]\!]^{picky}$ (cf. Schwarzschild 1992),
- an operation that allows the formation of derived binary predicates (cf. Scha (1984)),
- a lexical entry for *each other* that contains two variables with one stipulation on the binding of these variables (cf. Heim *et al.* 1991a, Roberts 1991),
- the ENOUGH operator, which can weaken the meaning of plural predicates (cf. Landman 1996).

The ideas underlying these elements on their own aren't novel in the study of plurality and reciprocity, though I sometimes modified or simplified their implementation. In most cases, I summarized the arguments already given in the literature in favor of the above ideas to show that they also argue for the specific implementation given here. The main goal of this paper though is to show is that a coherent picture emerges when these elements are put together. This, in turn, argues for the pieces of the puzzle and the way they're put together.

References

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2.83–135.
Bennett, Michael. 1974. *Some Extensions of a Montague Fragment of English*. UCLA dissertation.
Cooper, Robin. 1979. The interpretation of pronouns. In *Syntax and Semantics. Volume 10*, ed. by F. Heny and H. Schnelle, 61–92. New York: Seminar Press.
Cresti, Diana. 1995. *Indefinite Topics*. Cambridge: Massachusetts Institute of Technology dissertation.
Dalrymple, Mary, Makoto Kanazawa, Sam Mchombo, and Stanley Peters. 1994a. What do reciprocals mean? In *Proceedings of SALT IV, University of Rochester*, 61–78.
Dalrymple, Mary, Sam Mchombo, and Stanley Peters. 1994b. Semantic contrasts and syntactic similarities between Chicheŵa and English reciprocals. *Linguistic Inquiry* 25.145–163.
Fiengo, Robert, and Howard Lasnik. 1973. The logical structure of reciprocal sentences in English. *Foundations of Language* 9.447–468.
Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
Heim, Irene, Howard Lasnik, and Robert May. 1991a. Reciprocity and plurality. *Linguistic Inquiry* 22.63–101.
——, ——, and ——. 1991b. On "Reciprocal scope". *Linguistic Inquiry* 22.173–191.
Higginbotham, James. 1980. Reciprocal interpretation. *Journal of Linguistics Research* .
——. 1983. Logical form, binding, and nominals. *Linguistic Inquiry* 14.679–708.
——. 1985. On semantics. *Linguistic Inquiry* 16.547–593.

Kratzer, Angelika. 1993. On external arguments. In *University of Massachusetts Occasional Papers 17*, 103–130. Amherst: GLSA.

Kroch, Anthony S. 1974. *The Semantics of Scope in English*. Cambridge: Massachusetts Institute of Technology dissertation.

Landman, Fred. 1989. Groups, I. *Linguistics and Philosophy* 12.559–605.

——. 1996. Plurality. In *The Handbook of Contemporary Semantic Theory*, ed. by Shalom Lappin, chapter 16, 425–457. Oxford: Blackwell.

Langendoen, D. Terence. 1978. The logic of reciprocity. *Linguistic Inquiry* 9.177–197.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19.335–391.

Lasersohn, Peter. 1995. *Plurality, Conjunction and Events*. Dordrecht: Kluwer.

Lebeaux, David. 1983. A distributional difference between reciprocals and reflexives. *Linguistic Inquiry* 14.723–730.

Lewis, David K. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8.339–359.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice theoretical approach. In *Meaning, Use, and the Interpretation of Language*, ed. by R. Bäuerle, C. Schwarze, and A. von Stechow, 302–323. Berlin: de Gruyter.

——. 1991. Plural. In von Stechow and Wunderlich (1991), chapter 19, 418–440.

Moltmann, Friederike. 1992. Reciprocals and *same/different*: Towards a semantic analysis. *Linguistics and Philosophy* 15.411–462.

——. 1995. Exception sentences and polyadic quantification. *Linguistics and Philosophy* 18.223–280.

Nissenbaum, Jon. 1998. Movement, derived predicates, and the null operator construction: An interpretive approach to parasitic gap licensing. In *MITWPL 25. The Interpretive Tract*, ed. by Uli Sauerland and Orin Percus. Cambridge: MITWPL. (this volume)

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24.657–720.

Richards, Norvin. 1997. *What moves where when in which language?*. Cambridge: Massachusetts Institute of Technology dissertation.

Rizzi, Luigi. 1986. On chain formation. In *The Syntax of Pronominal Clitics*, ed. by Hagit Borer, 65–96. New York: Academic Press.

Roberts, Craige. 1987. *Modal Subordination, Anaphora, and Distributivity*. Amherst: University of Massachusetts dissertation.

——. 1991. Distributivity & reciprocal distributivity. In *Proceedings of SALT I*, 209–229. Ithaca: DMLL Publications.

Safir, Ken, and Tim Stowell. 1987. Binominal each. In *Proceedings of NELS 18*, 426–450. Amherst: GLSA.

Sauerland, Uli. 1994. The representation of reciprocals in grammar. In *Proceedings of the Eleventh Eastern States Conference on Linguistics (ESCOL '94)*, ed. by Janet Fuller, Ho Han, and David Parkinson, 270–281. Ithaca: DMLL Publications.

Scha, Remko. 1984. Distributive, collective, and cumulative quantification. In *Truth, Interpretation, and Information, GRASS 2*, ed. by Jeroen Groenendijk, Martin Stokhof, and Theo Janssen. Dordrecht: Foris.

Schwarzschild, Roger. 1991. *On the Meaning of Definite Plural Noun Phrases*. Amherst: University of Massachusetts dissertation.

——. 1992. Types of plural individuals. *Linguistics and Philosophy* 15.641–675.

——. 1994. Plurals, presuppositions and sources of distributivity. *Natural Language Semantics* 2.201–248.

——. 1996. *Pluralities*. Dordrecht: Kluwer.

Sternefeld, Wolfgang. 1993. Plurality, reciprocity, and scope. SfS-Report 13-93, Universität Tübingen, Tübingen.

——. 1998. Reciprocity and cumulative interpretation. Ms., Universität Tübingen.

von Stechow, Arnim, and Dieter Wunderlich (eds.) 1991. *Semantik: Ein internationales Handuch der zeitgenössischen Forschung* (Semantics: An International Handbook of Contemporary Research). Berlin: Walter de Gruyter.

Winter, Yoad. 1996. What does the strongest meaning hypothesis mean? In *Proceedings of SALT 6*. Ithaca: CLC Publications.

——. 1997. Dependency and distributivity of plural definites. Utrecht University Ms.

Yatsushiro, Kazuko. 1997. VP-structure in Japanese revisited. In *Proceedings of JK8*. Stanford. CSLI. to appear.

Uli Sauerland
Department of Linguistics and Philosophy
MIT, Room E39-249
Cambridge MA 02139

sauerlan@mit.edu
http://web.mit.edu/sauerlan/www/