This paper investigates different readings of plural and reciprocal sentences and how they can be derived from syntactic surface structures in a systematic way. The main thesis is that these readings result from different ways of inserting logical operators at the level of Logical Form. The basic operator considered here is a cumulative mapping from predicates that apply to singularities onto the corresponding predicates that apply to pluralities. Given a theory which allows for free insertion of such operators, it can then be shown that the lexical semantics of the reciprocal expressions each other/one another consists of exactly two components, namely an anaphoric variable and a non-identity statement. This receives further support from the observation that it is exactly these two components that can be focussed by only; all that remains to be done is to correctly manipulate these components at the level of LF.

1. Introduction

The semantic properties of expressions containing reciprocal elements and their antecedents have figured prominently in a number of recent discussions of linguistic theory. Nevertheless, these properties are not particularly well understood, despite recent pioneering efforts ... T. Langendoen (1978)

Langendoen – arguably the first to account for the “logic of reciprocity” in a really insightful way – identifies an intimate connection between what he calls “elementary plural relational sentences” (henceforth EPRSs) like (1a) and “elementary reciprocal sentences” (henceforth ERSs) like (1b):

(1) a. John, Barbara, and Cecile had relations with Jane and Bill.
    b. John, Barbara, and Cecile had relations with one another.

Langendoen’s analysis of the different readings of EPRSs suggests that these ambiguities correspond to ambiguities in ERSs. For example, given that EPRSs have the general format \(ARB\), one may distinguish between strong and weak distributivity as analyzed in (2), and analogously between strong and weak reciprocity as shown in (3):

(2) a. \((\forall x \in A)(\forall y \in B)(x R y)\) (SD)
    b. \((\forall x \in A)(\exists y \in B)(x R y) \land (\forall z \in B)(\exists w \in A)(w R z)\) (WD)
His article leaves open, however, as to how this parallelism can be accounted for in a formal theory, and how the above formulae can be derived from syntactic surface structures in a systematic and compositional manner.

In this paper I propose a simple way of bringing these semantic representations closer to the syntax of natural language. I will start with the traditional plural operator *, defined by Link (1983) as follows: For any \( C \), *\( C \) is the smallest property such that:

\[
\begin{align*}
\text{(4) a. } & P \subseteq \ast P, \\
\text{b. } & \text{if } a \in \ast P \text{ and } b \in \ast P, \text{ then } a \oplus b \in \ast P.
\end{align*}
\]

Usually, \( a \oplus b \) is understood as the mereological sum of \( a \) and \( b \). In section 2 I will follow Schwarzschild (1991; 1996) in adopting an alternative interpretation of \( \oplus \), which makes our plural semantics independent of mereological assumptions. In either case, note that for any \( X \) it holds that \( X \in \ast X \) and \( X \subseteq \ast X \).

The analysis of EPRSs then proceeds on the basis of the “cumulation” of a relation. This can be defined as the two-place counterpart of *, which is due to Krifka (1986): For any two-place relation \( CA \), let **\( CA \) be the smallest relation that satisfies (5):

\[
\begin{align*}
\text{(5) a. } & CA \subseteq \ast C, \\
\text{b. } & \text{if } \langle a, b \rangle \in \ast C \text{ and } \langle c, d \rangle \in \ast C, \text{ then } \langle a \oplus c, b \oplus d \rangle \in \ast C.
\end{align*}
\]

Assuming that \( R \) ranges over pairs of individuals of a certain domain \( D \) and presupposing that \( A \) and \( B \) are subsets of \( D \), it is clear that WD, as defined in (2b), is equivalent to \( A \subseteq \ast A \).

Weak reciprocity (WR) can now be analyzed by means of weak distributivity, as shown in (6):

\[
\begin{align*}
\langle A, A \rangle \in \ast \{ \langle x, y \rangle : \langle x, y \rangle \in R \land x \neq y \}
\end{align*}
\]

Again it can be seen that (6) is equivalent to (3b). I will return to SD and SR later.

Comparing the expressions of WD and WR in terms of **, it follows that the reciprocal NP is analyzed as making two contributions to the meaning of the whole sentence: Firstly, it expresses an anaphoric dependency of the object on the subject at the level of plural entities; secondly,
it states the non-identity between the subject and the object at the level of individuals. This treatment of reciprocals crucially differs from previous ones (cf. Chomsky (1973), Dougherty (1974), Lebeaux (1983), Heim, Lasnik and May (1991)) in not attributing any quantificational force to the meaning of the reciprocal NP. In particular, the quantification often associated with each of each other is not derived from a process like each-movement (which, in any case, seems unmotivated for the otherwise synonymous expression one another), but is analyzed here as part of the semantics of pluralization in general.

The analysis of reciprocals thus depends on an explicit semantics for EPRSs as developed in section 2. The general idea is that logical representations for plural sentences are obtained from the surface by augmenting representations at the level of Logical Form. Within the framework of “transparent” Logical Forms, syntactic representations can be enriched by logical material which serves as semantic “glue” to combine the meanings of constituents in a compositional way. In this paper the semantic operation * enters LF in two different ways. Firstly, * serves as the translation of the plural morpheme on nouns. Secondly, * (as well as **) can be attached to verbal projections as semantic glue, i.e., without presupposing any one-to-one correspondence to morphology. Accordingly, verbal plural morphology has no direct semantic interpretation at LF. Rather, operators like * and ** can be inserted freely whenever needed to express a particular thought.

For example, an iterated application of * will yield the representation in (7), which logically differs from \( A, B \) \( \in **R \) and in fact expresses SD as defined in (2a):

\[
(7) \quad A \in *\{x : B \in *\{y : R(x, y)\}\}
\]

In this way many of the semantic ambiguities of plural sentences simply follow from the option of inserting pluralization at different positions at LF.

The groundwork for dealing with plural objects and plural predication having been laid in section 2, section 3 discusses reciprocal expressions along the lines indicated above. In particular, our analysis shares one common feature with previous ones, namely that the denotation of the reciprocal NP involves discontinuous parts, and thereby relies on movement at the level of LF. A number of ambiguities will then be resolved as either ambiguities of scope or ambiguities of indexing. Finally, an attempt will be made to deal with the so-called Geach-Kaplan sentence Some critics
admire only one another. It will be shown that such sentences corroborate our theory in a straightforward way.

Section 4 analyzes sentences like *The men and the women had relations with each other.* We discuss whether or not the cumulative semantics proposed in the above analysis should be replaced by a pragmatic theory built on the notion of a cover – an idea pursued by Schwarzschild (1991; 1996). I will argue that although the concept of a cover is basic for some of the pragmatic aspects of his analysis, there is reason to believe that it should not entirely replace the operation of cumulation, which appears to be the more fundamental semantic notion.

2. Elementary Plural Sentences

2.1. Noun Denotations

Schwarzschild (1991; 1996) convincingly shows that plural NPs denote subsets of a domain \( D \) of individuals. For example, assuming that \( a, b, c \) are the girls in \( D \) and *John* is a name for \( j \in D \), *John and the girls* denotes \( \{ a, b, c, j \} \), rather than the two-element set \( \{ j, \{ a, b, c \} \} \). According to Schwarzschild, the denotations of referential (i.e., non-quantificational) plural terms and of singular common nouns are always subsets of \( D \). In consequence, NPs like *the group*, *the committee*, or *the collection* must denote elements in \( D \) rather than sets, so that there must be a “meta-linguistic” membership relation between elements of \( D \). The decision to interpret all noun extensions as sets of individuals enables us to establish a tight relation between plural morphology and the semantic type of nouns: Whereas the extension of a morphologically singular CN is always a set \( X \), the morphologically derived plural noun always denotes a set of sets, namely \( *X \).

It follows that the conjunction of referential NPs is set union. Implementing this theory in a formal way, Schwarzschild adopts “Quine’s innovation,” i.e., a set theory that identifies an individual with the singleton set containing it. Thus, for any *Urelement* \( j \) it holds that \( j = \{ j \} = \{ \{ j \} \} = \{ \{ \{ j \} \} \} = \ldots \). Conjunction of an individual with a set, as in *John and the girls*, proceeds by first identifying \( j \) with \( \{ j \} \) and then applying set union between the denotations of the terms.

Quine’s innovation proves extremely useful for stating semantic operations uniformly. In particular, it enables us to interpret the operator \( \oplus \) in (4b) and (5b) simply as set union. As another application, let us discuss the meaning of *the* when calculating the denotation of *the girls*. Given
the singular CN girl, which denotes \{a, b, c\}, the plural CN girls denotes $\star \{a, b, c\} = \{a, b, c, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$. In order to interpret the determiner, I follow Sharvy (1980) in assuming that the definite article the selects the set with the greatest cardinality (if there is one) from a noun denotation. Due to Quine’s innovation, this works no matter whether the CN is singular or plural. Accordingly, the maximal element of $\star \text{girl}$ is the set \{a, b, c\} again, so that the denotation of the girls is the same as that of girl.

2.2. Elementary Plural Predication

Turning next to the denotation of VPs, it is intuitively clear that a lexical item like gather cannot apply to singularities. However, given that NPs like the committee denote elements in $\mathcal{D}$ (usually called groups), elements of the denotation of gather cannot always be restricted to sets. On the other hand, sentences like John gathers do not make sense and should be classified as category mistakes. We must, therefore, presuppose a sorted domain that distinguishes between group “individuals” and true individuals, both being elements in $\mathcal{D}$. As mentioned above, we also assume a membership-relation which states that true individuals are members of certain groups.

In what follows, however, the sortedness of a domain will be irrelevant. All that counts is that verbal expressions can be predicated of different kinds of entities in $\star \mathcal{D}$. This contrasts with the extension of CNs, whose denotations are neatly divided between singular CNs having denotations in $\mathcal{D}$ and plural CNs having denotations in $\star \mathcal{D} \setminus \mathcal{D}$.

As a standard example of elementary plural predication, consider now Gillon’s (1987) case of a model $\mathcal{M} = \langle \mathcal{D}, f \rangle$, in which Rodgers and Hart, working as a team, and Rodgers and Hammerstein, working as a team, have written the musicals On Your Toes and Oklahoma! respectively. To simplify things a bit, I assume that the predicate write-musicals is a lexical expression (i.e., must not be decomposed into a transitive verb plus an object); moreover, I ignore tenses throughout the paper. Accordingly, the interpretation function $f$ in $\mathcal{M}$, which assigns semantic values to lexical expressions, assigns the set \{\{Rodgers, Hart\}, \{Rodgers, Hammerstein\}\} as the denotation of the VP write-musicals in $\mathcal{M}$. In this model, a sentence (8b) in a context like (8a) is intuitively true.

(8) a. Hammerstein, Rodgers, and Hart are composers.
   b. They have written musicals.
However, the set referred to by they in this context is neither a subset nor an element of the denotation of the predicate. As a way to make (8b) come out true, we may think of predication as being mediated by \textit{cumulation}. But note that cumulation is precisely what has been defined in (4), so that (8b) can be analyzed as \( f(\text{they}) \in *f(\text{write-musicals}) \). This is because \( f(\text{they}) = \{\text{Hammerstein, Rodgers, Hart}\} \in *f(\text{write-musicals}) = *\{\{\text{Rodgers, Hart}\}, \{\text{Rodgers, Hammerstein}\}\} = \{\{\text{Rodgers, Hart}\}, \{\text{Rodgers, Hammerstein}\}, \{\text{Hammerstein, Rodgers, Hart}\}\} \).

Apart from this cumulative reading, (8b) is often said to have a collective and a strong distributional reading as well. In Gillon’s model, these readings are false. Note that one cannot account for the cumulative reading by adding a meaning postulate to the effect that whenever \( \text{write-musicals} \) and \( \text{write-musicals} \), it holds that \( \text{write-musicals} \). On this account, there would be no way to prevent the collective reading from coming out true, contrary to fact. I will return to these readings further below.

2.3. \textit{Lexical Meanings}

Assuming as we have that the denotation function \( f \) assigns arbitrary subsets of \( *D \) to intransitive verbs, we must further restrict the possible denotations of lexical items. Let us first classify predicates as being either “merological” or “individual level” predicates. Mereological predicates denote sets that are closed under the part-whole relation and under cumulation; in particular, it holds that \( P = *P \). Individual level predicates have denotations only as subsets of \( D \), they do not apply to sets but only to true individuals. These will be called \textbf{D-based} predicates. Certain predicates, including all lexical CNs, are D-based by logical necessity; others, like VPs, may be D-based only in certain models.

Consider again the predicate \textit{write-musicals}. Clearly, this predicate is neither mereological nor D-based in Gillon’s model. Moreover, it exhibits a property that is presumably absent from true lexical predicates. To illustrate, assume counterfactually that Rodgers alone wrote \textit{Oklahoma!}. Since \( \{\text{Rodgers}\} \subset \{\text{Rodgers, Hart}\} \), the denotation of \textit{write-musicals} is not subset-free:

\begin{equation}
\text{(9) A collection of sets is subset-free iff it contains no sets } a, b \text{ such that } a \subset b .
\end{equation}

This possibility of not being subset-free seems to be limited to mereological and to certain complex predicates. In other words, my contention is
that all true lexical predicates of natural language are either subset-free or mereological. More precisely, I conjecture that (10) is a universal property of natural language:

(10) Given a lexical predicate denoting an \( n \)-place relation \( R \), the values for \( y \) that satisfy \( R(x_1, ..., y, ..., x_n) \) for any fixed values of \( x_1, ..., x_n \) form a set \( Y \) such that
   a. \( Y \) is subset-free or
   b. \( Y \) is mereological, so that if \( a \subset b \) with \( a, b \in Y \), it follows that \( \{ y : y \in b \land y \notin a \} \in Y \).

To see that the two-place relation write (simpliciter) satisfies (10a), select a certain musical \( c \) and ask whether it can be composed by a group and a subgroup thereof (or an individual) at the same time. To be concrete, assume a model such that \( D_1 \) has been written by \( a \) and \( b \) together, and (in addition) by \( a \), but not by \( b \). Such a situation is highly counterintuitive; in fact, it is ruled out by (10): If we want to claim that \( D_1 \) has been written by \( a \) and \( b \), and if in addition we also want to claim that \( D_1 \) has been written by \( a \), then the first claim can be understood only collectively, and it follows that \( m \) has also been written by \( b \).

2.4. Relational Predication

Some of the standard examples for relational predication originate from Jackendoff (1972) and Scha (1981) and constitute a major challenge for naïve compositionality. Let us begin with (11a), which is based on an example of Scha’s and which, as a first approximation, is paraphrased

---

1 A potential counterexample has been brought to my attention by Graham Katz (p.c.). Assume that Lisa and Sue are playing “May had a little lamb” (together) on a piano, each with one hand, and Lisa is playing it with one hand on a second piano. Imagine furthermore that Lisa is playing the bass on piano \( a \) with the right hand, and the melody on piano \( b \) with the left hand, while Sue adds the bass on piano \( b \). Then in fact Sue and Lisa together play the song (on one piano), and Sue alone plays the same song (on two pianos). Does this really contradict subset-freeness? One difficulty with this example — apart from a certain type-token ambiguity of the result of playing — seems to me that the description of such a state of affairs suggests that what actually goes on are two events that happen to coincide. If such a view is legitimate, and if events figure as arguments of predicates, the example is again consistent with (10). The subset property is indeed satisfied, because each set refers to argument roles in different minimal events or situations, so that one of the variables, namely the event variable is not kept constant. The entire situation described by Katz can then be interpreted as the mereological sum of the two more elementary events.
in (11b):


b. There are 500 Dutch firms and there are 2000 Japanese computers, such that for each firm there is a computer it uses, and for each computer, there is a firm which uses it.

The main difficulty presented by (11a) lies in the fact that (11b) uses six quantifiers, which correspond to only two NPs in (11a). Most importantly, the six NPs occur at various places in the paraphrase, so that their relative scopes cross each other. This seems to exclude a compositional account close to the surface of (11a). But suppose now that we analyze (11) as an instance of WD by using cumulation. It is straightforward that (12) – with DF denoting the set of Dutch firms and JC denoting the set of Japanese computers – expresses the intended truth conditions correctly:\[2\]

\begin{equation}
\exists X \in ^*\text{DF}\, \exists Y \in ^*\text{JC}\,(500(X) \land 2000(Y) \land \langle X, Y \rangle \in **[\text{use}]_s)
\end{equation}

Suppose now that we replace the verb use by own. The relevant difference is that “use” is presumably D-based, whereas own is not, i.e., some firms may only jointly own a single computer. Given our assumptions about noun denotations, WD as defined in (2b) now becomes inapplicable (or false) in such a situation. In fact, the equivalence of WD with \(\langle A, B \rangle \in **R\) holds only if \(R\) is D-based. Therefore, WD cannot work. However, since the difference between D-based and non-D-based predicates is immaterial for the semantics of cumulation, \(\langle A, B \rangle \in **R\) may still hold while WD is false. Accordingly, substitution of use by own in (12) does not necessitate any further change, i.e., the analogue of (12) still yields the correct truth conditions.

It has occasionally been claimed that the cumulative reading of the type exemplified above is likely to arise only with large cardinals; small cardinalities often invite a different analysis, which nonetheless is still non-compositional. One of the first to note the problematic nature of plural predication in this respect was Jackendoff (1972), who assumes that (13a) (= his example 7.55) has the paraphrase (13b):

\[\text{(13) } \langle 500 \rangle \text{ and } \langle 2000 \rangle\]

---

\[2\] As pointed out to me by Manfred Krifka (p.c.), Scha’s analysis logically implies that 500 and 2000 are the maximal numbers of firms and computers that make the sentence true. This aspect of meaning is not dealt with in my analysis, and I leave it open here as to whether it should be built into the semantics of numerals (such that 500 means “at most 500”) or whether it should be dealt with as an implicature.
(13) a. I told three of the stories to many of the men.
b. There is a fixed group of many men and a fixed group of three stories and each of the men heard each of the three stories.

(13b) is called the branching analysis of (13a). If, as many logicians assume, (13b) constitutes a genuine reading on its own, it seems that we need a new analysis. Observe, however, that (13b) is an instance of SD, and as already indicated in section 1, SD can be expressed by two iterative applications of one-place cumulation. To illustrate, assume that MM abbreviate the property of \( CG \) that holds of \( CG \) when \( CG \) is a large set of men. Let us furthermore agree to use all predicate expressions ambiguously, as denoting either sets or characteristic functions thereof. In particular, lambda abstraction may, as a notational convention, be read as set formation. This will greatly increase readability, because we don’t have to wait for functional application at the end of a formula. Consider now (14) as a possible LF of (13a):

\[
\exists X \in MM(\exists Y \in 3)(\text{story}(Y) \land X \in \lambda x[Y \in \text{tell}(I, x, y)])
\]

How does (14) capture the universal quantification expressed in (13b)? Let us define a distributional operator \( \Delta \) as in (15):

\[
\forall x \in X(x \in (D \cap P))
\]

If \( P \) is D-based, then \( *P = \Delta P \). (The converse also holds when \( P \) is subset-free.) On the plausible assumption that \( \text{tell} \) is D-based, \( * \) in (14) can be replaced by \( \Delta \), and by (15) this will allow for rewriting (14) by using standard quantification over individuals, as shown in (16):

\[
\exists X \in MM(\exists Y \in 3)(\text{story}(Y) \land (\forall x \in X)(\forall y \in Y) \text{tell}(I, x, y))
\]

This exactly represents (13b), which is an instance of SD.\(^3\)

Comparing the intuitive paraphrases of the Jackendoff sentence and the Scha sentence, it turns out that the Jackendoff paraphrase is less com-

\(^3\) Note also that SD can be defined by means of \( \Delta \). I leave it open here whether this is indeed necessary, i.e., I do not adjudicate between making \( \Delta P \) a different reading, or making it a special case of \( *P \). The difference matters only for \( P \)s that are not D-linked. Recall that \( \text{They write musicals} \) can be considered false if \( \Delta P \) is a reading. Towards this I have no intuitions to offer.
plex than Scha’s paraphrase as regards quantifier interaction. This lesser degree of complexity has a formal counterpart in the above analysis: Suppose that, by analogy to **, we define a new operator $\Delta\Delta$ with the desired strong truth conditions, so that the analysis of (13a) formally parallels that of the Scha sentences. As evidenced by (15), $\Delta\Delta$ can be defined by two iterations of $\Delta$, whereas ** crucially cannot be decomposed by two iterations of *. Because of this non-reducibility of **, the Scha sentence is more complex than the Jackendoff example — not only in terms of the number of quantifiers involved, but also in terms of definability.

2.5. Plural Predication at LF

I will now indicate how the above analyses can be represented at LF. This can be illustrated with another standard example of relational predication:

(17) Five men lifted two pianos.

Among the many different construals of (17) are the following:

(18) a. $(\exists X)(\text{five}(X) \land *\text{man}(X) \land (\exists Y)(\text{two}(Y) \land *\text{piano}(Y) \land \text{lift}(X, Y)))$

b. $(\exists X)(\text{five}(X) \land *\text{man}(X) \land X \in *\lambda x[(\exists y)(\text{two}(Y) \land *\text{piano}(Y) \land \text{lift}(x, Y))])$

c. $(\exists X)(\text{five}(X) \land *\text{man}(X) \land (\exists Y)(\text{two}(Y) \land *\text{piano}(Y) \land Y \in *\lambda y[\text{lift}(X, y)])])$

d. $(\exists X)(\text{five}(X) \land *\text{man}(X) \land (\exists Y)(\text{two}(Y) \land *\text{piano}(Y) \land X \in *\lambda x[\text{lift}(x, Y)])])$

e. $(\exists X)(\text{five}(X) \land *\text{man}(X) \land (\exists Y)(\text{two}(Y) \land *\text{piano}(Y) \land X \in *\lambda x[Y \in *\lambda y[\text{lift}(x, y)])])])$

f. $(\exists X)(\text{five}(X) \land *\text{man}(X) \land (\exists Y)(\text{two}(Y) \land *\text{piano}(Y) \land \langle X, Y \rangle \in **\lambda x y[\text{lift}(x, y)])])$

That is, we would translate (13a) as (i) and define $\Delta\Delta$ as in (ii):

(i) $(\exists X \in \text{MM} \land (\exists Y \in \exists)(\text{*story}(Y) \land \langle X, Y \rangle \in \Delta\Delta\lambda x y.\text{tell}(I, x, y))$

(ii) $\Delta_1\ldots\Delta_n \ast R := \{\{X_1, X_2, \ldots, X_n\} : X_1 \times X_2 \times \ldots \times X_n \subseteq R \land (\forall x) X_1 \neq 0 \land X, \subseteq D\}$

I hope the intuitive content of (ii) deserves no comment and can be grasped from the above example.
(18a) expresses a purely collective reading, where five men jointly lift two pianos, one piled on top of the other. Since one piano is already heavy enough, this is hard to imagine. (18b) is even less realistic. Here it is claimed that there may be subsets of a set of five men who jointly lift two pianos, one stacked on top of the other. This and the following analyses, except for (18e), mix the group reading with the cumulative reading. The analysis in (18c) is the most plausible one. Here five men jointly lift two pianos, and it is possible that each piano is lifted by five men at a time. (18d) is similar to (18b) but is again scopeless. This time the number of pianos cannot depend on the number of men, and hence the men will lift only two pianos at a time. (18e) is a kind of distributive analysis which implies that we can find subsets of men and of pianos such that each subset of men lifted each subset of pianos. Finally, (18f) is a purely cumulative analysis and hard to paraphrase. It means that each of the five men and each of the pianos is involved in some lifting, but it leaves it entirely open as to how many men (jointly) lift how many pianos.

At this point the question arises of whether these representations should be considered genuine readings of (17) and, if so, how they can be derived from a surface representation like (19):

\[
\text{[IP [NP five men ] [VP [V lift ] [NP two pianos ]]]}
\]

Current theories of the syntax-semantics interface differ somewhat in regard to the modes of semantic combination that are permissible as counterparts to syntactic rules or structures. It is clear from the multitude of readings that no such theory can maintain complete compositionality; in fact, all current systems (except classical categorial grammar) are only weakly compositional in that they make available a narrowly defined repertoire of semantic linking elements. These logical “morphemes” do

---

5 An analysis that enforces this interpretation must, however, be formulated in terms of \( \wedge \) instead of \( * \).

6 For some speakers of English there exist still further readings that rely on a reversal of the scope of the quantifiers, e.g. (i):

\[
(3Y)[\text{two}(Y) \wedge *\text{piano}(Y) \wedge Y \in *\lambda y[(3X)[\text{five}(X) \wedge *\text{man}(X) \wedge \text{lift}(X, y)]]
\]

According to (i) it would be possible that each of the pianos has been lifted by five men, which results in a maximal number of ten men. The derivation of this reading involves quantifier raising, but apart from that, the analysis proceeds analogously to the other examples without raising.
not correspond to particular lexical expressions, nor need there be a
one-to-one correspondence between the linking elements and particular
syntactic constructions. In general, however, syntax and semantics will
conspire and provide sufficient clues to enable one to disambiguate and
recover the semantic glue that has to be posited in order to make semantic
sense of a string of syntactic constituents.

Part of the evidence in favor of semantic glue has been gained from
Scha’s and Jackendoff’s examples. Given that strict compositionality has
proven untenable, there must exist rules that govern the insertion of glue,
i.e., of semantic material not overtly present in surface syntax. Most sys-
tems of semantic glue provide for at least four operations: intersection
(i.e., conjunction), existential quantification of indefinites, predication or
lambda abstraction, and functional application (cf. Heim (1982), Fanselow
(1985), v. Stechow (1993)). We here add another operation, namely plu-
ralization.

I am not concerned here with a precise implementation of rules that
insert semantic glue into syntactic representations, 7 but as an illustration
of the results, consider the Augmented Logical Forms (henceforth ALFs)
that correspond to the interpretations in (19). These are given in (20),
where the *-notation is used like brackets that embrace the variables
affected by pluralization:

\[
\begin{align*}
(20) \ a. \ & IP \ \ NP_1 \ VP \ \ V \ lift(x_1, x_2) \\
& b. \ & IP \ \ NP_1 \ \ *_{x_1} *_{VP} \ \ V \ lift(x_1, x_2) \\
& c. \ & IP \ \ NP_1 \ VP \ \ *_{x_2} *_{V} \ lift(x_1, x_2)
\end{align*}
\]

7 An explicit system has been developed in a previous version of this paper (cf.
Sternefeld (1993)), where it is guaranteed that morphological pluralization on CNs is
* obligatorily interpreted as semantic pluralization, whereas plural morphemes on verbs
are agreement phenomena left without direct semantic impact. Accordingly, * plays a
double role, namely as the semantic interpretation of plural nominal morphology, and
as freely insertible glue elsewhere in the system.
These representations are given a direct model-theoretic interpretation as follows. First note that predicates enter LF as open formulae, as in Heim (1982), so that pluralization can be defined as in (21):

\[(21) \text{ } ^*x_1 \ldots x_n \phi := \langle x_1, \ldots, x_n \rangle \in ^* \ldots ^* \lambda x_1 \ldots x_n.\]

Assuming that all variables range over entities in \(^D\) (which comprises \(D\) as a subset), we can use the same variable inside and outside the scope of the star operation. This will greatly facilitate exposition, and will play an important role in subsequent analyses, although, strictly speaking, it should not be considered an essential feature of the analysis.

Continuing with an analysis of the NPs in (20), two pianos can be represented at ALF as (22):

\[(22) \text{ NP}_3 \]
\[
\text{AP}(x_2) \quad \land \quad ^*x_2N
\]
\[
\text{two} \quad \text{N}(x_2) \quad \text{pl}
\]
\[
\text{piano} \quad \text{s}
\]

NP\(_1\) is analyzed analogously. Accordingly, all branchings in (20) are interpreted as conjunction. As in the standard treatment of indefinites as open formulae (cf. Heim (1982)), we finally add existential quantification as semantic glue at the roots of the trees.\(^8\)

\(^8\) Accordingly, I interpret glue here in the usual syncategorematic and therefore non-compositional way. This contrasts with the treatment of glue adopted in Sternefeld (1997) and in previous versions of this paper where glue is interpreted in a completely compositional way. However, this additional degree of compositionality does not contribute significantly to the particular analysis of plural phenomena proposed here.
In a system like this the rules that generate ALFs automatically generate different analyses for free, hence there is no need to worry about the multiplicity of readings. And even more readings can be generated once we allow for more operators, e.g. $\Delta$, which defines the strictly distributional construal. The relevant point is that there is no empirical evidence against assuming such representations, hence there is no reason not to permit them as possible readings of plural sentences.

3. Reciprocals

As already mentioned in the introduction, Langendoen’s (1978) paper has served as a benchmark for the considerations in this paper. His analysis of WD, repeated in (23a), entails my analysis in (24a), and the reader may verify that the same holds for WR, repeated in (23b) and (24b):

(23) a. $\forall x \in A(\exists y \in B)(x R y) \land (\forall z \in B)(\exists w \in A)(w R z)$
   b. $\forall x \in A(\exists y, z \in A)(x \neq y \land x \neq z \land x R y \land z R x)$

(24) a. $\langle A, B \rangle \in \text{**} \lambda x y[R(x, y)]$
   b. $\langle A, A \rangle \in \text{**} \lambda x y[R(x, y) \land x \neq y]$

We have also seen above that under current assumptions (24a) does not entail (23a) for relations $R$ that are not D-based. Likewise, (24b) does not entail (23b). Discussing a model with $A = \{a, b, c\}$ and a relation $f(R) = \{\langle a, b \rangle, c), \langle c, a \rangle, \langle c, b \rangle\}$, Langendoen indeed observes that, although it seems quite natural to say that the $As$ relate to each other, (23b) cannot apply in such a situation, i.e., the truth conditions of the reciprocal sentence cannot be captured by his formula.

Langendoen’s way out of this dilemma, a way also adopted by Moltmann (1992), is to state two versions of WR, one being WR as defined above and the other being a “type-shifted” generalization of WR, i.e., a formula that does not quantify over individuals, but over sets. Bearing in mind Quine’s innovation and our definition of cumulative predication, and recalling that all variables range over elements in $*D$, it can easily be verified that (24b) is still true in the problematic situation just described. Hence, a more general semantics for cumulation allows us to subsume both cases under the same analysis.

Taking this analysis of ERSs as a point of departure, I will move on to more complicated issues in the following sections.
A new degree of complexity comes into play when considering three place relations. The following construction constitutes a challenge for both the traditional theory of \textit{each}-movement and the view that the meaning of reciprocals could be captured compositionally by an operator that takes a two place relation as input and yields a one place relation as output:

(25) John read the letters they wrote to each other.

As pointed out by Irene Heim (p.c.) this sentence cannot be analyzed correctly on either assumption in a situation or model in which we cumulate writers, letters, and addressees simultaneously, and then state that the set of writers is identical to the set of the addressees. In particular, an analysis along the lines of \textit{each}-movement would imply an LF containing the proposition:

(26) Each of them writes these letters to the others

Interpreting \textit{each} as universal quantification yields too strong truth conditions and therefore does not allow us to analyze the cumulative reading correctly (for a more detailed exposition of the problem, see v. Stechow (1993, 82f) or Sternefeld (1993)).

Within the present framework, fortunately, no problems arise. The relevant part of (26) consists of an analysis of \textit{they write Y to each other}, where \textit{Y} is a variable to be bound by lambda abstraction, which is the standard semantics of relative clause formation. Simultaneous cumulation now requires a three place operator \(***\). For arbitrary \(n\)-place relations \(R\), the result of applying cumulative pluralization to \(R\) is the smallest relation \(R'\) which includes \(R\) and which is closed under accumulation of members of \(R'\). The general case is formalized in (27), where the arity of \(R\) is encoded by the number of stars preceding it.

\[
(27) \quad \star_{1} \ldots \star_{n} R := \bigcap \{Q : R \subseteq Q \land (\forall x_1, \ldots, x_n)(\forall y_1, \ldots, y_n) \[ ((x_1, \ldots, x_n) \in Q \land (y_1, \ldots, y_n) \in Q) \Rightarrow (x_1 \cup y_1, \ldots, x_n \cup y_n) \in Q] \}
\]

Translating \textit{they} as \(X\), the logical representation of (25) is straightforward:

(28) \(\langle X, Y, X \rangle \in \star_{1} \ldots \star_{n} \lambda x y z [x \neq z \land \text{write-to}(x, y, z)]\)

(28) implies that what they write to each other can be collected into a set \(Y\) such that for each element \(y\) of \(Y\), there are \(x, z \in X\) such that either \(x\) writes \(y\) to \(z\) or \(z\) writes \(y\) to \(x\). Moreover, the set \(X\) is such that for each
\[ x \in X, \text{ there are } y \in Y \text{ and } z \in X \text{ such that either } x \text{ receives a letter from } z \text{ or } x \text{ sends a letter to } z. \] These are the correct truth conditions and have been arrived at by simply applying the methods we have been using throughout this paper. 9

Note that natural language occasionally allows the expression of cumulative while using universal quantification at the surface:

\( (29) \quad \text{the letters every soldier received from home} \)

We here face the same problem that led us to conclude that (26) is inadequate as a LF for (25): It is not implied by (29) that each letter has been written by every soldier. As suggested by (29), universal quantification by *every* can occasionally be interpreted as forming a plural referential expression semantically equivalent to the description in (30):

\( (30) \quad \text{letters (all of) the soldiers received from home} \)

Interpreting *every* as a description that yields a plural term equivalent to *the soldiers* in (30) then yields the desired weak truth conditions. However, treating quantified NPs semantically as referential NPs creates a number of questions I cannot discuss here. Nonetheless I would like to discuss one problem which seems particularly threatening for such a proposal. Consider (31):

\( (31) \quad \text{the letters every soldier; received from his; mother} \)

The fact that *his* must be interpreted as a bound variable that ranges over individuals seems to contradict the above proposal not to interpret *every* in (31) as a quantifier. I will show how to resolve this conflict in the next subsection.

---

9 As has been observed by an anonymous reviewer, it is questionable whether all sentences of this form can be analyzed as involving cumulative predication. Thus, the reviewer asks whether (ia) can indeed “mean” (ib):

(i) a. John and Mary denied each other alcohol and cigarettes.
   b. John denied Mary alcohol and Mary denied John cigarettes.

Of course the more salient “reading” is the distributive one, which is the less interesting one, since it does not require three place cumulation. What might remain problematic in my theory is how to exclude such unwanted readings. Observe, however, that *we denied him cigarettes and alcohol* also might be considered weird in a context where I denied him cigarettes and you denied him alcohol. I conclude that the problem of salient readings is independent of the particular issue connected with reciprocity.
3.2. **Double Indexing and Dependent Plurals**

This section prepares the ground for the analysis of reciprocals in the following sections. I will discuss how certain ambiguities that arise with plural pronominals can be represented by indexing devices used at S-structure and at LF. These ambiguities have interesting repercussions for ambiguities that arise with reciprocals.

Consider first a sentence like (32):

(32) They love their parents

(32) can have a distributive reading, which claims that any one of them loves his own parents only. The logical paraphrase would thus require that the translation of their ranges over individuals, despite the fact that their is morphologically a plural NP.

Phenomena of this kind have sometimes been called “dependent plurals.” Within our system, the dependent reading of (32) can be analyzed as follows:

(33) $X_i \in *\lambda x_i [x_i \text{ loves } x_i \text{'s parents }]_s$

The point to be observed here is that the translation of their is still (by necessity — because of the morphological plural) a variable that ranges over elements in *D. But since love is D-based, the actual values of the variable will be elements in D only. Accordingly, whether or not a bound variable (plural) pronoun ranges over sets or over individuals seems to be a matter of scope, not of indexing.

On the other hand, despite having the same index in (32), they and their are not interpreted as coreferential. A possible reading with both expressions denoting the same set would be the distributive construal each of them loves each of their parents. As a first approximation, this could be represented as (34):

(34) $X_i \in *\lambda x_i [x_i \text{ loves } X_i \text{'s parents }]_s$

(34) suggests that we need two different variables although we have only one index at our disposal. But given that the notational difference between $X_i$ and $x_i$ serves mnemonic purposes only and has formal status in our theory, we in fact need two indices to spell out the ambiguity in a concise way.

Let us therefore adopt Heim’s (1993) distinction between outer and inner indices of NPs. Every NP has an outer index, its binding index. Referential NPs — in particular plural NPs and pronominal NPs — also
have an inner index, their referential index. If \( n \) is an outer index, the configuration \( \text{NP}_n \text{V} \) is to be interpreted as \( \text{NP} \in \lambda x_n \text{V} \). And if \( \text{NP} \) is a pronoun with the inner index \( i \), \( \text{NP} \in \lambda x_n \text{V} \) is in turn interpreted as \( x_i \in \lambda x_n \text{V} \).

Accordingly, \textit{they} in (32) must bear two indices. The outer one is picked up by \textit{their} in (33) as its referential intex; the inner one is picked up by \textit{their} in (34). In general, the referential index of an anaphoric pronoun can be coindexed with either the binding index or the referential index of the antecedent. Since the binding index appears inside and outside of the scope of pluralization, this will help to explain a number of ambiguities in the following sections. It should be mentioned here, however, that the choice between different indices alone is not always sufficient to disambiguate between readings. Consider, for example, the sentence \textit{They watch themselves}. In the dependent reading of the reflexive pronoun the sentence says that each one of them watched himself. The independent reading suggests a cumulative construal. At ALF these readings can be represented as in (35), where for reasons of readability it is assumed that the object precedes the verb:

\[
(35) \quad \begin{array}{c}
\text{a.} \\
\text{b.}
\end{array}
\]

Comparing the two representations it would seem that the essential dif-
ference lies in the different indexing of the anaphor. This, however, is incorrect. The reader may verify that replacement of all occurrences of $X_n$ by $x_1$ does not yield different truth conditions. Rather, the essential difference is the scoping of pluralization. In consequence, the indexing device alone is insufficient as a formal means to differentiate between dependent and independent readings; what matters is the semantic interpretation of the indexing, and, in particular, the question of having scope inside or outside of pluralization.

Returning to the end of the last subsection, it now emerges how to interpret his in (31). Recall that every cannot be translated as universal quantification here, and that the NP every soldier should be interpreted as a plural term. The problem is that his cannot be interpreted as coreferential with its antecedent, being interpreted as a plural term. However, the subject NP still has an outer index which can bind the possessive pronoun. The effect of quantifying over individuals thus results from the pronoun being interpreted inside the scope of pluralization. The latter serves as the hidden quantifier that enables his to be translated as a variable that ranges over individuals, as one would expect.

3.3. Reciprocals and Dependent Plurals

I will show in this section how the distinction between the referential and binding index applies to reciprocals. Consider, for example, the following sentence from Heim et al. (1991), which is ambiguous among the paraphrases listed in (37):

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

(37) a. John told Mary that he should leave, and Mary told John that she should leave.
    b. John told Mary that she should leave, and Mary told John that he should leave.
    c. John told Mary and Mary told John, “We should leave.”

These readings easily translate into the following formulae:

(38) $(\exists X)(X = \{j, m\} \land \langle X, X \rangle \in \lambda xy [x \neq y \land \ldots]$

\footnote{Observe, however, that a replacement of $x_1$ by $X_n$ in (35a) does make a difference. It would result in the implausible collective translation $X_n \in \lambda x_1. \text{watch}(x_1, X_n)$. Given our implicit assumption that the relation watch is D-based, such an ALF is necessarily false.}
These analyses reveal that we are in need of three different indices for the pronoun they. In fact, since the three possibilities for translating they correspond to the three different variables involved in the logical representation of the reciprocal, one might conclude that the analysis of the reciprocal itself is responsible for the threefold ambiguity.

Such a conclusion, however, would be premature. In the light of what we have said about dependent plurals it should be obvious that the ambiguity of they is totally independent of the analysis of the reciprocal. Above we have shown that in the context of pluralities, they can have a dependent reading, so that in a context like (36), the pronoun is in fact four ways ambiguous. It can have two dependent readings, one with respect to the matrix subject, and one with respect to the matrix object. In addition, we can also generate two corresponding independent readings. But due to the anaphoric nature of the reciprocal, these readings coincide, so that in sum we only get three readings instead of four.

How can the ambiguity be represented at LF? Having adopted Heim’s (1993) distinction between outer and inner indices of NPs, it is obvious that in (38i) the index corresponding to they (=x) is the outer index of the matrix subject; in (38ii) it (=y) is the outer index of the matrix object; and in (38iii) it (=X) is the inner index of either the matrix object or the matrix subject. What remains to be discussed is the representation of the reciprocal.

Let me illustrate this with a simpler example, by analyzing They watch each other, which differs from the cumulative analysis of They watch themselves in (35b) only in that we have to add a non-identity statement. Assume that non-identity is adjoined to a referential plural NP, so that the lexical items each other and one another are inserted into the ALF-tree as the indexed NP in (39):

(39) NP
    /    NP
        
        / x₂ ≠ x₁
          
X₁      2

We then have to move the inner NP into a position where it can bind x₂, as shown in (40):
As in the case of each-movement, the resulting traces of LF movement are not given any semantic interpretation. In both treatments movement obeys locality constraints. I will return to these in the following subsections.

3.4. LF Movement and Further Ambiguities of Scope

A further ambiguity discussed by Higginbotham (1980), Lebeaux (1983), and Heim et al. (1991) is shown in (41):

(41) John and Mary think they like each other.
   a. John thinks that he likes Mary, and Mary thinks that she likes John.
   b. John and Mary think they (that is, John and Mary) like each other.

In our theory, the reading paraphrased in (41b) is unproblematic; it is represented straightforwardly in (42):

(42) \( (\exists X)(X = \{j, m\} \land X \in \lambda x(\text{think}(x, [\{X, X\} \in \lambda y(x \neq y \land \text{like}(x, y)]))) \)

Here, the referential (inner) indices of all three NPs are identical. To get the alternative reading (41a) we cannot simply switch to the corresponding outer indices; this is because the corresponding dependent variable \( x \) ranges over individuals, and as soon as we replace \( \langle X, X \rangle \) in (42) by any of \( \langle X, x \rangle, \langle x, X \rangle, \) or \( \langle x, x \rangle \), the resulting formula expresses a contradiction. To get things to come out correctly we must in addition move each other into the matrix clause. The resulting logical analysis is shown in (43):

(43) \( (\exists X)(X = \{j, m\} \land X \in \lambda x(\lambda y(x \neq y \land \text{think}(x, [\text{like}(x, y)]))) \)
The fact that the reciprocal can have its antecedent in the matrix clause is predicted by the traditional analysis of binding in terms of Chomsky’s Specified Subject Condition. Recall from Chomsky (1973, 239) that subjects are “specified” if and only if they are lexical, but not if they are anaphoric (i.e., not coindexed with the potential antecedent). In (42), the subject they is coindexed with John and Mary, hence it is not “specified.” Therefore binding by the matrix subject, and thus movement into the matrix, is not blocked by the intervening subject.

This analysis differs from the traditional each-movement theory in that it is not only a distributive operator that moves into the matrix, but the entire reciprocal together with its referential variable. Some indirect evidence in favor of our analysis can be gained from the following observation. Compare (43), where they is translated as $x$, with the alternative construal shown in (44), where the pronominal translates as $X$:

(44) a. $(\exists X)(X = \{j, m\} \land \langle X, X \rangle \in *\lambda x y[x \neq y \land \text{think}(x, [X \in *\lambda x . \text{like}(x, y)])])$

b. John thinks that John and Mary like Mary, and Mary thinks that John and Mary like John.

Intuitively, such an “independent” analysis of the embedded they seems to be unavailable, although this kind of ambiguity is perfectly acceptable in previous examples. An explanation for the ungrammaticality of (44) emerges when taking into account that movement of the reciprocal might have caused a cross-over violation. This follows if we compare the referential index of the reciprocal with that of the pronominal. In both cases the translation is $X$ in the ungrammatical construal, so that the NPs must have identical referential indices. Assuming, as seems natural, that movement goes to an A-bar position that c-commands they, the pronominal becomes bound within the domain of movement. We therefore immediately encounter a cross-over violation.

Another difference from the traditional analysis consists in moving the non-identity statement out of the scope of the intensional verb think. According to the each-movement analyses of (41) proposed by Lebeaux (1983) or Heim et al. (1991), the non-identity statement remains in situ. I consider this a problematic feature of these analyses, since one normally does not entertain propositional attitudes towards (non-)identity, unless identity is explicitly addressed in the belief. It seems, then, that the reciprocal can move into the matrix clause only as a whole. In order to ensure this we need a further locality conditions at the level of ALF.
As indicated above, movement of the entire NP in (43) seems to be restricted by the Specified Subject Condition. Movement of the smaller NP, which has been so far an obligatory feature of the analysis (but see the next subsection), is also very local. Presumably, this kind of movement can only be minimal, being restricted to movement across pluralization operators, but excluding movement across other operators or lexical material.

3.5. **Strong Reciprocity**

As regards the examples so far, our formalization in terms of cumulations seems sufficiently strong to yield acceptable truth conditions. There is, however, exactly one stronger variant of reciprocity that might have intuitive plausibility, and this happens to be the strong reciprocal analysis which is presupposed in the theory of each-movement. However, as it happens, SR, as defined in (3a), cannot be expressed by any combination of (43) with pluralization.¹¹

This brings up the question of whether there is in fact such a strong “reading,” or whether the strong construal is just a special case of the weak reading. On the other hand, the reader might already have wondered whether or not there is a way to make the components of the reciprocal fit into one NP-meaning that could be interpreted in situ. I will now show that any reasonable attempt to achieve this will automatically generate SR.

Let us, therefore, try to interpret the reciprocal as one single NP. Of course, a certain rearrangement of the logical material in (43) is called for. The obvious change is to provide the entire NP with an outer index. The next step is to remove the type mismatch between $\mathcal{CG}$ and the non-identity statement by inserting semantic glue. The only way of doing this permitted by the theory of ALF is to insert set formation (or lambda abstraction) and intersection (or conjunction). The only reasonable outcome is displayed in (45):

```
There is yet another kind of construction in which WR seems too strong, as in The plates are stacked on top of one another. This is the rule with predicates that express so-called “asymmetric disconnected relations.” I will ignore relations of this type and refer the reader to Langendoen (1978) for some discussion.
```
In (45), \( k \) is the referential index of the reciprocal (to be bound by the antecedent), \( i \) is the outer index of the antecedent, and \( j \) is the outer index of the reciprocal. Inserting this into the syntactic context of an ERS, the result is (46a), which is equivalent to (46b):

\[
(46) \quad \text{a. IP} \\
\quad \text{NP} \quad \text{NP} \quad \text{NP} \\
\quad \text{NP} \quad \text{NP} \quad \text{NP} \\
\quad \text{A} \quad \text{A} \quad \text{A} \\
\quad \{y : x_1 \neq y\} \\
\quad \{y : x_2 \neq y\} \\
\quad \{y : x_1 \neq y\} \\
\quad \{y : x_2 \neq y\} \\
\]

Assuming now that \( A \) is D-based, (46b) is logically equivalent to SR.

Although this analysis shows that SR can be expressed at ALF in our theory, the strong reading requires such a different arrangement of the meaning components of the reciprocal expression that a more straightforward approach might be desirable. On the other hand, the decision is still pending on whether the strong construal should indeed be regarded as a "reading," or whether it should be considered only a special case of the weak construal. In the latter case, the more complicated analysis could be dispensed with entirely. I have no straightforward empirical evidence in either direction, but I will keep an eye on this question in what follows.

3.6. The Geach-Kaplan Sentence

In this section I will offer some additional evidence in favor of the view that the reciprocal expression splits up at ALF into two components. The evidence results from the analysis of (47), and I will show that it is exactly the two components of the reciprocal that can be focussed on by only.
Some critics admire only one another.

(47) has become known as the “Geach-Kaplan sentence,” and is first cited in the literature in Quine (1980, 293). Quine’s interest is the difference between set theory and second order logic, and in this context the sentence has figured prominently in a number of papers; cf., e.g., Lewis (1991, 63) and Cartwright (1993, 203) as the most recent references. My concern here is whether the interaction of the reciprocal with only necessitates an analysis of the sort proposed in the last subsection.

First note that in all the references cited above I found agreement as regards the logical analysis of (47), which is uniformly represented as (48a), which in turn is equivalent to (48b) in our notation:

(48) a. $$(\exists x)(\forall y \in x)(\forall y \in x) (C y \land (\forall y)(\forall z)(y \in x \land Ay z \rightarrow y \neq z \land z \in x))$$

b. $$(\exists X)((\text{critics}(X) \land (\forall y \in X)(\forall z)(\text{admire}(y, z) \rightarrow y \neq z \land z \in X)))$$

How, then, can the meaning of only, when applied to that of the reciprocal, yield (48b), so that the logicians’ analysis is predicted by our theory?

Consider first the strong reciprocal reading of (49a). As shown in the last subsection, this can be formalized by iterated cumulation as in (49b). Given that admire is D-based, this is equivalent to (49c):

(49) a. Some critics admire each other.

b. $$(\exists X)((\text{critics}(X) \land (\forall y \in X)(\forall z)(\text{admire}(y, z) \rightarrow y \neq z \land z \in X))$$

c. $$(\exists X)((\text{critics}(X) \land (\forall y \in X)(\forall z)(\text{admire}(y, z) \rightarrow y \neq z \land z \in X))$$

An immediate way to derive (48b) from (49c) would be to reverse the implication $\rightarrow$. Intuitively, “only($A \rightarrow B$)” implies “$B \rightarrow A$,” so this implication reversal would directly yield the desired result.

Unfortunately, however, there is to date no compositional formal semantics for only that would allow us to account for the reversal of the implication. Let us, therefore, adopt the focus semantics for only developed by Rooth (1985). By way of focussing on the property $\{y : y \in X \land x \neq y\}$, which figured as the translation of each other in the last subsection, we now get the following representation:

(50) $$(\exists X)((\text{critics}(X) \land X \in *\lambda x (\text{only}([\{y : y \in X \land x \neq y\}]_x \in *\lambda y . \text{admire}(x, y))))$$
Following Rooth, we first have to look for alternatives to the focussed set, i.e., the set of all elements in \( X \) except \( x \), and see whether or not they satisfy the above formula. If any of them does, the sentence is false; otherwise it is true. But we also have to be careful to restrict these alternatives appropriately: a little reflection will show that within plural semantics it is generally the case that subsets of the focussed set cannot qualify as alternatives. Given this “no subset condition,” (50) says that non-subsets \( Y \) of the focus that satisfy \( Y \in \lambda y. \text{admire}(x, y) \) cannot exist. In particular, the non-subset \( \{x\} \) cannot satisfy the property of being admired by \( x \). But some further reflection shows that this is exactly what the Geach-Kaplan truth conditions express. In consequence, it seems that Rooth’s focus semantics requires an analysis of reciprocals along the lines proposed in the last section. In other words, the strong construal of the reciprocal expression as one NP seems to be a prerequisite for the analysis of only(NP).

There are, however, a number of empirical problems with this analysis. First of all there are speakers for whom the entailment \( x \neq y \) appears too strong. Hence it would seem that this part of the reciprocal meaning is not necessarily focussed. Secondly, we encounter a problem with the presupposition of the Geach-Kaplan sentence, namely with (49a). This is not yet implied by our analysis of only. But now, adding the presupposition of only in the usual manner (by stipulating that \( Y\text{-only}-Z \) presupposes \( Y-Z \)), we necessarily presuppose strong reciprocity; i.e., we derive (49b) as the presupposition of (47). But now assume that \( \text{admire} \) is interpreted as in (51):

\[
\begin{array}{ccc}
(51) & a & a \\
b & & b \\
c & & c
\end{array}
\]

In this model only WR holds, but intuitively the Geach-Kaplan sentence still seems to be true. This, however, requires a new analysis, in fact one that starts off with weak reciprocity and then adds only with its associated focus.

Suppose, therefore, that we construe only in the way shown in (52), where the focus of only still has to be determined:

\[
(52) \quad (\exists X)(\text{\#critics}(X) \land \text{only}(\langle X, X \rangle \in \text{admire}(x, y) \land x \neq y)))
\]
Assume we focus on the second occurrence of $X$. Then the formula says that critics admire only critics, but it does not exclude self-admirers. This corresponds to the weak reading, which is accepted by speakers who do not exclude self-admiration. To get the strong reading proposed by Quine, Kaplan, Geach, and others, let us focus on $x \neq y$ (and optionally also on the second occurrence of $X$ again). What is a plausible candidate for an alternative to non-identity? It seems that the only reasonable alternative is identity, so that the alternatives to be considered have the form given in (53a). For the only-sentence to be true, no such alternative can be true, or, following Rooth’s semantics literally, (53a) must imply (53b):

\[
\begin{align*}
\text{(53) a. } & \langle X, Y \rangle \in \text{**}{\lambda}x,y(\text{admire}(x, y) \land x = y) \\
\text{b. } & \langle X, X \rangle \in \text{**}{\lambda}x,y(\text{admire}(x, y) \land x \neq y)
\end{align*}
\]

Assume that for some $a$, \text{admire}(a, a) is true. It is clear that the pair $\langle a, a \rangle$ satisfies (53a). However, (53b) is false for $X = a$, because, as a consequence of $x \neq y$, any value of $X$ must contain at least two members. It thus follows that in such a situation the Geach-Kaplan sentence is false. This is exactly what is implied by (48), so that focussing on $x \neq y$ yields the desired result.

Summarizing the discussion, I have shown that different readings of the Geach-Kaplan sentence can be obtained by focussing on different parts of the translation; moreover, the strong reading that excludes self-admiration is consistent with and can even be derived from WD. Therefore, the analysis does not support the claim that we need SR as a separate reading. SR as a presupposition has the same status as SR in general, being probably a special case of WR. I will continue to discuss the issue in the next section.

4. Semantics vs. Pragmatics

4.1. Elementary Plural Sentences

Hitherto I have also left open how to deal with SD, i.e., the distributive reading that renders \textit{They wrote musicals} false in Gillon’s model. I envisaged two solutions to the problem: One is to deny that such a reading exists; the other is to introduce new logical glue at the level of ALF. The latter can be done in two ways, either by modifying verb denotations by means of an invisible distributional operator $\Delta$ as defined in (15), or by modifying noun denotations by means of an invisible distributor, as in the traditional theory of \textit{each}-movement.
There is, however, a more interesting alternative pursued by Schwarzschild (1991; 1996). He proposes a unifying analysis of purported semantic ambiguities in which the availability of readings for EPRSs is attributed to properties of the context. In order to formally implement this idea he first introduces the notion of a cover of a plural NP denotation \( X \). According to Schwarzschild, a set of sets \( Y \) covers \( X \) iff \( Y \subseteq X \) and \( \emptyset \notin X \). If \( Y \) is a cover over \( X \), this will be written as \( C(Y, X) \), and I sometimes use \( C_X \) as a variable for a cover over \( X \). Given an arbitrary one-place predicate \( P \), the reader may verify that the following equivalence is logically true:

\[
X \in \star P \iff (\exists Y)(C(Y, X) \land (\forall z \in Y)(z \in P))
\]

Accordingly, the cumulative reading formerly expressed by (55a) can now be restated as (55b):

\[
\begin{align*}
(55) \ a. & \quad f(they) \in \star f(write-musicals) \\
\ b. & \quad \text{There is a cover } C_{f(they)} \text{ such that each element of } C_{f(they)} \text{ is} \\
\ & \quad \text{in } f(write-musicals).
\end{align*}
\]

In fact, all cumulative predications can straightforwardly be reformulated in terms of covers.\(^\text{12}\)

The next step is to impose certain restrictions on the choice of a cover. For example, distributive predication, which is paraphrased as \( they \text{ each wrote musicals} \), can now be expressed by the requirement that \( C_{f(they)} = f(they) \). Likewise, the group reading, which can be paraphrased as \( they \text{ all wrote musicals together} \), follows from the assumption that \( C_{f(they)} = \{ f(they) \} \).

Schwarzschild argues that the availability of certain readings depends on context; in some situations the cumulative construal may be too weak and the distributive construal too strong; cf. the discussion in

\[\text{12} \quad \text{For example, the truth conditions of the modified Scha-sentence, repeated here as (ia), can be restated in terms of covers as in (ib):}\]

\[
\begin{align*}
\text{(i) a.} & \quad (\exists X \in \star DF)(\exists Y \in \star JC)((500)(X) \land 2000(Y) \land \langle X, Y \rangle \in \\
\ & \quad \star \text{[own]}) \\
\ b. & \quad (\exists X \in \star DF)(\exists Y \in \star JC)((\exists C_X)(500)(X) \land 2000(Y) \land \\
\ & \quad (\forall x \in C_X)(\exists y \in Y)(\langle x, y \rangle \in \text{[own]} \land \\
\ & \quad (\forall y \in Y)(\exists x \in C_X)(\langle x, y \rangle \in \text{[own]})) \\
\end{align*}
\]

The lack of compositionality exhibited by (ib) will lead to the definition of pair-covers further below.
Schwarzschild (1991, 107ff; 1996, 67ff.). However, if covers can be determined contextually, truth conditions come out correctly. In particular, covers are extremely useful in the context of union theory. Recall that according to that theory, conjoined NPs like the boys and the girls can have the same denotation as the younger and the older children, provided that the union of the younger children with the older children is just the set of boys and girls. Nevertheless, there is a distributive reading in which the boys and the girls met in the park may be true, but the younger and the older children met in the park is false. These readings can be distinguished by assuming that, although both sentences have the same truth conditions (56a), with $\mathcal{C}$ being defined as in (56b), the two NPs induce different covers of $X$, one that divides the children into the old and the young ones and one that divides them into boys and girls. Clearly, these different contexts make one sentence true and the other false.

$$\text{(56) a. } X \in \mathcal{C}(f(\text{meet})), \quad \text{b. } X \in \mathcal{C}P \text{ iff the contextually determined cover } \mathcal{C}_X \text{ is such that each element of } \mathcal{C}_X \text{ is in } P.\]$$

Turning again to the analysis of EPRSs, Schwarzschild discusses Scha’s sentence (57a), which describes the situation shown in (57b). Since parallel is D-based, our analysis (57c) is logically equivalent to (57d):

$$\text{(57) a. The sides of rectangle 1 are parallel to the sides of rectangle 2.} \quad \text{b.}$$

$$\quad \text{c. } \langle [\text{the sides of R1}]_m, [\text{the sides of R2}]_m \rangle \in \mathcal{C}[\text{parallel}]_m, \quad \text{d. } (\forall x \in [\text{side of R1}]_m)(\exists y \in [\text{side of R2}]_m)(\langle x, y \rangle \in [\text{parallel}]_m) \wedge (\forall x \in [\text{side of R2}]_m)(\exists y \in [\text{side of R1}]_m)(\langle x, y \rangle \in [\text{parallel}]_m).$$

Now, Schwarzschild argues that, intuitively, the semantics of (57a) involves a pairing of two covers. Each cover separates the horizontal lines from the vertical lines in each rectangle, so that the paired cover has to pair together the horizontal lines of R1 with those of R2, and the vertical lines of R1 with those of R2. More formally, we quote his definition 190)
and restate his truth conditions as in (58):

\[ PCov \text{ is a paired cover of } \langle A, B \rangle \text{ if } PCov \text{ is a set of pairs constructed as follows: take a cover of } A \text{ and a cover of } B \text{ and form a set of pairs such that every element of the } A \text{ cover is a first member and every element of the } B \text{ cover is a second member of a pair in } PCov. \]

\[ \langle A, B \rangle \in \otimes \otimes R \iff \text{the pragmatically determined } p\text{-cover } P\text{ of } \langle A, B \rangle \text{ is such that for all } \langle x, y \rangle \in P_{AB}, \langle x, y \rangle \in R. \]

Applying (58) to the case at hand, let \( A = \) “the sides of R1,” \( B = \) “the sides of R2,” and \( R = \) “are parallel.” Accordingly, the analysis states that, given an appropriate cover, the horizontal lines of R1 are parallel to the horizontal lines of R2, and the vertical lines of R1 are parallel to the vertical lines of R2. However, as the reader may have noticed, the relation \( R \) still holds between sets rather than individuals. Hence \( R \) must already be “pluralized” — an operation which Schwarzschild calls Paired-Part-PCov and which is defined in (59):

\[ \text{PPart}(PCov) \in \otimes \otimes \text{parallel} \]

Observe that PPart(PCov) works like a two place cumulation, except that the pairs that cumulate must be determined by a PCov. Closer inspection will also reveal that the condition stated in (58) parallels that in (59), so that Schwarzschild’s ultimate analysis is \( \langle R_1, R_2 \rangle \in \otimes \otimes (\Delta \text{parallel}) \)

There is no doubt that the truth conditions come out correctly, provided that the contextual determination works correctly. But this is exactly the point where intuitions become vague. In the above example it seems to me that the basic intuition that governs us in choosing a particular paired cover is our desire to make \( \langle a, b \rangle \subseteq \langle c, d \rangle \) if \( a \subseteq c \) and \( b \subseteq d \).

Observe that PPart(PCov) works like a two place cumulation, except that the pairs that cumulate must be determined by a PCov. Closer inspection will also reveal that the condition stated in (58) parallels that in (59), so that Schwarzschild’s ultimate analysis is \( \langle R_1, R_2 \rangle \in \otimes \otimes (\Delta \text{parallel}) \)

There is no doubt that the truth conditions come out correctly, provided that the contextual determination works correctly. But this is exactly the point where intuitions become vague. In the above example it seems to me that the basic intuition that governs us in choosing a particular paired cover is our desire to make \( \langle R_1, R_2 \rangle \in \otimes \otimes (\Delta \text{parallel}) \) true. In other words, we set up the partitions the particular way we do just because we want to find a way of making \( \Delta \text{parallel} \) true. Put differently, the reason for choosing a pragmatically determined partition is most probably not another pragmatically determined partition that accidentally turns out to be equivalent to \( \Delta \text{parallel} \). Rather, the pragmatic choice is determined by semantic truth conditions that exist independently of the pragmatics of covers.

The same point can be made with Schwarzschild’s discussion (1991, 131f; [1996, 87f]) of the following sentence from Scha and Stallard
The frigates are faster than the carriers.

Imagine, for example, that (193) is uttered in a context in which it is clear that these ships are sent out in teams to different areas of the globe with each team consisting of frigates and carriers. It may be that one area calls for very fast action while another requires a rather sluggerish response. If that were the case, I would judge (193) true just in case the frigates in a given area were faster than the carriers of that area, regardless of what speed relations obtained between ships of different areas. In this situation the universal-universal reading is too strong and the other reading is too weak. A semantics that incorporates the notion of a contextually determined partition accounts for these facts without having to drum up new translations.

Here again we would need two covers, one that partitions the ships into operation areas, and a second one which guarantees that for each operation area it must be true that all the frigates (or perhaps the typical frigates) must be faster than all carriers (or perhaps the typical carrier). Accordingly, this second cover must be equivalent to SD, but again this does not seem to be a matter of context dependence. In fact, it seems reasonable to say that a cover is “neutral” if the context does not impose any special restrictions on the cover. This means that any such neutral cover $\circ \circ$ is equivalent to WD. However, for WD to be satisfied it would suffice that in each operation area each frigate must be faster than at least one carrier, and each carrier must be slower than at least one frigate. These conditions are inadequate; in all of Schwarzschild’s examples only SD is correct.

This outcome cannot be accidental. But how can an analysis in terms of $\circ \circ \circ \circ$ be ruled out? And what does this tell us about the status of SR as a genuine reading? Before trying an answer to the second question let me address the first question and discuss some examples in which $\circ \circ \circ \circ$ still makes sense.

4.2. Elementary Reciprocal Sentences

Recall that the main objective of Schwarzschild’s dissertation is to support union theory, i.e., the thesis that a conjoined NP like the men and the women denotes the union of the set of men and the set of women. This set is unstructured. If it should turn out that we need to have access to the conjuncts separately, we can refer to them only as parts of a cover that is induced by the linguistic structure of the NP. Although this claim is central to Schwarzschild’s (1991) dissertation, it has not been checked.
against examples like the following, which seem to challenge the union theory: 13

(60)  
  a. The men and the women had relations with each other.
  b. The men and the women hate each other.

The considerations of sections 3 and 4.1 suggest an analysis along the lines of (61), with \{M, W\} as the relevant partition of X into the sets of men and women, which is contextually induced by the subject expression.

(61) \( \langle X, X \rangle \in \oplus \{ \forall x y (x \neq y \land \text{hate}(x, y)) \} \)

There are different ways to build a p-cover of \{M, W\}. Suppose we group the men with the men and the women with the women. This will yield the distributive reading which is spelled out in (62): The men hate each other, and the women hate each other.

(62) \( \forall X \in \{M, W\} (\langle X, X \rangle \in \forall x y (x \neq y \land \text{hate}(x, y)) \)"

Alternatively, one might pair the men with the women and vice versa, which yields (63):

(63) \( \forall \langle X, Y \rangle \in \{\langle M, W\rangle, \langle W, M\rangle\} (\langle X, Y \rangle \in \forall x y (x \neq y \land \text{hate}(x, y))) \)

As the non-identity statement in (63) only says that we may not take hermaphrodites into account, this part of the formula becomes more or less trivial, and (63) simply means that the men hate the women and the women hate the men.

Clearly, these truth conditions seem empirically correct. Nevertheless, trouble arises with further possibilities for building p-covers, e.g. \{\langle M, M\rangle, \langle M, W\rangle, \langle W, W\rangle\}, which do not correspond to intuitively possible analyses. The problem intensifies when we turn to more than two conjuncts, as in (64):

(64)  
  The men, the women, and the children hate each other.

An investigation of all covers of \(X \times X\) will reveal that we have to rule out “mixed readings,” e.g., an analysis in which the children hate the children, the men hate the women, and the women hate the men.

---

13 Two anonymous reviewers brought Schwarzschild (1992) to my attention, where examples of this type are indeed under consideration. As the reader may verify, however, his discussion in no way touches on the point I want to make in this section.
In fact, only “reciprocal” p-covers can be allowed. It follows that the above analysis of the reciprocal has misrepresented the semantic role of reciprocity: The intuitively disallowed covers should be ruled out by the semantics of the reciprocal rather than by the “pragmatics” of p-covers.

Let us now determine what the correct truth conditions are. The above discussion suggests that the non-identity statement has to apply to sets rather than to individuals. Thus, the correct analysis must involve a different scoping, as given in (65):

$$
(\exists X)(X = \{M \cup W \cup C\} \land \langle X, X \rangle \in \otimes \lambda X_i X_j [X_i \neq X_j \land \langle X_i, X_j \rangle \in \lambda x_i x_j**\text{hate}(x_i, x_j)])
$$

By first applying $\otimes$ we identify pragmatically salient paired covers, which are induced by the structure of the subject. The reciprocal is the non-identity statement that excludes reflexive pairs. The relation of hatred now holds cumulatively between these pairs. Thus, the analysis hinges on the pragmatic impact of the subject expression, captured by $\otimes$, but then goes on with the usual semantic strategies of pluralization, represented here by $**$. (65) also suggests a rather strict interpretation of the locality constraint for the movement of the referential movement. Above we assumed that this kind of movement can only skip pluralization. In order to rule out (61) we must strengthen this condition by saying that it cannot skip more that one “pluralization” operation. 14

Thus far I have represented the weak version of the reciprocal construals; we still have to account for the distributive reading, which says that the men hate each other, the women hate each other, and the children hate each other. This requires covering the subject by the obvious trivial partition $C = \{M, W, C\}$. We then have to interpret the reciprocal construction as in (66a). After eliminating $\otimes$, this is equivalent to (66b):

$$
(66) \begin{align*}
(66a) & \{M \cup W \cup C\} \in \otimes \lambda X. \langle X, X \rangle \in **\lambda y z [y \neq z \land \text{hate}(y, z)] \\
(66b) & (\forall X \in \{M, W, C\})(\langle X, X \rangle \in **\lambda y z [y \neq z \land \text{hate}(y, z)])
\end{align*}
$$

Note that Schwarzchilds analysis also accounts for restrictive relative clause constructions like (i), called “hydras” in Link (1984):  

(i) the men and the women who hate each other

The only difference from (65) is that the partition induced by the NP-head must not cover the meaning of the men and the women, but rather the variable bound by the relative pronoun who.
These analyses show that it is absolutely crucial to apply various sorts
of pluralization iteratively. A theory of ALF which allows insertion of
logical material in a fairly unrestricted way will automatically account for
these readings. Let us briefly look at the ALF of (66):

\[
(67) \quad \text{IP} \quad \text{NP} \quad \text{VP} \quad \text{NP} \quad \text{VP}
\]
\[
\{M \cup W \cup C\} \quad \text{1} \quad \text{NP} \quad \text{VP} \quad \text{NP} \quad \text{VP}
\]
\[
X_1 \quad 2 \quad \text{NP} \quad \text{V}(x_1, x_2)
\]
\[
t \quad x_1 \neq x_2 \quad \text{hate}
\]

Due to the double application of pluralization operations the anaphor can
take up an index of the antecedent which is interpreted inside the scope
of cumulation by \(\odot\). This is in fact crucial for the truth conditions to
come out correctly. Suppose the anaphor has the referential index of the
antecedent. We then get the wrong analysis in (68):

\[
(68) \quad \{M \cup W \cup C\} \in \odot \lambda X.\langle X, \{M \cup W \cup C\} \rangle \in \ast \ast \lambda y z[y \neq z \land \text{hate}(y, z)]
\]

Of course such an analysis has to be blocked. Recall that in all previous
cases the difference between taking an inner or an outer index was relevant
for pronouns but not for reciprocals, i.e., both possibilities yield the
same truth conditions. However, as (68) shows, there are cases where the
difference matters; in accordance with (68) we may therefore stipulate
that only the outer index of the antecedent should be available as the
referential index of the reciprocal.

4.3. Conclusion

In the above sections I tried to show that a theory of ALF is capable
of analysing the many ambiguities that arise with plural sentences in
a principled way, by inserting semantic glue. As this process applies
fairly unrestrictedly, I occasionally looked at overgenerations and asked
how they can be blocked. In some cases I proposed \textit{syntactic} principles,
while in other cases we found that the forms generated express \textit{semantic}
contradictions, and hence are likely to be automatically understood in an
alternative way. Schwarzschild added some further \textit{pragmatic} points of
view, although I tried to make clear that a pragmatic determination of covers is unlikely when cumulation applies at lower levels, i.e., at his level of $\text{PPart}(\text{PCov})(\alpha)$. It seems, therefore, that at deeper levels proper semantic operations are at work, operations that are not influenced by pragmatic constraints on covers.

This brings us back to the question of different readings. In Schwarzschild’s theory, different readings should not exist; we only have different kinds of context dependencies. In our theory we have multitudes of different readings, but as in all theories of this kind the stronger “readings” might just be a special case of the weaker ones. One theoretical argument in favor of readings is that once we start with some minimal assumptions about compositionality and the role of ALFs, there is no reason to block the representations that correspond to these “readings.”

Another argument is based on our discussion in section 4.1. Recall that Schwarzschild’s examples uniformly required an analysis of the form $\circ\odot (\text{SD})$ rather than $\circ\odot (\text{WD})$. This suggests that SD is indeed a genuine reading. On the other hand, WD must be ruled out, and I have already indicated why WD is inadequate: the weak construal is not informative enough. Saying that the frigates are faster than the carriers is misleading when the truth conditions require that only one frigate must be faster than all carriers. But given that WD is inadequate, how can we correctly interpret the sentence? The answer I would like to propose is that we try a different reading, namely SD. And so far there is no other sensible choice that could be made. In contrast, Schwarzschild’s notion of a cover would allow for indefinite other ways of making the sentence true and more informative. Intuitively, however, only SD seems to be correct, and this cannot be explained by a theory that is based on the context dependency of covers.

ACKNOWLEDGEMENTS

This paper’s history reaches back to my Habilitationsvortrag in June 1992. Its long development makes it difficult to list all the people who have been involved, but the ones who haven’t escaped my memory are Sigrid Beck, Urs Egli, Manfred Krifka, Manfred Kupffer, Claudia Nohl, Heike Winhart, and four reviewers. Kirsten Brock, Graham Katz, Irene Heim, Kimberly Miller, Renate Musan, Ulrich Sauerland and Arnim v. Stechow commented on previous drafts of the paper (one of them prepublished as Sternfeld (1993)) and drew my attention to numerous inconsis-
tencies, confusions, and awkwardnesses of presentation. I would also like to thank the participants of the 1992 Tübingen Plural Conference. Originally having been submitted as a contribution to the proceedings of the conference, a referee’s comments led to substantial revisions in January 1994. Lack of progress on the side of the editors caused my retraction from the volume and explains the long history of the paper. Submission to NLS led to further revisions, and I would like to thank Irene Heim again for her constant invaluable encouragement over the years and for pointing out to me serious flaws in my first attempts to define a reciprocal operator.

Needless to say, even revised manuscripts do not always ameliorate over the years, and since I have not always taken the advice I have been offered, responsibility must be placed on me alone for whatever has been changed and for what has not.

REFERENCES


Lebeaux, David: 1983, ‘A Distributional Difference Between Reciprocals and Reflex-
Rooth, Mats: 1985, Association with Focus, Ph.D. dissertation, University of Massachusetts at Amherst.

Seminar für Sprachwissenschaft
Universität Tübingen
Wilhelmstr. 113
D-72074 Tübingen
Germany
Sternefeld@uni-tuebingen.de