The Representation of Reciprocals in Grammar*

Uli Sauerland

Massachusetts Institute of Technology

1 Introduction

The idea that the English reciprocal *each other* has a complex structure is an old one in syntax (Lebeaux (1983), Baker *et al.* (1989)). Heim *et al.* (1991) proposed a more radical account of the reciprocal: All the semantic and syntactic properties of the reciprocal should be explained as the interaction of its complex strcture with independently needed principles of grammar. In this paper I will briefly summarize problems of the original account of Heim *et al.* (1991), and then present a novel analysis of the reciprocal that closely follows the orginal intuition of their analysis.

My proposal disagrees with the recent analyses of Moltmann (1992), Sternefeld (1993), and Dalrymple *et al.* (1994b), who employ rules that specifically refer to the reciprocal. Their argument for such specific reciprocalization-rules is for each of them based on one of the following examples, for which the analysis of Heim *et al.* (1991) does not provide the interpretations paraphrased in b):

(1) a. Byron and Chandos wrote these letters to each other. (Sternefeld (1993))

b. Each of these letters was either written by Byron to Chandos or by Chandos to Byron, and both Byron and Chandos wrote at least one of these letters.

(2) a. The children follow each other into the room. (Dalrymple *et al.* (1994a))

b. For a big number of the children: They followed another child into the room.

*The largest part of this paper is a spelling out of classnotes from Irene Heim's spring 1994 *Advanced Semantics* course at MIT. It is fair to say that she contributed more to this paper than I have. Furthermore I profited from the comments of Noam Chomsky, Diana Cresti, Piroska Csuri, Viviane Déprez, and the audiences at MIT, ESCOL '94 at the University of South Carolina, and WECOL 1994 at UCLA. All remaining errors are my own. A less abbreviated version of this paper is going to appear in *MIT Working Papers in Linguistics*, Volume 25. Financial support was provided from the German academic exchange service DAAD with a grant in the second Hochschulsonderprogramm HSP II/AUFE. To appear in: Proceedings of ESCOL '94 held at the University of South Carolina. Cornell University Working Papers in Linguistics.

For both of the above sentences the only interpretation that Heim *et al.*'s (1991) analysis provides is too strong. For Sternefeld's example (1) they only get an interpretation that demands that all the letters went both ways. This interpretation is paraphrased as: Byron and Chandos each wrote these letters to the other. For the example (2) the only interpretation on their account is the one where the children go around in a circle. The paraphrase of this interpretation is: The children each followed the others into the room. In this paper I will provide an analysis for both sentences without surrendering the intuition that no principle of grammar refers specifically to reciprocals.

The structure of the paper is the following: In the rest of the introduction I will briefly describe the semantic framework and Schwarzschild's (1991) semantics of plural, which I will then use. Section 2 motivates a general distributivity operator that is needed independently of the reciprocal. Section 3 first introduces the lexical entry of the reciprocal and then displays how it interacts with binding theory, quantifier raising, and the insertion of the distributivity operator. In section 4 I will lay out how pragmatics enters into the semantic description and derive the effects of the *strongest meaning hypothesis* of Dalrymple *et al.* (1994a).

The general model this investigation is based on the following assumptions: Semantic interpretation takes as input the logical form of a principles and parameters style syntax, which is a binary branching tree. On the semantic side the possible expressions are given by a functional type theory, where for my purposes the two basic types e for individuals and groups and t for truth values are sufficient. Each terminal node of the tree is mapped either onto an expression of this type theory, or onto a λ-abstractor. The interpretation of a non-terminal node α is determined by the values of the two daugthers β and γ: If one of the daughter-nodes has the appropiate type to function as an argument for the other one, then the mother-node is interpreted by functional application $\beta(\gamma)$ or $\gamma(\beta)$. If both daughter-nodes have an identical type $\langle\delta, t\rangle$ that is a function into truth values, the mother-node is interpreted by intersecting the two as $\lambda x^{\delta}(\beta(x) \wedge \gamma(x))$. If one of the daughter-nodes is an abstractor λn, the mother-node derives from the other daughter-node as abstraction over this variable $\lambda n\beta$ or $\lambda n\gamma$. In all other cases the logical form is semantically ill-formed

The plural ontology I assume is, with some notational differences, the *union theory* that Schwarzschild (1991) argues for extensively. What he concludes is that all plural DPs denote sets of individuals, since all the reasons that lead e.g. Link (1991) and Landman (1989) to postulate structured groups seem to be merely pragmatic effects, whereas binding facts undermine the structured groups approach. I will make one notational simplification of Schwarzschild's system, namely that I write groups as mereological sums, not as sets. In a mereological setting the basic assumption of the union theory can be expressed as the postulate that the mereological sum operation $\oplus$ is associative. Now, calling the type e that of individuals is somewhat misleading because groups are contained within the same type-domain, but I will continue with

this use. In addition I assume that the mereological sum operator also applies to n-tuples of individuals, where $(a_1, \ldots, a_n) \oplus (b_1, \ldots b_n)$ is defined as $(a_1 \oplus b_1, \ldots, a_n \oplus b_n)$.

2 (Co-)distributivity with Sternefeld's $\star$-operator

Since Scha (1984) it is known that multiple plural noun phrases in a sentence like (3-a) can give rise to codistributive[1] readings, namely the reading paraphrased in (3-b).

(3) a. The women face the men. (cf. Schwarzschild (1992))
b. For each of the women there is a man who she faces, and for every man there is a woman who faces him.

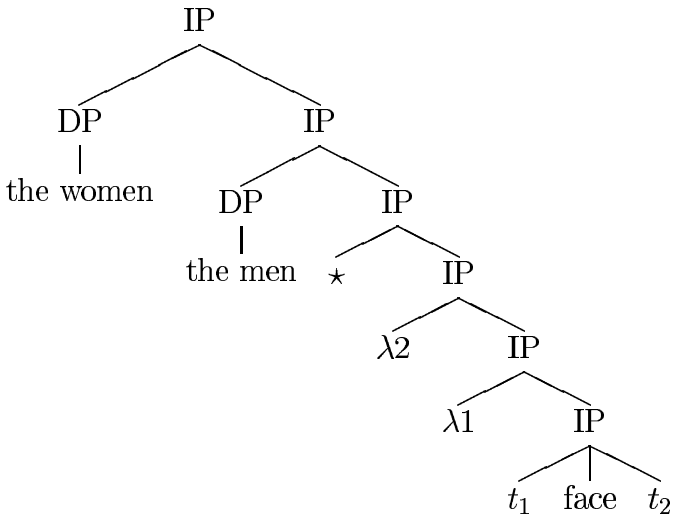
Following Sternefeld (1993) I will subsume these examples under a general distributivity operator that applies to predicates. This operator is defined for sets of n-tuples as follows:[2]

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

Intuitively this operator can be understood as closing the set under the operation $\oplus$, the result M' is a collection of all items that can be constructed from elements of the original set M by applying the mereological sum operation $\oplus$.

Using this operator we can represent the codistributive reading of the sentence in (3-a) as follows:

(5)



[1]Scha actually uses the term *cumulative*, but in my opinion *codistributive* reflects better that these readings involve distribution over two arguments 'in parallel', as explained below.

[2]Since I use a functional type theory the actual definition would be not the one given here, but its (less transparent) equivalent using predicates instead of sets.

Before we turn to the derivation of such a logical form, let us check that it is indeed true in a situation where Mary faces John, Carol faces Martin, and Lucy faces Tim, and these are all men and women present. The crucial step of the calculation is the application of the $\star$-operator given in (6). 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 (3-a) is true in the described situation.

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

Now we need to describe how the logical form could be derived from the surface structure of the sentence (3-a). For this derivation the following two rules are needed:

(7) **Quantifier Raising:** Target a segment of a maximal projection XP to which first an abstractor then the raising DP are adjoined.[3]

(8) **optional $\star$-insertion rule:** Insert a $\star$-operator above any predicate.[4]

Two properties of the above rule of quantifier raising are usually not explicitly assumed, but are clearly needed for the generation of the logical form in (5). Firstly, the assumption that, along with the raising of a DP, an abstractor is generated that binds the trace that the raising operation leaves behind. Secondly, that raising cannot only target the topmost segment of a maximal category, but can adjoin to any position between the segments of a maximal category. Only these two assumptions enable us to generate the logical form (5). The steps of this derivation are the following:

1. Adjoin the abstractor $\lambda 1$ to IP and then raise *the men* to the position above it.

2. Now quantifier raising targets the position between the abstractor generated before and *the men*. Between these two, the abstractor $\lambda 2$ and then *the women* are adjoined.

3. Insert a star immediately above the two abstractors $\lambda 1$ and $\lambda 2$.

The use of quantifier raising for these examples could be described as forming the right predicate – in this case a two place predicate – for $\star$-insertion.[5] Note that the quantifier raising between an abstractor and its binder as in step 2 above has no semantic effect unless the $\star$-operator is inserted.

[3] Of course QR may apply only if the relevant locality and/or economy conditions are obeyed. These are, however, of little concern here.

[4] Instead of having this rule optionally applying, there is a possibility of having the $\star$ as an entry in the lexicon, especially in a system of incremental phrase structure generation as described in Chomsky (1994).

[5] The view that distributive interpretation is quantificational whereas collective interpretation is not, was recently supported by the findings of Avrutin (1994).

At this point the reader may wonder why quantifier raising is necessary in example (3-a) at all, since we would achieve the same interpretation by simply applying the $\star$-operator to the predicate *face*, which is already the necessary two-place predicate, and moreover that it would be sufficent to always apply the $\star$-operator to verb-heads. This is in fact Sternefeld's (1993) proposal, and works fine for the above example. But for examples like (9-a), such an account cannot generate a reading where distribution takes place twice, over two different argument positions of the verb.

(9) a. The fathers taught the ten commandments to the eldest sons.

b. For each pair of a father and his eldest son the father taught each of the ten commandments to his eldest son.

The view that definite plurals, if interpreted (co-)distributively, act similar to quantifiers has been questioned by Krifka (1992) and Srivastav (1992). Their argument is supported solely by fact that the quantifier in (10) doesn't allow a list-reading, whereas the plural in (11) does. From this contrast between a true quantifier and a definite plural they conclude that the list readings in with quantifiers is syntactically derived, whereas with plurals it is a pragmatic effect. So in (10) the list-reading is not allowed because either movement of the quantifier in Krifka's (1992) approach is illicit or binding of a variable in the complex trace in Srivastav's (1992) system is not possible. In contrast, for (11) no such syntactic restriction is at work. Hence they both conclude that the list-reading of (11) cannot be annotated at the level of logical form, and must come from somewhere else in a way they do not explain. However, the only conclusion that is really necessary is that the list reading in the two examples (10) and (11) must be represented differently, but the list-reading of (11) may still be annotated in the logical form. The above system offers a way of doing it. The list-reading of (11) can be represented by interpreting *bars* and *beers* codistributively by applying a star operator to the two-place predicate *serve*, which yields the reading in (12).[6]

(10) #Which bars serve every beer? The 'Middle East' serves Heinecken, 'Christopher's' serves Pilsner Urquell.

(11) Which bars serve these beers? The 'Middle East' serves Heinecken, 'Christopher's' serves Pilsner Urquell.

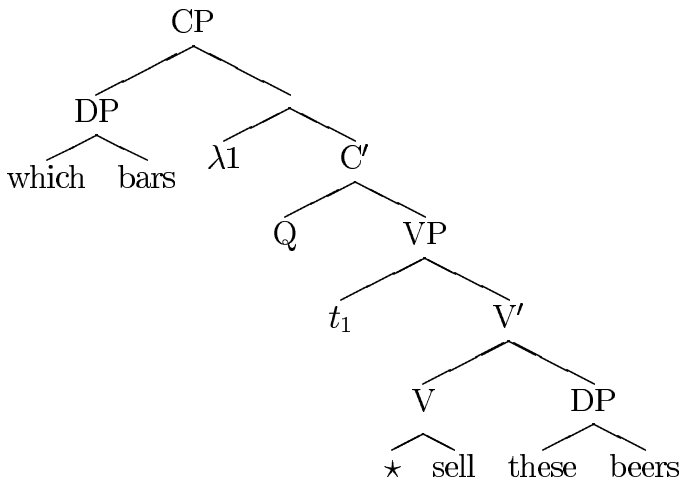
[6] A possible argument against annotating codistributivity at LF would be a lack of sensitivity to islands. The example (i-a) which is true of the situation in (i-c) could be the case in question. Actually however for a number of speakers (i-a) is degraded, which strongly supports my analysis. For the other speakers we could assume that the movement generating the necessary two-place predicate after $\star$-insertion could be focus movement.

(i) a. If the first numbers are odd, then the second numbers are even.

b. For each pair of numbers: If the first on is odd, the second one is even.

c. $(1,2),(2,3),(3,8),(4,4),(5,4),(6,7)$

(12)

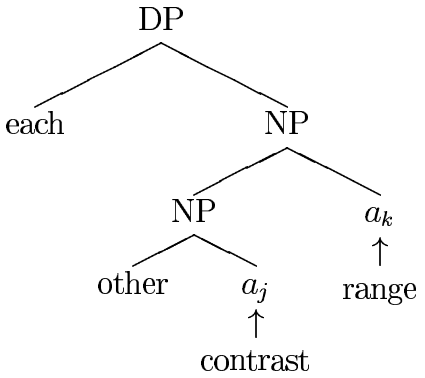


Before we start with the analysis of the reciprocal, let me point out that so far nothing specific to the reciprocal needed to be said, and with the one difference pointed out, the above analysis is in agreement with Sternefeld (1993).

3 The Representation of the Reciprocal

The internal structure I propose for the reciprocal is given in (13). It can be paraphrased as: *each one other than himself$_j$ among themselves$_k$*. The two arguments of *other* are the contrast argument a_j and the range argument a_k, where a stands for a base-generated empty anaphoric element (like a trace of DP-movement). The question whether this complex is the actual lexical entry of the reciprocal or generated in the syntax from the parts which correspond to lexical entries, is difficult to answer and of no concern here. For now I will just assume that it is a grammatical necessity for an item that has the complex referential properties of a reciprocal to have a correspondingly complex structure.[7]

(13)



[7]Under this assumption it is not surprising that a reciprocal-anaphor with a radically different surface realization like Chichewa *an* shows exactly the same behavior as English *each other* (cf. Dalrymple *et al.* (1994b)).

The semantic interpretation of *each* and *other* in this structure does not differ from freestanding *each* or *other*, namely it is:

(14) a. $[\![\text{other}]\!](x)(y)(z) = 1$ iff z is part of y and z is not part[8]of x

b. $[\![\text{each}]\!](X)(Y) = 1$ iff $\forall x(x$ smallest element of $X \Rightarrow x \in Y)$

In the following I will abbreviate the structured representation in (13) with *each* $[other(t_1, t_2)]$, and also I will never represent in the logical forms that the reciprocal, as it is headed by a quantifier, actually raises to a scope position.

The binding properties of *each other*, or rather of the two anaphors are an important part of my proposal. The contrast argument has to be bound like on ordinary reflexive, which accounts for the anaphoric behavior displayed by the complex. The range argument has to be bound by the closest binder without violating the path crossing constraint – i.e. the closest binder outside the chain of the contrast argument.[9] Let us examine how these binding statements account for the ambiguity of (15), without a need for Heim *et al.*'s (1991) *each*-movement. The sentence (15) has the two readings paraphrased in (16).

(15) John and Mary thought they liked each other.(Higginbotham (1980))

(16) a. J thought (J liked M, M liked J) and M thought (J liked M, M liked J)

[8]In Heim *et al.* (1991) the semantics for *other* are given as: z is part of y and z is not equal to x. However if we accept that the *other* that is part of *each other* and the adjectival *other* are the same lexical entry, the move from *not equal to* to *not part of* is clearly necessary. This is shown in the following example of adjectival *other* where the contrast argument, taken from the context, is a group of two people, not just a single individual.

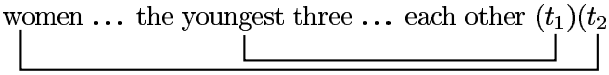
(i) Two of the three students lived in Cambridge. The other student lived in Somerville.

[9]These precautions about the binding of the range argument are necessary to avoid the following infelicitious interpretations. In (i) the binder of the range argument is not the closest possible one, in (ii) we see a violation of the path crossing constraint. Even though the binding condition on the range argument looks like a stipulation, I assume that in fact it is derived from a theory of LF-movement of the anaphors and bound variables.

(i) a. The women told the youngest three of them to give lectures to each other. (in Heim *et al.* (1991) attributed to Rooth (p.c.))

b. **The women told the youngest three of them to give lectures to each of the other women.*

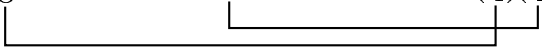
c. the women ... the youngest three ... each other $(t_1)(t_2)$



(ii) a. The youngest three of them heard them speak to each other.

b. **Each of the youngest three of them heard them speak to each of the women other than herself.*

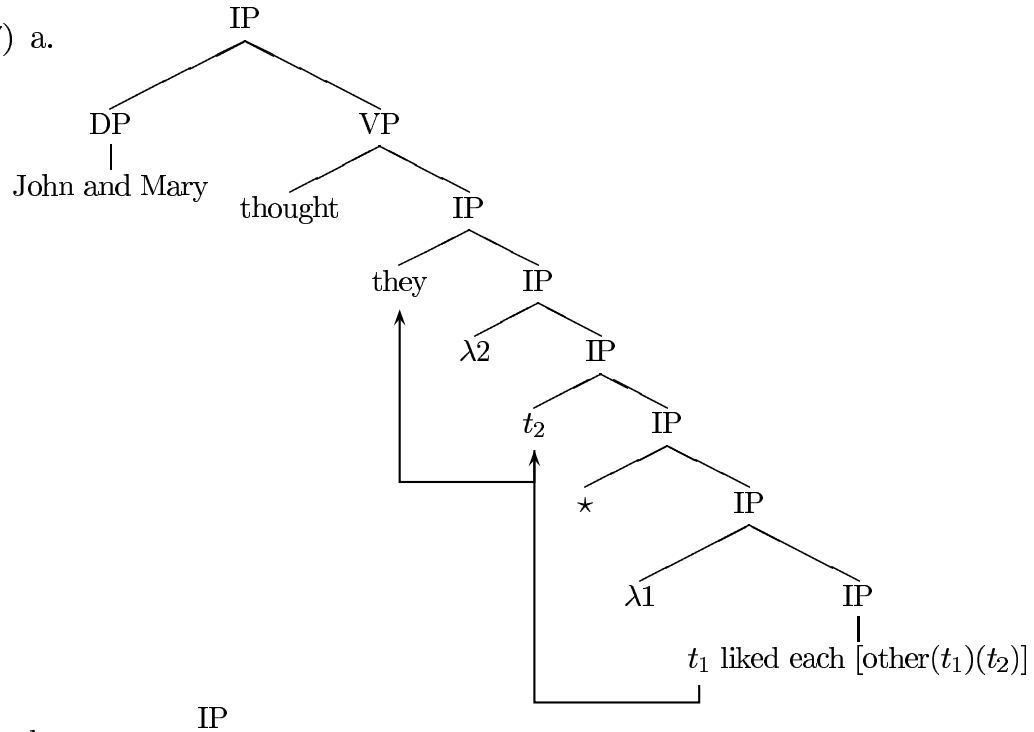
c. the youngest three ... the women ... each other $(t_1)(t_2)$



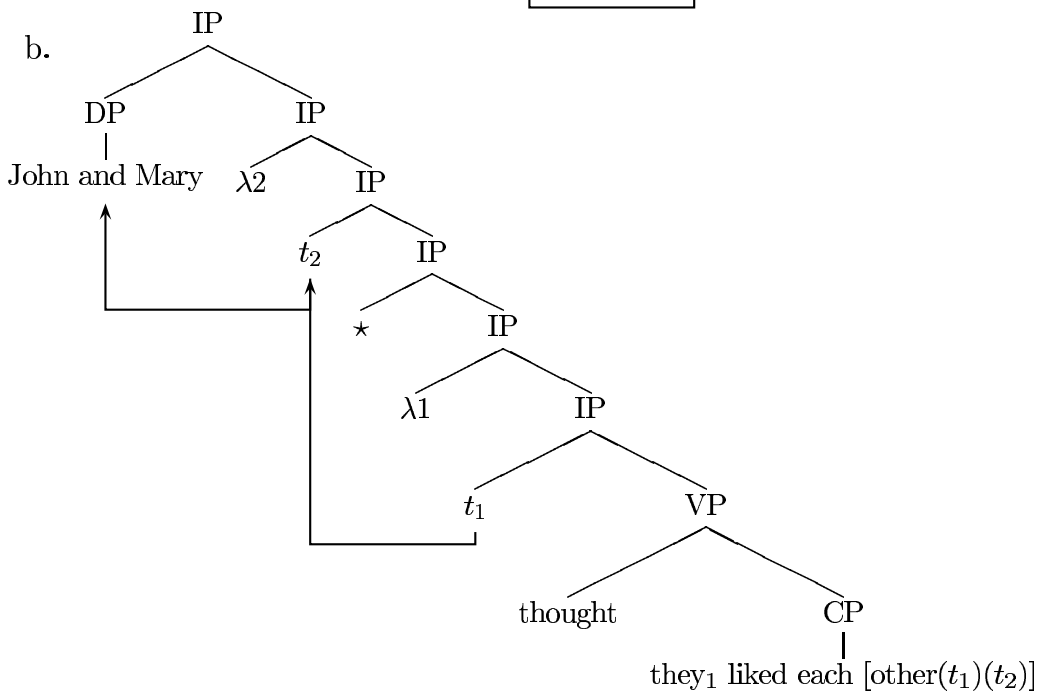
b. J thought (J liked M) and M thought (M liked J)

The (simplified) logical forms for the two readings of example (15) are the following. The important point here is that in (17-b) the local binder of the contrast argument, the pronoun *they*, is itself a bound variable.

(17) a.



b.

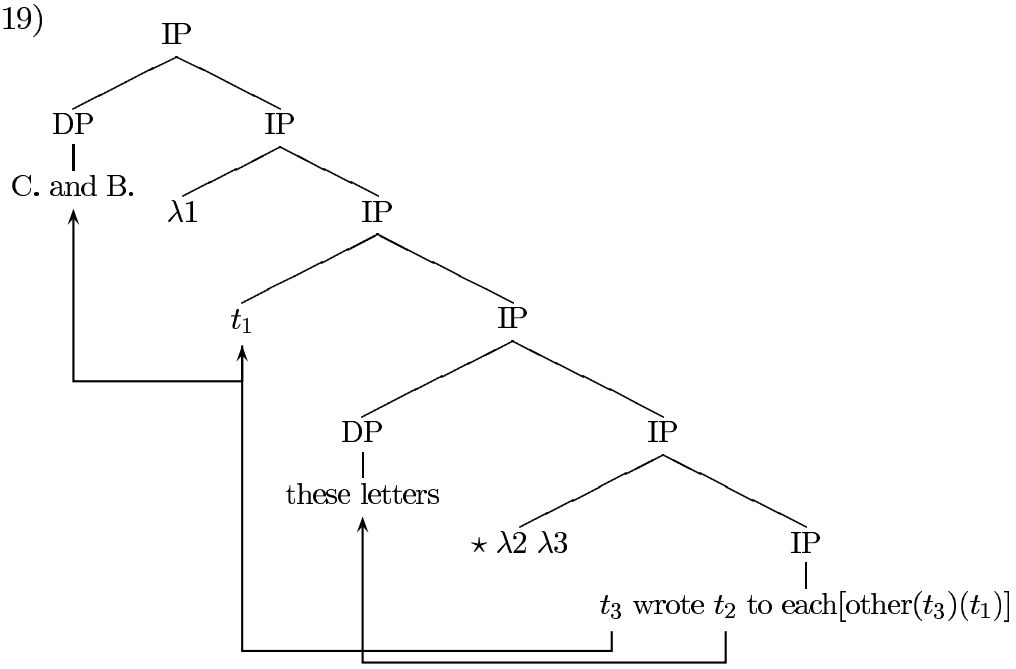


To account for Sternefeld's example (1) (repeated in (18)) all we have to do now is to put the account of codistributivity from section 2 and the

above proposal for the reciprocal together. The logical form that receives the desired interpretation is given in (19).

(18) Chandos and Byron wrote these letters to each other.

(19)



4 Capturing Pragmatic Effects

Schwarzschild (1991) 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 (20) (repeated from (3-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. Schwarzschild captures the contextual influence by defining context sensitive operators. I will here employ a simplification of Schwarzschild's account suggested by Heim (p.c.). Instead of complicating the operator definitions with context-sensitive parts, she assumes contextual restrictors that are functions from individuals or tuples of individuals onto truth values. The contextual restrictors are true of the contextually relevant individuals or tuples of individuals. In the logical form the contextual restrictors will be represented as κ_n, that adjoin to the predicates they restrict.[10] Using this idea we can acount for the contextual influence on the interpretation of (20) using the logical form in (21-a). If we assume the contextual relevance expressed by the function in (21-b) the desired interpretation arises.

[10]Note that these restrictors make a proposition logically stronger if they intersect with a part of the nuclear scope as in (21-a), but logically weaker if they intersect with the restrictive clause. Hence they are hopefully versatile enough to capture a wide range of pragmatic effects

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

(21) a. [the women] [the men] $\star\kappa_{12}\lambda 2\lambda 1[t_1 \text{ face } t_2]$

b. $\kappa_{12}(x, y) = 1$ iff x and y are a couple.

Now the question arises in which positions at logical form such contextual restrictors can occur, or rather in which positions we are driven to assume them. I will leave an answer to this question for future research (cf. Sauerland (n.d.)), and will for now assume that contextual can occur in every position. As the only restriction I will assume that due to the economy principle superfluous restrictors will never arise.

The next question is how the value of such a restrictor gets set. I will assume that there are two possibilities for this. One is that, as illustrated with (20), the restrictor reflects what is relevant or prominent in the extralinguistic context. The second possibility to set the value of a restrictor I assume is similar to the mechanism of presupposition accomodation, 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 appropiate restriction.

This mechanism enables us to account for the example (2) (see the similar (23)) from the introduction, that was a problem for the theory of Heim *et al.* (1991). More general, it offers a way to account for the *strongest meaning hypothesis* of Dalrymple *et al.* (1994a) to the extent that it is correct. 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. An example of how this works is the following: The contradictory feeling that example (22) has in contrast to (23), is explained as the fact that *know* expresses a relation that is not necessarily asymmetric, whereas *follow* expresses a asymmetric relation. Hence for the interpretation of (22) the strongest possible interpretation for the simple reciprocal sentence is chosen; i.e. the one where all pairs of non-indentical willow-school-fifth-graders have to stand in the relation *know*. For the interpretation of (23) 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.

(22) #The willow school fith graders know each other, but the oldest doesn't know the youngest.

(23) 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, independently of what my proposal forces me to say. A sketch of how this effect, to the extent that it is correct, can be derived from pragmatic principles, goes as follows: The two pragmatic principles that are relevant are roughly stated the following: Be charitable; try to enable a true interpretation. The antagonist of this principle is: Be economical; don't insert pragmatic operators. The interplay of these two principles then has to ensure that in a neutral context no restrictor is inserted for example (22), whereas for example (23) the relevant restrictors are inserted. Then for example (23) the above process of restrictor-accommodation takes place, which results in an interpretation of (23) that says that the children went in an order, but this is order is not specified.

References

Avrutin, Sergey. 1994. Distributivity and binding in child grammar. *Linguistic Inquiry* 25.165–171.

Baker, Mark C., Kyle Johnson, & Ian Roberts. 1989. Passive arguments raised. *Linguistic Inquiry* 20.219–251.

Chomsky, Noam. 1994. Bare phrase structure. Occasional Papers in Linguistics 5, Massachusetts Institute of Technology, Cambridge.

Dalrymple, Mary, Makoto Kanazawa, Sam Mchombo, & Stanley Peters. 1994a. What do reciprocals mean? In *Proceedings of SALT IV, University of Rochester*, 61–78.

——, Sam Mchombo, & Stanley Peters. 1994b. Semantic contrasts and syntactic similarities between Chichewa and English reciprocals. *Linguistic Inquiry* 25.145–163.

Heim, Irene, Howard Lasnik, & Robert May. 1991. Reciprocity and plurality. *Linguistic Inquiry* 22.63–101.

Higginbotham, James. 1980. Reciprocal interpretation. *Journal of Linguistics Research* .

Krifka, Manfred. 1992. Definite NP's aren't quantifiers. *Linguistic Inquiry* 156–163.

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

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. Reprinted in Lewis (1983).

—— 1983. *Philosophical Papers*, volume I. Oxford: Oxford UP.

Link, Godehard. 1991. Plural (Plural). In *Semantik: Ein internationales Handuch der zeitgenössischen Forschung* (Semantics: An International Handbook of Contemporary Research), ed. by Arnim von Stechow & Di-

eter Wunderlich, number 6 in Handbücher zur Sprach- und Kommunikationsforschung, chapter 19, 418–440. Berlin: Walter de Gruyter.

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

SAUERLAND, ULI, n.d. Reciprocals and codistributivity. forthcoming in the Proceedings of WECOL 1994.

SCHA, REMKO. 1984. Distributive, collective, and cumulative quantification. In *Truth, Interpretation, and Information, GRASS 2*, ed. by Jeroen *et al.* Groenendijk, 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.

SRIVASTAV, VENEETA. 1992. Two types of universal terms in questions. In *Proceedings of NELS 22 at the University of Delaware*, 443–457, Amherst. GLSA.

STERNEFELD, WOLFGANG. 1993. Plurality, reciprocity, and scope. SfS-Report 13-93, Universität Tübingen, Tübingen.

Author's address: Uli Sauerland
Room 20C-231B
MIT, Cambridge MA 02139
sauerlan@mit.edu