Plurality, Reciprocity, and Scope

Wolfgang Sternefeld, Universität Tübingen, March 1994

1. Introduction

1.1. Aims and Goals

"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 ...". Dating back more than fifteen years, Langendoen's (1978) article, from which the above quote is taken, is a benchmark for the considerations in this paper. This is so "despite recent pioneering efforts" in present day literature, which, to my mind, all but reveal that the above quote still proves true today.

Langendoen – arguably the first to account for the "logic of reciprocity" in a really insightful way – successfully identifies an intimate logical relation between what he calls "elementary plural relational sentences" like (1-a) and "elementary reciprocal sentences" like (1-b) and (1-c):

(1) a. John and Barbara had relations with Jane and Bill.
b. John and Barbara had relations with one another.
c. The men and the women had relations with each other.

This relation can superficially be brought to light by replacing *one another* by *Barbara and John* and *each other* by *the women and the men.* It can then be seen that the resulting sentences can be interpreted in much the same way. By pointing to the fact that the interpretative mechanism to be seen at work in (1-a) also carries over to the interpretion of (1-b) and (1-c), Langendoen correctly hints at the importance of the "cumulative" reading of (1-a) (which, in the case at hand, can be enforced by adding *respectively*). It is this property of plural semantics that will be recognized as central to the semantics of reciprocals discussed in this paper.

Although I have unduly oversimplified, the above sketch of Langendoen's analysis should nourish the intuition that there is something essential in the interpretation of plural sentences that all the constructions of (1) have in common, regardless of whether or not they contain a reciprocal operator. The lesson to be learned from this parallelism is that, contrary to what has been proposed in more recent analyses (cf. e.g. Moltmann 1992), the reciprocal adds only relatively little to the meaning of the whole. I regard this insight as a decisive clue for a proper understanding of these constructions, and one of the major tasks of the present article is to clarify exactly how the meaning of the reciprocal should be analyzed and how it fits into a compositional analysis.

Enlightening though Langendoen's article is, it cannot be claimed that his analysis already solves all problems. In particular, the relation between surface expressions and the proposed logical paraphrases has been left unaccounted for, and, indeed, the phenomena under discussion have often proven a paradigm case for those who cast doubts on compositional analyses altogether. And it is true that a strictly compositional semantics is impossible. This does not mean, however, that any analysis must be ad hoc; another important task of the paper is to show how non-compositional features of lexical representations can be embedded into a general account of Logical Form and semantic representation.

As a large number of different factors always need to be sorted out when translating from language into logics in a systematic way, it is important to isolate what is specific to the constructions under discussion. In that respect the following intriguing problems arise:

Since reciprocal sentences are invariably plural sentences, how can we identify the specific semantic contributions of "pluralization" vs. "reciprocalization" to the meaning of the sentence? Or put differently: What is the "differential" contribution to meaning made by the reciprocal expression – a contribution not yet contained in the semantics of "pluralization"? How does "pluralization" combine with anaphoric binding in reciprocal constructions? Since both reciprocal sentences and plural sentences are often felt to be multiply "ambiguous," is there an implicational relation between these ambiguities? Is the purported ambiguity semantic in nature, or can it be resolved by pragmatic considerations? How can we deal with the problems of compositionality that arise with plural morphology in general, and with the problems posed by the reciprocal constructions in particular?

Answers to all these questions will have a direct impact on the semantic analysis of reciprocal expressions, and it will be the task of this paper to disentangle some of these interacting factors and clarify their role within an explicit theory about the syntax/semantics interface. In particular, the above mentioned phenomena will shed light on the problem of how tight the relation between surface structures and semantic representations can be, even when taking into consideration intermediate levels like Logical Form (LF). To illustrate, consider a model for the sentence in (1-a), which (in one of its construals) is true because John had relations with Jane, and Barbara with Bill. Such a model also supports the truth of (2):

(2) Two men had relations with two women.

Under this construal, the two quantificational expressions *two men* and *two women*, when treated as generalized quantifiers, cannot have scope with respect to each other. Now, as far as LF is concerned, one of the most elementary problems – albeit one that to the best of my knowledge has not yet been tackled in the literature – is this: How can the scopeless reading be represented at that level of representation?

Examples like (2) reveal that, contrary to generalized quantifier theory as pursued, e.g., by van der Does (1992) among others, the distributional force of plural quantifiers should not be analyzed as a compositional part of the NP-meaning. The representation of "scopeless quantification" will then depend on a proper treatment of pluralization and scope in general. The proposal put forth in this paper is that what is generally considered the scope inducing part of the numerical expressions cannot be considered part of the meaning of *two*, but must be regarded as a part of the meaning of pluralization.

This thesis has far-reaching repercussions with respect to virtually every question raised so far. For instance, it will turn out that the distributive part usually associated with the meaning of *each other* cannot be so analyzed, i.e. cannot be part of the meaning of the reciprocal. Rather, the quantificational force attributed to *each* must be taken as part of the meaning of pluralization in general. We will see below that it is essential to *not* tie together the quantificational force of the construction with the meaning of the reciprocal operator. In particular, as pointed out to me by Irene Heim (p.c.), the traditional way of treating *each* as a distributive operator (as, e.g., in Heim, Lasnik, & May 1991) will not permit for any correct semantic analysis of simple sentences like (3):

(3) I read the letters they wrote to each other.

We will see below why this should be so. On the other hand, even if quantificational ingredients of meaning are not part of the NPs involved, they cannot be part of the predicate either, as we will see. Thus, the basic claim to be defended in this article is that "pluralization" is kind of an a-morphous semantic process that is not compositionally tied to plural morphology. Rather, semantic pluralization is only partially dependent on the presence of morphology – a point that will be elaborated upon in the sections to follow. Thus, as pointedly expressed by Urs Egli (p.c.), the basic idea is to conceive of plural morphology as a kind of "suprasegmental" phenomenon: the semantic operation of pluralization will be "stuffed in" where needed in semantic interpretation (e.g. as a modification of predication), rather than being induced by a rule-by-rule process that interprets syntactic structure in a Montegovian manner.

The above theorizing automatically leads us halfway between model theory and surface syntax: it seems that there must be an intermediate level of representation where appropriate operators are inserted before a semantic interpretation can be achieved. Of course, this level is the level of LF. But unlike most proposals of similar provenance, I will propose that this level can directly be interpreted by unambiguously representing truth conditions. The resulting kind of representation has sometimes been called *transparent logical form* (cf. v. Stechow 1993); it will be dubbed *Augmented Logical Form* (ALF) in the present paper. Intuitively, ALFs take ordinary LFs (or rather variants thereof) as a point of departure and add to them all that is needed in order to get a disambiguated, fully interpretable semantic representation.

This conception of LF differs from the standard account in more than only disambiguating ordinary LFs. To grasp the qualitative difference, let us restate how non-compositionality motivates this model of grammar. Strictly compositional semantics is driven by word meaning and the interpretation of syntactic configuration alone. Under standard assumptions, LFs are usually obtained from S-structure by purely syntactic manipulation, e.g. by quantifier raising, reconstruction, copying, etc., but not by supplementing any kind of genuine semantic operations. In contrast, the construction of ALFs is not governed by structure and lexical meaning; rather, it is driven by the attempt to find some kind of semantic glue that helps to put the meanings of two adjacent LF-constituents together. This glue is meaning beyond word meaning, it does not deterministically depend on syntactic structure, and it comes for free, i.e., it is part of Universal Grammar, in being a part of the general UG-process of building up ALFs.

Thus, the role (non-)compositionality plays in the theory is twofold. First, non-compositional phenomena shape the present conception of ALF as being augmented by elements drawn from linguistic context rather than from the lexicon. Second, it will be seen that the way ALFs are interpreted is completely compositional (in the same sense as ordinary predicate logic is). Among semanticists, this is usually taken for granted; but what is usually called LF in the literature cannot be so interpreted, unless we add substantial restructuring and dubious semantic operations like "absorption," which are essentially non-compositional in nature (cf. Higginbotham & May 1981, May 1985, Keenan 1987, or Chierchia 1992). In what follows it will be presupposed that all this is unnecessary once we have constructed the level of ALF.

1.2. Overview

The organization of the paper is as follows. Section 2 is devoted to the development of a semantic theory that can serve as the basis for the following discussion. As usual, it is most economical to adopt and modify an already existing theory, and the one that seems most suitable for our present purposes is that of Schwarzchild (1991). I will briefly summarize this work, without much discussion, but with a few elaborations and additional remarks that will prove useful for later developments.

Having laid the grounds for dealing with plural objects and plural predication, Section 3 analyzes elementary reciprocal sentences. It will be shown that the meaning of the reciprocal splits up ("non-compositionally") into two parts: an anaphoric relationship and a non-identity statement $x \neq y$. Beyond these two components the only additional ingredients to represent meaning are in fact syntactic, i.e., we have to relate the two parts in a proper way, e.g., by the use of appropriate indices. A certain non-compositionality of reciprocal binding has traditionally been acknowledged in theories of "*each*-movement," cf. Heim, Lasnik, and May (1991) for recent discussion. As pointed out above, however, the present theory differs substantially from this and previous accounts in attributing the alleged distributive function of *each* to the semantics of plural predication, rather than to (part of) the meaning of the reciprocal NP.

Section 4 is devoted to a description of ALFs. Although the concept as such seems neither original nor new,[1] I believe that the approach taken here involves a change of perspective that makes it worthwhile to reformulate traditional ideas in a more theory governed context. In particular, I will explicitly account for the relation between surface representations and ALFs. This will be done in the spirit of "fragment linguistics," i.e., it will be carried out (incompletely) for a very minimal fragment of English only, ignoring all kinds of problems that may arise once we strive away from the narrow concerns pursued in the foregoing sections.

Section 5 applies the theory to yet another area of problematic scope interaction, exemplified by plural questions like the following:

(4) a. Which books did most students read?
b. Which books did a majority of students read?
c. Which book(s) did every student read?

The basic contrast to be accounted for is one of scope: the quantified NP *most students* in (4-a) must have a wide scope interpretation with respect to the *wh*-operator, whereas only a narrow scope reading is available for *a majority of students* in (4-b). This contrast will be discussed against the background of list readings and the Hamblin/Karttunen semantics of questions. I will try to motivate that there are two types of quantifying-in. Both *most students* in (4-a) and *every student* in (4-c) can have wide scope over *which books*, but for reasons to be explained the scopal domains of universal and existential quantifiers must differ.

In consequence, the semantic interpretation of question formation must be split up into several components, each of which is delimited and determined by the sco-

[1] Cresswell (1973) uses logical notation in much the same way as I do, but offers no rules on how to generate them; in particular, there is no syntactic component in his theory. Fanselow (1985) develops the idea of semantic glue in his semantics of compounds. Stechow (1993) uses somewhat impoverished ALFs (or logically enriched LF), but does not explicitly state the rules that make his representations different from ordinary LFs. Generative Semanticists have stated rules of that sort, but have had a somewhat different conception of the role of LF in the organization of grammar.

pal domains of certain operators. One domain is identified by the scope of *wh*-operators so that *a majority of students* in (4-b) must be interpreted within that domain.[2] Another domain is identified by the scope of distributive operators when being "quantified-in," as exemplified by the relevant reading of (4-c).[3] The semantic representation thus obtained will define a paired list answer much as in "which student read which book." And finally a domain is delimited by quantifying-in existential plural NPs that do not give rise to list answers, as illustrated in (4-a).[4]

I do not want to close this overview without expressing my gratitude to all who have read and commented on various parts and versions of the manuscript. The contribution of most people will be made explicit in the text or in footnotes, but there are a few whose invaluable influence still remains concealed. Most importantly, I would like to thank Irene Heim for her encouragement and for pointing out to me serious errors in early but infelicitous attempts to define a reciprocal operator. Manfred Kupffer and Sigrid Beck deserve special thanks for having detected several flaws in my presentation of the semantics of questions. Renate Musan, Arnim v. Stechow and Kirsten Brock commented on major parts of the manuscript and drew my attention to further inconsistencies, confusion, and awkwardness of presentation.

Needless to say, revised manuscripts do not always ameliorate, and in many cases I have not taken the advice I have been offered so generously. Thus, responsibility must be placed on me for whatever has been changed, and what has not.

---

[2] The associated meaning component of question formation can be represented by an abstract question morpheme located in C.

[3] The associated meaning component cannot be expressed as a morpheme, but nonetheless it can intuitively be associated with the performative paraphrase "I ask" as in "for each student *I ask* which book(s) he read."

[4] This domain is different from the one described in the last footnote, since "for most students I ask which book(s) each of them read" is not a felicitous paraphrase; it should not imply a list answer as in the case of quantifying in *every student*. Rather, the semantic rules laid down below will imply something like "given a certain majority of students who read the same books I ask which books they read."

2. Basic Assumptions

2.1. Simple Noun Phrases

Assume that a, b, c are girls in a domain D of entities and *John* is a name for $j \in D$. Schwarzschild (1991) has convincingly shown that a plural NP like *John and the girls* denotes a subset of D, namely $\{a, b, c, j\}$, rather than a two element set containing John and the set of girls, i.e., $\{j, \{a, b, c\}\}$. According to Schwarzschild, the denotation of plural terms like the one above is always a subset of D, and the semantics of NP-conjunction is argued to be set union: if *John* denotes the set containing John, and *the girls* denotes a set of girls, term conjunction is set theoretical union.

An important ingredient of Schwarzschild's union theory is what he calls "Quine's innovation," i.e., a set theory that identifies an individual with the singleton set containing it. Thus, $j = \{j\} = \{\{j\}\}$, etc. Although nothing really hinges on it, Quine's innovation will prove extremely helpful in uniformly stating the rules we will come across in what follows. In particular, term conjunction can be described uniformly as set union, e.g. as $j \cup \{a, b, c\}$, regardless of whether the conjuncts are individuals (i.e. elements of D) or proper sets or collections.[5]

Another central part of the theory is that all lexical items of category N denote subsets of D. In particular, NPs like *the group*, *the committee*, or *the collection* denote individuals, rather than sets. One of the consequences of this decision, in fact one that has carefully been motivated in Schwarzschild's dissertation, is that there must be a "metalinguistic" membership relation between elements of D. Within the theory to be developed in this paper, this has another important consequence: the decision to interpret all noun extensions as simple sets makes it possible to maintain a very tight relation between plural morphology and the semantic type of denotation of nouns. Below it will be argued that the same cannot be true for verbs, so that the plural semantics must be anchored in the morpho-semantic pluralization of nouns, rather than of verbs. In other words, plural semantics for nouns will be largely

[5] I have been warned repeatedly that Quine's set theory is apparently inconsistent with Lewis' (1991) *Parts of Classes* and rests on shaky formal grounds. These objections can of course be overcome by simply reformulating the rules and translating them into one's favorite ontology. A. v. Stechow has done so in class lectures. For him, the result was fine since he has shown that the theory is compatible with his ontological prejudices. For me, the result was so cumbersome that I became convinced not to presuppose the mereological theory that motivated all the reformulations. However, there is no accounting for likes and dislikes.

compositional, i.e., directly induced by plural morphology, but plural semantics for verbs is an epiphenomenon that cannot be handled in a strictly compositional way, as we will see.

In general, then, plural common nouns denote sets of sets.[6] In order to compute the denotation of a plural CN from a singular CN, I assume a function * that works as follows. If $\{a, b, c\}$ is the extension of *girl*, the extension of *girls* is $*\{a, b, c\} = \{a, b, c, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$. The function * can be defined recursively as follows:

(5) For any set P, $*P$ is the smallest set such that
  a. P is a subset of $*P$, and
  b. if $x \in *P$ and $y \in *P$, then $(x \cup y) \in *P$.

The function thus described is only a subcase of what I call "cumulative pluralization" (CP). The generalization of CP will be carried out for arbitrary n-place relations and can be stated by the following condition: the result of applying cumulative pluralization to a relation R is the smallest relation which includes R and is closed under accumulation of members of the relation. Given a two-place relation R, its pluralization will be denoted as $**R$. This is the smallest relation R' which contains R and whenever $R'(a, b)$ and $R'(c, d)$, then $R'(a \cup c, b \cup d)$. The general case is formally defined in (CP), where the arity of the relation R is encoded in the number of stars preceding the relation.[7]

(CP) $$*_1 \cdots *_n R := \bigcap\{Q : R \subseteq Q \wedge (\forall x_1, ..., x_n)(\forall y_1, ..., y_n)[(\langle x_1, ..., x_n\rangle \in Q \wedge \langle y_1, ..., y_n\rangle \in Q) \rightarrow \langle x_1 \cup y_1, ..., x_n \cup y_n\rangle \in Q]\}$$

For $n = 1$, I will illustrate (CP) in the next section. The case of $n = 2$ will become relevant in section 2.3., and $n = 3$ is of utmost importance in section 3.2.2.

---

[6] However, in a model with only one girl, an expression like *the girls* will still be meaningful: it denotes the one girl of the model, which turns out to be identical with the extension of the CN *girls*. Schwarzschild (1991) argues that expressions like *no girls* can be handled properly only if the term *girls* is allowed to denote a singleton set, i.e., an individual. The above claim that plural CNs always denote sets of sets must therefore be modified. I will take up the issue in sections 2.2.2. and 3.2.1.

[7] Two anonymous reviewers of *Linguistische Berichte* attribute (CP) to Link (1983). They are wrong. First of all, Link only defines * for one-place predicates. Moreover, he simply defines $*P$ as the power set of P minus the empty set. Although this is extensionally equivalent when P denotes a set of individuals, the results will crucially differ as soon as P is more complicated; cf. below. –

I got the idea of generalizing cumulative pluralization from Manfred Krifka (p.c.), but it seems to have been around for a while, cf. Krifka (1986) for an early reference.

For the remainder of the present section however, let us briefly take up again our first example *John and the girls* and calculate its denotation. We already know that the result is $\{j, a, b, c\}$, which has been obtained by conjoining j with the meaning of *the girls*. Thus, it only remains to be shown how the meaning of *the girls* can be calculated from the meaning of *girl*. We stipulated above that *girls* denotes $^*\{a, b, c\} = \{a, b, c, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$ in the given model, hence all that remains is to describe the semantics of the definite article. Following Sharvy (1980), the definite article *the* selects from a noun denotation – no matter whether in the singular or plural – the set with greatest cardinality if there is one. Due to Quine's innovation, an individual has cardinality one, so that the denotation of a singular NP like *the girl* only exists if the extension of *girl* (a subset of D) meets the uniqueness requirement; this will be the case only if it contains exactly one element, i.e., an individual. If the argument of *the* is a plural CN, as in *the girls*, one has to look for the maximal element in $\{a, b, c, \{a, b\}, \{b, c\}, \{a, c\}\{a, b, c\}\}$, which is $\{a, b, c\}$. Hence, the denotation of *the girls* equals the denotation of *girl*.

Since I do not intend to develop a semantics for NP meanings in general, I leave aside quantificational NPs here. The above remarks and basic assumptions will suffice for dealing with simple predication in the next section.

2.2. Simple Predication

2.2.1. Semantics

Let us turn now to the denotation of VPs. Clearly, a verb like *gather* cannot apply to singularities, although its denotation will contain elements of D: recall that D also contains groups (e.g., the denotation of the CN *committee*). Hence, predicates like *gather* or *lift a piano* will apply to all sorts of entities in $^*D$, rather than to members of D alone.

For an illustration of standard examples, I follow Gillon (1987) in considering a model $\mathcal{M} = \langle D, f\rangle$ in which Rodgers and Hart as a team, and Rodgers and Hammerstein as a team have written musicals. To simplify things a bit, I assume that the predicate *write-musicals* is a lexical expression (i.e., is not decomposed into a transitive verb plus an object); moreover, I ignore tense throughout the paper. The interpretation function f of $\mathcal{M}$, which assigns semantic values to lexical expressions, is supposed to assign the set {{Rodgers, Hart}, {Rodgers, Hammerstein}} as the denotation of the VP *write-musicals* in $\mathcal{M}$. In this model, a sentence like (6-b)

(6) a. Hammerstein, Rodgers, and Hart are composers.

b. They have written musicals.

is intuitively true. Unfortunately, however, it does not hold that the set referred to by *they* is a subset or element of the denotation of the predicate.[8] In order to analyze (6-b) as true, Gillon proposes thinking of predication as being mediated via cumulative predication, i.e., he proposes to analyze (6-b) as (7):

(7) $f(\mathit{they}) \in {}^*f(\mathit{write\text{-}musicals})$

Clearly, by adding the cumulative operator, the predication (or functional application, which – as a matter of convenience – we represent as "$\in$") will become true. This is because ${}^*f(\mathit{write\text{-}musicals}) = {}^*(\{\{\text{Rodgers, Hart}\}, \{\text{Rodgers, Hammerstein}\}\}) = \{\{\text{Rodgers, Hart}\}, \{\text{Rodgers, Hammerstein}\}, \{\text{Hammerstein, Rodgers, Hart}\}\}$. We will see below that some kind of "addition" to the usual predication relation not only gives the correct truth conditions in cases like cumulative predication, but will also serve as the semantic basis for plural predication in general.

Although the example chosen by Gillon nicely illustrates cumulative predication, it is nevertheless somewhat artificial, since the predicate *write-musicals* can be lexically decomposed. That this might make a crucial difference will emerge when considering some formal properties of verb denotations. To do so we first introduce the notion of a cover. A set of set Y *covers* X iff $X = \bigcup Y$ and $\emptyset \notin X$. A cover is minimal if it does not contain any sets a and b such that $a \subset b$. Assume now that a predicate expression P and the interpretation function of a model satisfy the following condition:

(8) There is an $X \subseteq D$ and a minimal cover Y of X such that $f(P) = Y$.

In other words, if $f(P) \neq \emptyset$ and $X \in f(P)$, then there are no subsets $Y \subset X$ or supersets $Y \supset X$ such that $Y \in f(P)$. If this property holds of P for any interpretation function and model, I shall say that P is *C-reducible*.[9]

---

[8] One might reject f as inadequate and add a meaning postulate to the effect that whenever $X \in f(\mathit{write\text{-}musicals})$ and $Y \in f(\mathit{write\text{-}musicals})$ it holds that $X \cup Y \in f(\mathit{write\text{-}musicals})$. This method has a number of drawbacks, however, among them the well-known fact that under this analysis it would become impossible to express the group reading, which would require that all three composers jointly wrote musicals. In this reading, the sentence would (correctly) turn out false in Gillon's model, but if one adopted the above meaning postulate, it would come out true, contrary to what the group reading should intuitively imply.

[9] The term "reducible" has originally been derived from considering denotations of plural predicates which may be *reduced* to that of singular predicates with certain properties. For instance, the term D-reducible which will be introduced below was meant to hold for a plural property P iff the statement "$P(X)$" can be reduced to (or logically implies) the first order quantificational sentence that each $x \in X$ has a corresponding singular property Q. Since reducibility is in effect not a specific property of the plural predicate P but rather one of a corresponding predicate Q, I started to use the term "reducibility" for the denotation of Q. Although "representability" or "expressibility" might be more appropriate terms I suggest to keep to the use of "reducibility" in order to indicate

My contention is that all lexical items of natural language are C-reducible, perhaps with the exception of some that have to do with measurement. What rules out *write-musicals* as a lexical item is that it is not C-reducible. For suppose Rodgers and Hart together wrote the musical *On Your Toes*, but, counterfactually, Rogers alone wrote *Oklahoma!*. Since {Rogers} $\subset$ {Rogers, Hart}, the denotation of *write-musicals* is not C-reducible. But of course the denotation of the real lexical item *write* is C-reducible in the generalized sense:

(9) An n-place relation R is C-reducible iff the set of values for y that satisfy $R(x_1, ..., y, ...x_n)$ for any fixed values of $x_1, ...x_n$ forms a minimal cover.

Once we fix, e.g., a certain musical, either a group has composed it together, or a subgroup or an individual has composed it, but not both. The same is true when fixing a (set of) composer(s): as regards musicals, it seems that the result of composing at a time is always an individual, rather than a set. And so it is in general. Thus we conclude that *write* is C-reducible.[10]

Let us say that a predicate is D-reducible if in every model its denotation is a subset of D. Above we hypothesized that all genuine (non-derived) nouns are D-reducible. A two place relation is D-reducible if it denotes a subset of $D \times D$; an n-place relation is D-reducible if it is a subset of D^n. Relational nouns need not be D-reducible, e.g., *owner-of*, and some even might not be C-reducible. Moreover, it may happen that relations are D-reducible only with respect to certain argument positions. For instance, we just conjectured that *write* is D-reducible with respect to the object position, besides being C-reducible with respect to both argument positions.

Let us turn now to the relationship between cumulative predication and the existence of a *cover* of the subject denotation. The statement that Y is a cover over

---

that it is this property that enables one to state logical relations between the denotations of certain plural and singular properties.

[10] Note that the verb in

(i) The babies weigh less than 280pds.

– although belonging to the realm of measurement – is still C-reducible, since *less than 280pds.* itself is to be decomposed. Likewise, the degenerate case of non-reducability with a sentence like

(ii) John and Mary make \$2000 per month

is only apparent. (ii) might be considered true even if Mary doesn't earn anything, so that the VP will also be true of John alone. But if the above statement is the cumulative result of "John makes \$2000 per month" and "Mary makes \$0 per month," the basic predicate involved is again reducible. Some predicates like *heavy* appear not to be reducible, but again it seems that an atomistic analysis can decompose it into C-reducible parts.

X will be written as $C(Y, X)$ and I will sometimes use C_X to refer to a particular cover Y. Given an arbitrary one-place predicate P it is obvious that the following equivalence is logically true:

(10) $X \in {}^*P \Leftrightarrow (\exists Y)\, C(Y, X) \wedge (\forall z \in Y)\, z \in P.$

We therefore may rephrase the truth conditions (11-a) of *they write musicals* as (11-b):

(11) a. $f(\textit{they}) \in {}^*f(\textit{write-musicals})$
b. There is a cover $C_{f(\textit{they})}$ such that each element of $C_{f(\textit{they})}$ is in $f(\textit{write-musicals})$.

The transition from (a) to (b) may turn out to be advantageous, since *distributive predication*, which is paraphrased as *they each wrote musicals*, can now be expressed by formulating a certain restriction on (11-b). Suppose we say that the cover whose existence is ascertained in (11-b) has to take a special form: in the distributive construal the relevant cover over $f(\textit{they})$ must consist of singularities only. Given Quine's innovation this means that the truth conditions for the distributional construal can be obtained via (11-b) by adding the requirement that $C_X = X$.

The "group reading," which can be paraphrased as *they all wrote musicals together*, can be expressed by assuming that C_X contains only the one element X, so that $C_X = \{X\}$. Of course, this amounts to the same as simply analyzing the sentence as $f(\textit{they}) \in f(\textit{write-musicals})$. And of course, the distributional reading could also be described by a new operator Δ as defined in (12-a):

(12) a. $X \in {}^{\Delta}P$ iff $(\forall x \in X)\, x \in (D \cap P)$.
b. $f(\textit{they}) \in {}^{\Delta}f(\textit{write-musicals})$

The relevant point here is that these readings need not be expressed by separate operators, but can be *subsumed* under a unifying approach that uses trivial kinds of covers. I will come back to and comment on these alternative ways of representing different readings in the next subsection.

2.2.2. Pragmatics

According to what we have said so far, elementary plural sentences may translate into various logical representations, e.g. the ones given in (13):

(13) a. The boys lifted a piano
b. $\iota({}^*f(\text{boy})) \in [\![\text{lifted a piano}]\!]_{\mathcal{M}}$
c. $\iota({}^*f(\text{boy})) \in {}^*[\![\text{lifted a piano}]\!]_{\mathcal{M}}$

d. $\iota(*f(\text{boy})) \in {}^{\Delta}[\![\text{ lifted a piano }]\!]_{\mathcal{M}}$

The cumulative reading is (13-c), and (13-b) is the group reading. It is worth mentioning that (13-b) does not exhibit any semantic counterpart of pluralization on the verb; hence, (CP) can apply to verbs optionally.[11] Likewise, it seems important to realize that there is nothing in the truth conditions stated above that would make (13-a) false in a case where there is only one boy. The fact that (13-a) is felt inadequate in such a situation can be derived from the Gricean maxim of manner: if there is reason to believe that an alternative but more informative sentence, i.e., *the boy lifted a piano*, is true, this sentence must be uttered in order to convey that meaning. We therefore assume that related singular and plural NPs form expression alternatives on a quantitative scale, such that *the boys* will conventionally implicate "more than one boy." This implicature may, however, be overridden by other factors, one of them being the phenomenon of dependent plurals, to which I return in section 3.2.1.

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. To illustrate such an intermediate case, imagine

> [...] a situation in which two merchants are attempting to price some vegetables. The vegetables come in various varieties and they are piled in baskets. To determine their price, the vegetables need to be weighed. Unfortunately, our merchants do not have an appropriate scale. Their grey retail scale is very fine and is meant to weigh only a view vegetables at a time. Their black wholesale scale is coarse, meant to weigh small truckloads. Realizing this, one of the merchants truthfully says:
>
> 159) The vegetables are too heavy for the grey scale and too light for the black scale.
>
> In order to save space in our explanation, let us reword his utterance:

---

[11] As pointed out above, one might assume that even the group reading requires some (vacuous) kind of pluralization operator defined by the special cover $\{\{X\}\}$. It turns out, however, that in the system to be described below some additional effort would be necessary to exclude (13-b) as a possible analysis. I therefore will not adopt this alternative.

If pluralization on the verb can apply without a morphological trigger, (i-a) also has two analyses:

(i) a. The boy lifted a piano

b. $\iota(f(\text{boy})) \in [\![\text{ lifted a piano }]\!]_{\mathcal{M}}$

c. $\iota(f(\text{boy})) \in *[\![\text{ lifted a piano }]\!]_{\mathcal{M}}$

But since $P \subseteq *P$, we do not get wrong truth conditions. Pluralization in (i-c) simply applies vacuously, and as such it can be ignored.

160) a. The vegetables are too heavy for the grey scale.
b. The vegetables are too light for the black scale.

> (160a) is false on its distributive reading [...] It is true on its collective reading but that is not what the merchant intended. (160b) is false on the collective reading [...] It is true on its distributive reading, but again that is not what the merchant intended to say. The physical arrangement of the vegetables in baskets suggests a plurality-cover of the vegetables with cells of the cover corresponding to baskets full of vegetables. (159) is true and informative on the intended intermediate reading because the verb is true of every member of that cover. (p. 107f)

Schwarzschild reformulates the truth conditions of simple plural predication in such a way as to make the cover part of the context. Modifying slightly his notation, I rename his contextual variant of * as "©" and rephrase his truth conditions as (14-a), which, by assuming C to be part of the model at a context, can be abbreviated as (14-b):

(14) a. X $\in$ ©P iff there is a contextually determined cover $\mathcal{C}_X$ such that each element of $\mathcal{C}_X$ is in P.
b. $X \in$ ©P iff $(\forall x \in \mathcal{C}_X)\, x \in P$.

What is usually called a particular "reading" can now be considered a particular choice of a context, i.e., of a cover.

It is not altogether clear to me, however, whether or not the three dominant readings – distributional, cumulative, and group reading – should really be described as involving the kind of context dependence discussed above. For example, the cumulative reading represented by $X \in {}^*P$ seems to involve no specific choice of a cover at all, and the group reading can simply be formalized as $X \in P$. Differentiating these readings might as well be a matter of Logical Form, i.e., whether or not predication should be translated in one way or another. Similarly, distributivity – although it can be described as a subcase of (14) when choosing the trivial cover – might as well be expressed by choosing $^\Delta$ as defined in the last section, as a possible choice within the translation procedure.[12] As context dependence is omnipresent in ambiguity resolution – and one of the tasks of LF is ambiguity resolution – one might argue that the pragmatics involved should be located in the process of translation, rather than transporting it into the semantic representation itself.

On the other hand, this argument ignores the striking family resemblance between the different "readings." As shown above, all construals can be derived from

[12] Note that the choice of $^\Delta$ seems possible even with "non-distributive" predicates and does not yield a contradiction for sentences like *they gathered*. Recall that *they* can co-refer with and therefore denote a set of two groups, e.g. two committees. Since each committee is an element of D, the sentence does have the distributional reading, contrary to precipitate conclusions.

a single semantic representation via restrictions on C, so that only the conceptually most elementary and simple restrictions are called "readings." No such uniformity can be guaranteed under the "define a new operator" approach we exemplified with $^{\Delta}$. Furthermore, expressions like *each* and *together*, which disambiguate these readings, can easily be described *semantically* as natural language terms that express the above described properties of covers. In the alternative approach, where different readings become available by using different operators, it is unclear how to ascribe any identifiable meaning of their own to these expressions. In that approach their contribution to meaning is to select certain LFs and exclude others, but although their *function* is fairly clear, it is unclear how this function can be turned into what is called a *denotation* for these expressions.

Another argument in favor of contextually dependent covers derives from Schwarzschild's thesis that the way one refers to sets by using particular expressions will induce particular covers over that set. Recall that according to the union 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 are analyzed as $X \in {}^{©}(f(met))$, the NP expressions induce different covers over 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 can render one sentence true and the other false.

My tentative conclusion is that not all occurrences of cumulative predication depend on context, so that there is still a motivation for adopting * as a purely semantic operation. On the other hand, we do have to take into account pragmatic considerations and semantic operators that specify properties of covers, which implies that we need both * and © independently and for different purposes. This conclusion will gain substantial support from the analysis of more complicated constructions. As regards the distributive reading, I continue to use $^{\Delta}$ and simply leave it open whether or not it should be reduced to a particular choice of ©.

2.3. Relational Predication

2.3.1. Semantics

Since our main objective is compositionality, there is no way to avoid a discussion of the standard examples from Jackendoff (1972) and Scha (1981) which constitute a major challenge for naïve compositionality. Let us begin with Scha's famous example (15-a) and its paraphrase in (15-b):

(15) a. 500 Dutch firms own 2000 Japanese computers.

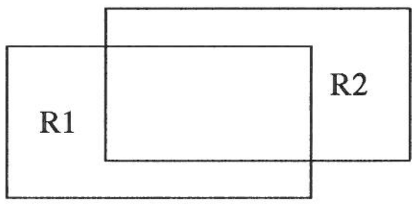
b. There are 500 Dutch firms and there are 2000 Japanese computers, such that for each firm there is a computer it owns, and for each computer, there is a firm which owns it.

The difficulty illustrated by (15) lies in the fact that the six quantifiers used in the logical paraphrase (15-b) constitute the meaning components of only two NPs in (15-a), and appear at various places in the paraphrase, so that their relative scopi penetrate each other.

The same kind of problem arises with Scha's second example (16-a), which describes the situation shown in (16-b) and can be analyzed as in (16-c):

(16) a. The sides of rectangle 1 are parallel to the sides of rectangle 2.

b.



c. $(\forall x \in [\![\text{side of R1}]\!]_{\mathcal{M}})(\exists y \in [\![\text{side of R2}]\!]_{\mathcal{M}})\ (\langle x, y\rangle \in [\![\text{parallel}]\!]_{\mathcal{M}}) \wedge$
$(\forall x \in [\![\text{side of R2}]\!]_{\mathcal{M}})(\exists y \in [\![\text{side of R1}]\!]_{\mathcal{M}})\ (\langle x, y\rangle \in [\![\text{parallel}]\!]_{\mathcal{M}})$

It is obvious that (16-c) cannot be derived from the syntactic structure of (16-a) by compositional methods. No matter how one construes the semantics of the determiner, the paraphrase with four quantifiers will always exhibit a kind of "cross-over effect" so that the quantifiers get in one another's way, excluding a compositional analysis.

Scha's method of dealing with these examples is to state meaning postulates for the verbs. As already noted in footnote 8, meaning postulates are problematic in cases where their applicability depends on the kind of construction at hand and the desired analysis only represents one reading of several possible construals. Thus, a better alternative to formalize (16-a) is (17):

(17) $\langle [\![\text{the sides of R1}]\!]_{\mathcal{M}}, [\![\text{the sides of R2}]\!]_{\mathcal{M}}\rangle \in {}^{**}[\![\text{parallel}]\!]_{\mathcal{M}}$

In order to demonstrate that (17) is adequate, first note that *parallel* is D-reducible. It is easy to see, then, that (17) is logically equivalent to (16-c), and this is sufficient for showing that the analysis is correct.

Turning next to (15-a), an analogous analysis – with "DF" denoting the set of Dutch firms and "JC" denoting the set of Japanese computers – is this:

(18) $(\exists X \in {}^{*}\text{DF})(\exists Y \in {}^{*}\text{JC})\ 500(X) \wedge 2000(Y) \wedge \langle X, Y\rangle \in {}^{**}[\![\text{own}]\!]_{\mathcal{M}}$

(18), however, is not logically equivalent to my tentative paraphrase in (15-b), because the predicate *own* is presumably not D-reducible, i.e., some firms may only jointly own a single computer. Conversely, however, it seems that *own* is nevertheless D-reducible in its second argument position, which means that if a set of computers is owned by some x, then each of its elements is owned by x. On the assumption that *own* is "partially" D-reducible, it can now be seen that (18) is equivalent to (19):

(19) $$(\exists X \in {}^*\mathrm{DF})(\exists Y \in {}^*\mathrm{JC})(\exists \mathcal{C}_X)\, 500(X) \wedge 2000(Y) \wedge$$
$$(\forall x \in \mathcal{C}_X)(\exists y \in Y)\, \langle x, y\rangle \in [\![\, \mathrm{own}\,]\!]_{\mathcal{M}} \wedge$$
$$(\forall y \in Y)(\exists x \in \mathcal{C}_X)\, \langle x, y\rangle \in [\![\, \mathrm{own}\,]\!]_{\mathcal{M}}$$

Intuitively, each element of the cover is a set of firms who share a computer. I use lower case letters here to denote these elements; it should be recalled, however, that all variables denote non-empty elements in *D*. I distinguish between lower and upper case variables for mnemonic purposes only, the general convention being the following: arguments of pluralized nouns and verbs are written as $X, Y, \ldots$, but arguments of non-pluralized nouns and verbs are written as $x, y, \ldots$. It follows that a lower case variable denotes an entity in D only if the argument position it fills is D-reducible. This is always the case with singular CNs, but need not be the case with verbs. Bearing this in mind, it should be easier to decipher (19) and grasp the correctness of these truth conditions.[13]

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 (20-a) (= his example 7.55) has the analysis given in (20-b):

(20) 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.

---

[13] The above analysis is the same as the one in Krifka (1990). However, Krifka (p.c.) pointed out to me that this is not exactly the paraphrase given by Scha, because (18) does not logically imply that 500 and 2000 are *maximal* numbers of firms and computers respectively that make the sentence true. I leave it open, here, whether this aspect of meaning 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. In any case, it seems to me that the problem of compositionality is sharpened, since one has to manipulate two quantifiers "simultaneously." The problem seems even more interesting for the pragmatic approach, since it shows that even here – in an adequate formalization of pragmatics – do we have to take non-compositionality into account.

(20-b) is called the branching reading of (20-a). It is easy to verify that this reading cannot be derived from a pragmatic strengthening of $**P$: there is no way to turn cumulative predication into the non-compositional distributive construal of (20). On the other hand, in order to represent the Jackendoff reading, it is not necessary to define a new operator. In order to demonstrate this, it will be helpful to introduce some pieces of notation.

Henceforth, I will use lambda abstraction as a more convenient way to express set formation. Along with $(\lambda x\varphi)(y)$, I will also write $y \in \lambda x\varphi$, hoping that no confusion of semantic types will arise.[14] The property of X which will hold true of X when X is a large set of men will be abbreviated as $\mathbf{MM}(X)$ (= $\mathbf{many\text{-}men}(X)$). Applying these conventions, the Jackendoff reading can be represented as in (21-a). By taking into account that *story-telling* is D-reducible, and that $*P$ and $^{\Delta}P$ are equivalent for D-reducible predicates, we derive that (21-a) is equivalent to (21-b), which I abbreviate as (21-c).[15]

(21) a. $(\exists X)(\exists Y)\ \mathbf{MM}(X) \wedge \mathbf{3}(Y) \wedge *\mathbf{story}(Y) \wedge X \in *\lambda x.Y \in *\lambda y.\mathbf{tell}(\mathbf{I}, x, y)$

b. $(\exists X)(\exists Y)\ \mathbf{MM}(X) \wedge \mathbf{3}(Y) \wedge *\mathbf{story}(Y) \wedge X \in {}^{\Delta}\lambda x.Y \in {}^{\Delta}\lambda y.\mathbf{tell}(\mathbf{I}, x, y)$

c. $(\exists X)(\exists Y)\ \mathbf{MM}(X) \wedge \mathbf{3}(Y) \wedge *\mathbf{story}(Y) \wedge \langle X, Y\rangle \in {}^{\Delta\Delta}\lambda xy.\mathbf{tell}(\mathbf{I}, x, y)$

Since $^{\Delta}$ can be interpreted as a special case of $^{©}$, we can also substitute $^{©}$ for occurences of $^{\Delta}$ in (21-b), thus generating a blend of different analyses that are irrelevant for the above example. For reasons to be discussed in the next subsection I will define "$^{©©}$" as a genuine two place operator in such a way that $^{©©}$ is *not* equivalent to two iterated applications of $^{©}$. It should therefore be kept in mind that our notational convention will not allow for substituting $^{©}$ for $^{\Delta}$ in (21-c).

Clearly, the two place operator $^{\Delta\Delta}$ can easily be defined directly, in the same way as we defined (CP).[16] But as shown above there is no need to do so, because $^{\Delta\Delta}$ –

[14] In fact, I will use $\lambda x\varphi$ ambiguously, i.e., both in a context like $y \in \lambda x\varphi$ and in the context of functional application $(\lambda x\varphi)(y)$. The motivation comes from word order in natural language and the desire to have a more conspicuous notation for $(x)f$, i.e., for functional application to the left as in subject-verb constructions; cf. section 4.

[15] Henceforth, I will often write α instead of $f(\alpha)$ or $[\![\alpha]\!]_{\mathcal{M}}$ and assume that "α" is a constant of the language of semantic representation, i.e., of predicate logic. These constants will be written bold face.

[16] The following definition does the job:

(i) $^{\Delta_1 \ldots \Delta_n}R := \{\langle X_1, X_2, \ldots, X_n\rangle : X_1 \times X_2 \times \ldots \times X_n \subseteq R \wedge (\forall i)\, X_i \neq \emptyset \wedge X_i \subseteq D\}$

Unfortunately, (i) seems hard to understand, but I hope its intuitive content has been made clear by the above example.

unlike ** – can be reduced to two iterative applications of $^{\Delta}$.[17] This strikes me as an important result concerning compositionality. Compare the intuitive paraphrases of the Jackendoff sentence and the Scha sentence. Intuitively, the Jackendoff paraphrase is less complex than Scha's paraphrase as regards quantifier interaction. This observation is formally reflected in the above analysis: $^{\Delta\Delta}$ can be decomposed into two operators, but ** cannot. This shows that, in a sense, the Scha sentence is less compositional than the Jackendoff example.

2.3.2. Pragmatics

Schwarzschild points out in his dissertation that the Jackendoff-type analysis in terms of $^{\Delta\Delta}$ is often too strong, but the Scha-type analysis in terms of ** can be too weak. Let me quote his discussion (p. 131f) of the following example from Scha and Stallard (1988):

> 193) 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 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.

In order to arrive at such an analysis one has to take into account covers that work on pairs. Thus, before being in a position to state truth conditions we first have to extend the notion of cover, so that it establishes a relation between the two covers that correspond to the two argument places of a predicate. This can be done as follows (p. 127):

> 190) PCov is a paired cover of $\langle A, B \rangle$ if PCov is a set of pairs constructed as follows: take a cover of A and a cover of B and form a set of pairs such that every element of the A cover is a first member and every element of the B cover is a second member of a pair in PCov.

Having defined a cover over a relation as indicated above, the next thing to do is to formulate contextually relativized truth conditions. Schwarzschild's condition (191)

---

[17] I owe this insight to Arnim von Stechow (class lectures), who analyzed scope independent readings simply as scope phenomena that involve interaction between existential quantification and lambda abstraction. It should be clear that this method cannot be used to describe the semantics of **.

says that there must be a pragmatically determined paired-cover of $\langle X, Y \rangle$ such that each element of the paired-cover is in R. Applying the proposal to the example given above, this means that we have to partition the frigates and the carriers according to operation areas and pair them together in such a way that each pair in the paired-cover includes all the ships of one operation area. Schwarzschild now claims that in each operation area the frigates are faster than the carriers.

Unfortunately, however, our precise apparatus developed above will reveal that this analysis, which is reformulated as (22-a), cannot work properly as soon as the predicate involved is C-reducible. But as seems plausible enough, the relation *faster than* is even D-reducible. Given the paired-cover $\mathcal{PC}$ as induced by operation areas, (22-a) fails because it is not the elements of $\mathcal{PC}$ (which are sets) that stand in relation R – as should be the case according to (22-a) – but rather the elements of these elements must be paired by R in a certain way.

At this point of the analysis, there are several ways to proceed; one is to weaken the strong branching analysis; another is to strengthen the weak cumulative analysis. Using obvious notational abbreviations, these analyses are formalized in (22-b) and (22-c), respectively:

(22) a. $\langle A, B \rangle \in {}^{\copyright\copyright}R$ iff there is a pragmatically determined p-cover $\mathcal{PC}$ over $\langle A, B \rangle$ such that $(\forall \langle x, y \rangle \in \mathcal{PC}_{AB})\, \langle x, y \rangle \in R$.

b. $\langle A, B \rangle \in {}^{\copyright * \copyright}R$ iff $(\forall \langle x, y \rangle \in \mathcal{PC}_{AB})\, \langle x, y \rangle \in {}^{**}R$

c. $\langle A, B \rangle \in {}^{\copyright \Delta \copyright}R$ iff $(\forall \langle x, y \rangle \in \mathcal{PC}_{AB})\, \langle x, y \rangle \in {}^{\Delta\Delta}R$

Although Schwarzschild is agnostic about which formalization is the correct one (probably because he didn't notice the problem), it turns out that all of his examples favor the last analysis. As concerns the above scenario, it is clear that $^{**}[\![$ faster-than $]\!]_{\mathcal{M}}$ only requires the fastest ship to be a frigate and the slowest to be carrier, which is compatible with a majority of carriers being faster than most frigates, in some or all operation areas. But clearly, under these circumstances the sentence should be judged false. Turning therefore to the branching analysis in (22-c), we see that $^{\Delta\Delta}$ requires the Cartesian product of the carriers and frigates to be in R, which means that in each operation area, each frigate must be faster than each carrier. Prima facie, then, only (22-c) correctly reflects the intended reading and therefore should count as superior to (22-b). I will show in the next section, however, that there is reason to adopt both (or all) of the proposals in (22), and make the choice between them a matter of pragmatics again.

Summarizing so far, we have seen that (CP) is essential for formulating correct truth conditions. On the other hand, we adopted a competing pragmatic theory of covers. As shown in (22), the generalization of this theory to two place relations involves both p-covers and semantic specifications, at least in (22-b) and (22-c). We will see in the next section that the thesis of having both semantic and pragmatic conditions at the same time will receive support from the analysis of reciprocal constructions.

3. Reciprocals

3.1. Elementary Reciprocal Sentences

3.1.1. Semantics

Imagine a line of elephants in a circus: "Each elephant holds tightly in his trunk the tail of an elephant who lumbers ahead of him" (Dougherty 1974: 17). In this situation one can truthfully utter: *The elephants are holding each other*. Given that the elephants are arranged in a circle, the situation can adequately be described by the following formula:[18]

(23) $(\exists X)\, X = [\![\text{the elephants}]\!]_{\mathcal{M}} \wedge \langle X, X \rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{hold}(x, y)]$

To the best of my knowledge, Langendoen (1978) was the first to notice that there is a tight connection between plural semantics and reciprocity. He first defines "weak reciprocity" (=WR) as in (24-a) and shows that it is semantically adequate in being general enough to handle almost all elementary reciprocal sentences.[19] Next he defines the semantics of what he calls "elementary plural relational sentences," i.e., sentences of the form $A\,R\,B$, where R expresses a relation between the sets A and B. This is stated in (24-b):

(24) a. $(\forall x \in A)(\exists y, z \in A)\, x \neq y \wedge x \neq z \wedge xRy \wedge zRx$ (WR)

b. $(\forall x \in A)(\exists y \in B)\, xRy \wedge (\forall z \in B)(\exists w \in A)\, wRz$ (EPR)

[18] If the elephants are lined up in a row, such that there is a first and a last elephant, the formula presented in the text is false, i.e., it is too strong. This is always the case with predicates that express "asymmetric disconnected relations," as in (i):

(i) The plates are stacked on top of one another.

I will simply ignore relations of this type and refer the reader to Langendoen (1978) for some discussion.

[19] There are only two exceptions to WR: the one mentioned in the last footnote, and a certain generalization of WR, namely WR for sets, to which I will return later in this section.

(24) is familiar from the analysis of the Scha sentences; in fact, for D-reducible predicates, EPR is logically equivalent to $**R(A, B)$.

Now, Langendoen's important insight is that one can "derive" WR from EPR by identifying A with B and adding to (24-b) the non-identity statements $x \neq y$ and $w \neq z$. This connection between WR and EPR directly carries over to our analysis of EPR-sentences in (25-a). However, we can arrive at WR in an even simpler way, namely by adding only one non-identity statement instead of two. Thus, our semantic representation of "the A's R each other" amounts to (25-b):

(25) a. $\langle A, B\rangle \in **\lambda xy[R(x, y)]$

b. $\langle A, A\rangle \in **\lambda xy[x \neq y \wedge R(x, y)]$

(25) reveals that the semantics of the reciprocal reduces to nothing more than the non-identity statement $x \neq y$, plus a suitable anaphoric relationship $A = B$, the latter being encoded as part of the syntax of anaphoric indices to be explained in section 4.4.

As a further illustration, let us discuss (26) from German. Note first that the reflexive form *sich* in German is morphologically neutral between singular and plural. In its plural use, *sich* can be understood as an indication of reflexivity, but it can also be used like English *each other/one another* as an indication of reciprocity. Thus, (26-a) is ambiguous between the reflexive reading expressed in (26-b) and the reciprocal reading in (26-c):

(26) a. Sie betrachteten sich
They looked-at REFL

b. Jeder von ihnen betrachtete sich selbst
Each of them looked-at REFL self

c. Sie betrachteten sich gegenseitig
They looked-at REFL mutual

It seems to me that the ambiguity is not exactly on a par with other lexical ambiguities, perhaps because (26-a) does not enforce a disambiguation into either (26-c) or (26-b). This might point to the existence of a "neutral" construal which simply leaves it open whether or not the action involved is reflexive or reciprocal.[20] This

[20] One might also consider an analysis in terms of pragmatically determined p-covers that disambiguate readings by context dependence. Note, however, that a context might leave it completely open which reading is salient, in which case (27-a) could be judged true even if some of them look at themselves, but others look at one another. The same is true, e.g., for *The women released each other*, in a situation where some women were not released by others, but themselves. Hence, it seems justified to take (27-a) as the neutral construal that does not depend on pragmatic

"neutral" construal will be refferred to as "purely anaphoric." All three variants of (26-a) can be described by means of CP in a simple way. Translating the pronoun *sie* ('they') as a variable X, (27-a) represents the purely anaphoric case, (27-b) adds reflexivity, and (27-c) adds reciprocity, i.e., non-reflexivity:[21]

(27) a. $\langle X, X\rangle \in {}^{**}\lambda xy[\mathbf{look\text{-}at}(x, y)]$

b. $\langle X, X\rangle \in {}^{**}\lambda xy[x = y \wedge \mathbf{look\text{-}at}(x, y)]$

c. $\langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{look\text{-}at}(x, y)]$

Let us discuss these formulas in some detail. Since *betrachten* ('look-at') is D-reducible, we can mechanically transform these formulas into Langendoen's EPR format, with R representing *betrachten*:

(28) a. $(\forall x \in X)(\exists y \in X)\, R(x, y) \wedge (\forall y \in X)(\exists x \in X)\, R(x, y)$

b. $(\forall x \in X)(\exists y \in X)\, x = y \wedge R(x, y) \wedge (\forall y \in X)(\exists x \in X)\, y = x \wedge R(x, y)$

c. $(\forall x \in X)(\exists y \in X)\, x \neq y \wedge R(x, y) \wedge (\forall y \in X)(\exists x \in X)\, y \neq x \wedge R(x, y)$

These formulas can be simplified in the obvious way:

(29) a. $(\forall x \in X)(\exists y, z \in X)\, R(x, y) \wedge R(z, x)$

b. $(\forall x \in X)\, R(x, x)$

c. $(\forall x \in X)(\exists y, z \in X)\, x \neq y \wedge x \neq z \wedge R(x, y) \wedge R(z, x)$

Since (29-c) is the same as (24-a), WR is a special case of EPR.

As already noted by Langendoen, these reductions do not hold for predicates which are, in our terms, not D-reducible. Discussing a model with $D = \{a, b, c\}$ and a relation $f(R) = \{\langle\{a, b\}, c\rangle, \langle c, a\rangle, \langle c, b\rangle\}$, he observes that it seems quite natural to say that the Xs relate to each other. But since (29-c) is false in this situation – regardless of whether $X = D$ or $X = D \cup \{\{a, b\}\}$ – the truth conditions of the EPR-sentence cannot be captured by his formula.

Langendoen's way out, also adopted by Moltmann (1992), is to state two versions of WR: the first is WR as defined above, and the second is a "type shifted"

evaluation of a given situation.

[21] The existence of the purely anaphoric analysis appears to conflict with the Gricean principle adopted above, namely that the less informative reading as encoded in (27-a) is blocked if there are more informative alternatives like (26-b) and (26-c). In order to regain consistency with our previous use of the principle, one would have to say that the expressions *sich* vs. *sich selbst* and *sich* vs. *sich gegenseitig* do not form quantitative scales in the sense of Horn (1972), and hence cannot be compared in the same way as the grammaticalized difference between singular and plural.

generalization of WR, i.e., a formula that does not quantify over individuals but over sets. Bearing in mind, however, that our semantics of ** is equivalent to Langendoen's EPR only if the predicate involved is D-reducible, our analysis might not be flawed by the inadequacy of (29-c) for the case under debate. And indeed, Quine's innovation and our definition of CP allow us to subsume both cases under the same analysis (30): since the small case variables do not range over individuals only, it can easily be verified that (30) is true in the model described above.

(30) $(\exists X)\, X = \{a, b, c\} \wedge \langle X, X\rangle \in {**}\lambda xy[x \neq y \wedge R(x, y)]$

Hence, a more general semantics for pluralization obviates any duplication of reciprocity conditions and therefore corroborates our thesis that there is an intimate connection between reciprocity and plurality.[22]

3.1.2. Pragmatics

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 just the union set of men and 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 dissertation, it has not been checked against examples like the following, which seem to challenge the union theory:

(31) The men and the women hate each other.

Assuming that M and W denote the sets of men and women respectively, our previous analysis suggests the following truth condition:

(32) $(\forall\langle X, Y\rangle \in \mathcal{PC}_{\langle M\cup W}, M \cup W\rangle)\, \langle X, Y\rangle \in {**}\lambda xy[x \neq y \wedge \mathbf{hate}(x, y)]$

In order to evaluate (32), consider first the partition $\{M, W\}$ contextually induced by the subject expression. There are different ways to build a p-cover over $\{M, W\}$. Suppose we group the men with the men and the women with the women (i.e., we use $\mathcal{PC}_{X^{\mathrm{id}}}$). This will yield the distributive reading spelled out in (33): the men hate each other and the women hate each other.

(33) $(\forall X \in \{M, W\})\langle X, X\rangle \in {**}\lambda xy[x \neq y \wedge \mathbf{hate}(x, y)]$

[22] These findings count against isolated treatments of reciprocity like Moltmann's (1992), which tend to make the semantic rules for reciprocal expressions more complicated than necessary, while still being unable to cope with certain empirical problems to which we return in section 3.2.2.

Alternatively, one might pair the men with the women and vice versa. In that case (32) yields:

(34) $(\forall\langle X, Y\rangle \in \{\langle M, W\rangle, \langle W, M\rangle\})\, \langle X, Y\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{hate}(x, y)]$

As the non-identity statement in (34) only says that we may not take hermaphrodites into account, this part of the formula becomes more or less trivial, and (34) simply means that the men hate the women and the women hate the men.

Clearly these truth conditions are empirically correct. Nevertheless they are not satisfying. One problem results from the fact that there are further possibilities for building p-covers, e.g. $\{\langle M, M\rangle, \langle M, W\rangle, \langle W, W\rangle\}$, that generate "mixed readings" but do not seem to correspond to intuitively possible analyses. We observe, however, that all of these additional covers are extensions of the two minimal covers described above. Suppose, then, that as a rule these covers must be ignored. We can now identify the real problem, which emerges only with more than two conjuncts, as in (35):

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

Let $X = \{M, W, C\}$ be the relevant cover induced by the subject expression, with C as the set of children. We now have to choose a p-cover, i.e., a subset of $X \times X$. Choosing $\mathcal{PC}_{X^{\mathrm{id}}}$, i.e., the identity relation $\{\langle M, M\rangle, \langle W, W\rangle, \langle C, C\rangle\}$, yields the distributional construal in the same manner as above. In order to get the strong reciprocal reading (36),

(36) The men hate the women and vice versa; the women hate the children, and vice versa; the men hate the children, and vice versa.

we make the cover itself reciprocal, i.e., we take it to be $(X \times X)\backslash\mathcal{PC}_{X^{\mathrm{id}}}$. Investigating other covers of $(X \times X)$, however, it happens that some of these have to be ruled out. If we allow for all combinations of p-covers, we not only get correct readings, but also generate impossible "mixed readings" again, e.g., an analysis in which the children hate the children, the men hate the women, and the women hate the men. It is also clear that the corresponding cover $\{\langle C, C\rangle, \langle M, W\rangle, \langle W, M\rangle\}$ is minimal and therefore cannot be excluded by the kind of additional consideration we mentioned when dealing with (31). Nonetheless it is evident that such a p-cover does not represent a possible reading. Only "reciprocal" p-covers can be allowed. From this we must infer that our analysis of the reciprocal cannot be correct, because we have missed the semantic role of reciprocity: the intuitively disallowed covers should be ruled out by the syntax and semantics of the reciprocal, rather than by the "pragmatics" of p-covers.

It can readily be seen that the analyses proposed in (32) and (34) are indeed unacceptable LFs. Why? Note that in all previous applications of the theory, an occurrence of the same variable X represented both the subject and the object. This feature of our analysis does not carry over to (32). Nonetheless, it evidently constitutes an essential part of the semantics of the reciprocal, one that encodes the anaphoric nature of the reciprocal. We may express this by saying that its meaning components are $x \neq y$ and $X = Y$. Clearly, the second component is missing in the above analyses. But adding it would only allow us to get the distributive reading. Clearly this is not sufficient for a proper analysis of the different construals of (35).

Let us see how to rescue the situation. The above discussion already suggests that we have to state the reciprocal contribution with respect to sets, i.e., as $X_i \neq X_j$ rather than $x_i \neq x_j$. Thus, the correct analysis must involve a different scoping, as given in (37), where non-reflexivity must hold between pluralities rather than individuals:

(37) $(\exists X)\, X = \{M \cup W \cup C\} \wedge \langle X, X\rangle \in {}^{\circledcirc\circledcirc}\lambda X_i X_j[X_i \neq X_j \wedge {}^{**}\mathbf{hate}(X_i, X_j)]$

Now, it is easy to see that (37) correctly analyzes all possible reciprocal construals, although it excludes the distributive reading, which requires a separate analysis to which I will return immediately. By first applying $^{\circledcirc\circledcirc}$ 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 employs both pragmatic and semantic strategies of pluralization, relying thereby on a crucial double application of pluralization.

Now we still need to derive the distributive construal. Here, we first have to cover the subject by the obvious trivial partition $\mathcal{C} = \{M, W, C\}$. This induces quantification over its elements as shown in the first half of (38) up to $^{\circledcirc}$. In the second half, we interpret the reciprocal construction as usual:

(38) $(\exists X)\, X = \{M \cup W \cup C\} \wedge X \in {}^{\circledcirc}\lambda x[\langle x, x\rangle \in {}^{**}\lambda yz[y \neq z \wedge \mathbf{hate}(y, z)]]$

By thus instantiating $^{\circledcirc}$ we get (39-a), which is paraphrased in (39-b):

(39) a. $(\forall X \in \{M \cup W \cup C\})\, \langle X, X\rangle \in {}^{**}\lambda yz[y \neq z \wedge \mathbf{hate}(y, z)]$

b. The men hate each other, the women hate each other, and the children hate each other.

We conclude, then, that the union theory is in fact capable of accounting for conjoined subjects in the way indicated. However, it is crucial for this theory to be able to correctly translate surface expressions into adequate logical representations.

In particular, we have had to apply pluralization twice. Before elaborating on this aspect of the theory, I would like to discuss some further examples that illustrate the correctness of the proposed semantics for reciprocals.[23]

3.2. Non-Elementary Reciprocal Sentences

In this section I briefly analyze two constructions that have been regarded as puzzles for any theory of reciprocity. In both cases it will be seen that they pose no special problems for our semantic analysis, although they are particularly telling with respect to problems of compositionality and the syntactic component of our theory.

3.2.1. Non-Local Antecedents

(40) is first discussed by Higginbotham (1980) who points out that (40-a) is ambiguous between (40-b) and (40-c):

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

The ambiguity is resolved by assuming that *each other* may have different scope in LF. The narrow scope reading paraphrased in (40-c) is unproblematic and is analyzed in (41):

(41) $$(\exists X)\, X = j \cup m \wedge X \in {}^*\lambda x(\mathbf{think}(x, {}^\wedge[\langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{like}(x, y)]]))$$

Here, the subject of the embedded sentence is translated as the first occurrence of X in $\langle X, X\rangle$. The analysis of (40-b) is even shorter, but we have to move *each other* into the matrix clause (cf. Heim et al. (1991) for a recent discussion):

(42) $$(\exists X)\, X = j \cup m \wedge \langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{think}(x, {}^\wedge[\mathbf{like}(x, y)])]$$

---

[23] Note also that essentially the same analysis goes through for restrictive relative clause constructions like (i) called "hydras" in Link (1984):

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

The partition induced by the NP-head here must cover the semantic value of the relative pronoun rather than that of the argument expression *the men and the women* directly. Apart from this trivial difference to the non-relativized construction (31), there is no special problem involved here.

This time, however, the subject pronoun must translate as a variable x which is co-referential with the external argument of *think*, and since *think* is D-reducible in its subject position, this variable ultimately ranges over elements in D. Let us briefly discuss the nature of this phenomenon, which is called "dependent plural."

Consider first a non-reciprocal sentence like (43):

(43) They love their children.

The logically weakest analysis is given in (44), where the iota-operator is the definite article as described in section 1: it serves to pick out the largest element of a set, so that *their children* is in effect translated as *all their children*:

(44) $(\exists X)\, X = [\![\text{they}]\!]_{\mathcal{M}} \wedge {}^{**}\mathbf{love}(X, \iota(\lambda Y[{}^{**}\mathbf{child}(Y, X)]))$

Inspecting the truth conditions of (44) will reveal, however, that the analysis in terms of two-place cumulative predication does not imply that any person loves any of his own children. A logically stronger analysis would result from replacing ****love** by $^{\Delta\Delta}$**love**, but this time the truth conditions might be considered too strong, because everyone has to love his own and everyone else's children, although the sentence can very well be judged true when everyone loves only his or her own children.[24]

Of course, such a correct analysis is (45-a), which is D-reducible to (45-b):

(45) a. $(\exists X)\, X = [\![\text{they}]\!]_{\mathcal{M}} \wedge X \in {}^{*}\lambda x[\, \iota({}^{*}\lambda y.\mathbf{child}(y, x)) \in {}^{*}\lambda y.\mathbf{love}(x, y)]$

b. $(\forall x \in [\![\text{they}]\!]_{\mathcal{M}})(\forall y)\, [\mathbf{child}(y, x) \rightarrow \mathbf{love}(x, y)]$

A close look at (45-a) will reveal that it is crucial to "translate" *their* as lower case x, not as X. Technically, this does not really mean that there were two different translations of the pronoun. In fact, plural pronouns invariably take elements in $^{*}D$ as denotations. The only relevant difference between x and X in the above formulas is scoping, i.e., whether or not the variable is in the scope of a certain associated CP. If it is, it is bound and may happen to quantify over individuals only, even though the morphological form is plural.

[24] This judgment could be captured by a pragmatic resort to a p-cover over *love*, such that mothers and fathers together are paired with only their own children. But now observe that this p-cover conveys exactly the same information as the relation "child-of." Recalling what we have said in section 3.1.2. about importing too much semantics into the construction of p-covers, we must conclude, however, that such an analysis misses the correct logical representation – one that should give us the intuitively most salient reading of (43) without resorting to any pragmatic accessories.

As another illustration of the phenomenon, consider (46-a) with its two analyses in (46-b) and (46-c):

(46) a. They think their children are ill.

b. $(\exists X)\, X = [\![\text{ they }]\!]_{\mathcal{M}} \wedge X \in {}^*\lambda x[\textbf{think}(x, {}^\wedge[{}^*\textbf{ill}(\iota(\lambda Y.{}^{**}\textbf{child}(Y, X)))])]$

c. $(\exists X)\, X = [\![\text{ they }]\!]_{\mathcal{M}} \wedge X \in {}^*\lambda x[\textbf{think}(x, {}^\wedge[{}^*\textbf{ill}(\iota({}^*\lambda y.\textbf{child}(y, x)))])]$

In (46-b) everyone has a belief about all children. Upon the dependent plural construal in (46-c), however, the sentence can be uttered truthfully even if a) none of them has an opinion about anyone else's children, and b) each of them has no more than one child. According to the pragmatic theory of the singular/plural distinction, this is okay and *their* can so to speak "quantify" over individuals only, as long as there is no expression alternative that conveys the same information by using the third person singular pronoun. But clearly the pronoun cannot, for reasons of agreement, be in the singular, nor can the noun *children*, since the meaning of *they love their child* is completely different.

The formulas in (46) exhibit an ambiguity that will also show up in the analysis of (40-a), repeated as (47-a):

(47) a. John and Mary think they like each other.

b. $(\exists X)\, X = j \cup m \wedge \langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \textbf{think}(x, {}^\wedge[\textbf{like}(x, y)])]$

c. $(\exists X)\, X = j \cup m \wedge X \in {}^*\lambda x[\textbf{think}(x, {}^\wedge[\langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \textbf{like}(x, y)]])]$

The analysis in (47-c) translates *they* as one would expect, namely as the same variable as the one used to interpret the subject of the matrix clause. The reading (47-b), however, is the dependent one, with the second occurrence of *they* being interpreted in the scope of the distribution over the co-indexed upper case variable.

Given our remarks about dependent plurals, all this is straightforward. The more surprising fact to be explained is this: Why can't we get an "independent" analysis of embedded *they* as shown in (48-a) and paraphrased in (48-b):

(48) a. $(\exists X)\, X = j \cup m \wedge \langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \textbf{think}(x, {}^\wedge[X \in {}^*\lambda x.\textbf{like}(x, y)])]$

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

Recall that lower case variables and upper case variables have the same logical type; we must, therefore, distinguish between (48-b) and (47-b) in syntactic terms.

In order to do so, let me introduce some terminology that will also prove useful later. Let us say that in a formula like $\langle X, X\rangle \in {}^{**}\lambda xy[...R(x, y)...]$ the variable x

is *p(lural)-linked* to the first occurrence of X, and the variable y is p-linked to the second occurrence of X. Clearly, the reciprocal pronoun not only translates into $x \neq y$, but an additional requirement must be satisfied: both variables must be p-linked to occurrences of the same variable X – the first occurrence of X being interpreted as the argument filled by the antecedent, the second one being interpreted as the argument representing the anaphor.[25]

An explanation for (48) will emerge only when we compare the derivations from S-structure to the logical representations of (47) and (48). In the former, the reciprocal can be interpreted by (locally) p-linking its variables x and y, but in the latter the reciprocal has been moved across the variable X which might have served as a potential local p-binder. Thus, the configurations to be compared are the following:

(49) a. $X \;\ldots\; x \neq y \;\ldots\; x \;\ldots\; t$

b. $X \;\ldots\; x \neq y \;\ldots\; X \;\ldots\; t$

It seems, then, that movement of the reciprocal cannot skip a potential p-binder; hence, the ultimate reason for the ungrammaticality of (48) is illegitimate cross-over (or perhaps a version of minimality or economy, depending on one's favorite syntactic theory).[26]

Heim, Lasnik and May (1991) refer to the ambiguity displayed in (47) as the "puzzle of scope." They also discuss a second problem, the "puzzle of grain," exemplified by (50), which is ambiguous among the paraphrases listed in (51):

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

(51) a. John told Mary that *he* should leave and Mary told John that *she* should leave.

---

[25] Alternatively, one might require that y is p-linked to a variable Y and add to the translation of the reciprocal the formula $X = Y$.

[26] There is yet another analysis, namely

(i) $(\exists X)\, X = j \cup m \wedge \langle X, X \rangle \in {}^{**}\lambda xy[\mathbf{think}(x, {}^\wedge[x \neq y \wedge \mathbf{like}(x, y)])]$

which does not constitute a possible reading. (i) shows that if the reciprocal moves into a matrix clause, it has to move as a whole; which implies that the two components of the reciprocal meaning can only be separated by pluralization. Although it is straightforward to state the generalization descriptively, I do not have an explanation for it.

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."

I leave it to the reader to verify that these readings easily translate into the following formulas:

(52) a. $(\exists X)\, X = j \cup m \wedge \langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{tell}(x, y, \hat{}[\mathbf{leave}(x)])]$

b. $(\exists X)\, X = j \cup m \wedge \langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{tell}(x, y, \hat{}[\mathbf{leave}(y)])]$

c. $(\exists X)\, X = j \cup m \wedge \langle X, X\rangle \in {}^{**}\lambda xy[x \neq y \wedge \mathbf{tell}(x, y, \hat{}[{}^{*}\mathbf{leave}(X)])]$

The alleged puzzle only arises for the indexing mechanism of GB-theory. As the paraphrases reveal, we must give three different indices to the pronoun *they*. Coindexation with the subject expression (i.e., X) as for (52-c) is straightforward. In the remaining examples we have to distinguish between the logical variable that appears in the argument position of the plural subject (which is x, rather than X) and the variable that appears in the argument position of the plural anaphor, which is y. Binding of *they* is to the former in (52-a) and to the latter in (52-b). Making the required distinction at the level of LF is a simple matter, as we will see in the next section.

What might appear less trivial is the fact that here again the plural pronoun can be bound to what seems to be a variable that ranges over individuals only. But according to what we have said above about dependent plurals, this doesn't really come as a surprise – hence there is no real puzzle to solve here.

3.2.2. Non-Compositionality

Hitherto we have been considering only two place relations and a reciprocal object. The following construction involves a three place relation, and I will show that it constitutes another challenge for strictly compositional semantics.[27]

(53) John read the letters they wrote to each other

Let us agree that relative pronouns are simply lambda abstractors so that relativization is conjunction of properties, i.e., set theoretical intersection of the sets formed by our extended use of the notation of lambda abstraction.[28] Ignoring tense as usual,

[27] I was challenged by Arnim von Stechow (p.c.), who presumably was referred to the problem by Irene Heim (p.c.). Now see v. Stechow 1993: 82f.

[28] In the fragment of LF defined in section 4 there is no intersection but only conjunction of open formulas. One can either add to the rules of LF (which seems unnecessary), or change the

the matrix clause can be analyzed as follows:

(54) $\iota(*[\![\text{ letter }]\!]_M \cap \lambda Y[\![\text{ they write } Y \text{ to each other }]\!]_M) \in *\lambda x[\mathbf{read}(j, x)]$

We still need to analyze the relative clause *what they write to each other*. This can be done straightforwardly in the way suggested above. Translating *they* as X again, we get:

(55) $\lambda Y \langle X, Y, X\rangle \in ***\lambda\langle xyz\rangle[x \neq z \wedge \mathbf{write\text{-}to}(x, y, z)]$

(55) 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 $x \in X$ there are $y \in Y$ and $z \in X$ such that either x receives a letter from z or x sends a letter to z. These seem to be the correct truth conditions. Observe that we arrive at them by simply applying the methods we already used throughout the paper.

In order to see why the sentence was judged so problematic, one has to ask how to formulate adequate truth conditions without following the road we paved. Hence, the relevant question we pursue in this section is whether one could proceed in a more compositional way. Above, we factored out the meaning of the reciprocal into the variable X and the statement that $x \neq z$. Necessarily, x has to be p-linked to X, so that another operator, namely pluralization, intervenes between these meaning components. We will see below how to integrate this into an LF-like representation. The problem now is how to avoid such an analysis, i.e., how to find a way of representing the meaning of the reciprocal in (53) in such a way as to circumvent any splitting up of meaning components.

For a strictly compositional analysis to work properly, one would have to reformulate (55) in such a way that it would be composed of only the familiar operations like lambda abstraction, functional application, pluralization, etc., plus a single meaning component that corresponds to the reciprocal. Intuitively, this must be a kind of operator that applies to a two place relation (e.g. the relation of writing something) by transforming it into a one place reciprocal predicate (i.e., writing something to each other). Given our previous analysis of EPR-sentences, namely $\langle X, X\rangle \in *\lambda xy[x \neq y \wedge R(x, y)]$, it is easy to see that for a two place relation R such an operator would be $\lambda R \lambda X[\langle X, X\rangle \in *\lambda xy[x \neq y \wedge R(x, y)]]$. The present difficulty arises with the assumption that the basic predicate to be modified is three place, rather than two place, with all of its arguments being pluralized.

---

representation in (53); in the present context, however, nothing hinges on such technicalities.

Of course, it would be possible to design a reciprocal operator that takes an n-place predicate as input and reduces its arity by one. Although there is technically no problem with such a definition, three things should be pointed out: first, the meaning of the operator would depend on the arity of the predicate it applies to, i.e., it needs additional contextual information that specifies exactly which argument places of which relations are affected. In other words, we have to define not one, but a family of operators, one for each arity and for each possibility of identifying argument positions.[29] Second, although such an analysis is feasible, the lexical meaning would still be far more complex than our proposal in terms of the non-identity statement $x \neq z$ plus the correct choice of X as anaphoric variable to be co-indexed with the antecedent. And third, it would no longer follow, as it does in our theory, that reciprocal anaphors have a plural antecent. Rather, this semantic fact would have to be integrated into the meaning of the envisaged operator(s), and hence could only be stipulated instead of being derived.

In the remainder of this section I will try to demonstrate that a strictly compositional analysis, i.e., one that operates on a two place relation, cannot work. In order to get an idea of why such an account must fail, we have to look at the two place relations we can generate by lambda abstraction and pluralization and see whether or not we can use these as the argument of a two-place operator whose meaning yields the intended truth conditions. The aim is to show that none of these two place relations can serve as input for a reciprocal operator that can work correctly.

Our first observation is that for C-reducible predicates it is irrelevant whether we consider $\lambda xyR(x, y)$ or $\lambda XY{*}{*}\lambda xyR(x, y)$, since both formulas convey the same information. Thus, it is immaterial whether or not the two place relation to which we want to apply a reciprocal operator has already undergone pluralization. Assume it has not. In consequence, the only *prima facie* candidates to be investigated are $\lambda xz[\textbf{write-to}(x, y, z)]$ and $\lambda xz[Y \in {*}\lambda y[\textbf{write-to}(x, y, z)]]$. A moment's reflection will reveal, however, that neither formula can serve as a basis for reciprocalization. The first one depends on a value assignment for y, which is an individual, since *write-to* is D-reducible in its second argument position. Writing y to each other would mean that one single letter must be sent and received by a group of persons. This does not correctly capture the intuitive truth conditions which are much

[29] Heim (p.c.) proposes defining a family of operators $R^{i,j,n}$ for n-place predicates P, such that the non-identity statement holds between x_i and x_j in $P(x_1, ...x_i, ...x_j, ...x_n)$. The operator $R^{i,j,n}$ is defined as $REFL^{i,j,n}({*}^n({\neq}^{i,j,n}(P)))$, where ${\neq}^{i,j,n}(P)$ extracts from P the subrelation whose i^{th} and j^{th} members are different, and $REFL^{i,j,n}(Q)$ defines a subrelation of Q whose i^{th} and j^{th} members are identical and reduces it to an $(n-1)$-place relation. The meaning of *what they write to each other* can now be represented as $\lambda Y[\langle X, Y\rangle \in R^{1,3,3}(\textbf{write})]$. I am indebted to Irene Heim for having pointed out to me that it is impossible to define reciprocalization as a two place operator.

weaker: if a writes a letter x to b and b writes a letter y to a, then $\{x, y\}$ are the letters they (=$\{a, b\}$) write to each other. This cannot follow from any semantic rule that takes $\lambda xz[\textbf{write-to}(x, y, z)]$ as the basis for reciprocalization.

Although $\lambda xz[Y \in {}^*\lambda y[\textbf{write-to}(x, y, z)]]$ looks more promising, it is easy to see that it cannot do the job; here we encounter a problem for Y. The relation described here depends on Y, but in the model described above, there is only one letter a that x writes to z and another letter b that z writes to x. Unfortunately, Y cannot be the set $\{a, b\}$, since the formula requires that each individual wrote these two letters, instead of only one. It follows rather that in the described circumstances Y is empty, which shows that we are on the wrong track again.

Some further reflection will be necessary in order to grasp that the *prima facie* candidates discussed indeed exhaust the logical possibilities of forming a basis for compositional reciprocalization. But once it is clear that no other choice can reasonably be made, we must conclude that no compositional operation can exist. It seems to me, then, that we have no choice but to analyze reciprocals as a cohesion of two meaning components, as we have done throughout the paper.

3.3. On 'Only One Another'

3.3.1. The Geach-Kaplan Sentence

In this section I will give further support to the thesis that reciprocals have to be decomposed at LF. The corroborating evidence is drawn from the behavior of *each other/one another* when interacting with the operator *only*, as exemplified in

(56) Some critics admire only one another

(56) has become known as the "Geach-Kaplan sentence," which has first been cited in the literature by Quine (1982: 293). Quine's objective is the difference between set theory and second order logic, and in this context the sentence figured prominently in a number of papers; cf. e.g. Lewis 1991, p. 63 and Cartright 1993, p. 203 as the most recent references. In all references I found agreement as regards the logical analysis of (56), which has consistently been represented as (57-a), which is equivalent to (57-b) in our notation:

(57) a. $(\exists x)(\exists y) y \in x \wedge (\forall y)(y \in x \to Cy)) \wedge (\forall y)(\forall z)(y \in x \wedge Ayz \to y \neq z \wedge z \in x))$

b. $(\exists X){}^*\textbf{critics}(X) \wedge (\forall y \in X)(\forall z)(\textbf{admire}(y, z) \to y \neq z \wedge z \in X)$

It will be our task in this section to show that the meaning of *only* when applied to that of the reciprocal will give us (57-b), so that the logicians' analysis is predicted by our theory.

I would first like to point out again that the reciprocal cannot be analyzed as an operator, this result being endorsed in the present context by the fact that it does not make semantic sense to apply *only* to an operator. Instead, we are confronted here with the problem of applying *only* to a morpheme that splits up at LF into two meaning components. Thus, the interesting question that arises is the following. Is the usual meaning analysis of *only* applicable at all and can we resolve the special problem posed by the reciprocal by assuming that *only* applies to both of its meaning components, i.e., to the purely anaphoric variable X_i and the non-identity $x_i \neq x_j$? Our analysis will reveal that this is in fact the case; indeed, there seems no other way to get (57)out of (56). Before giving a formal proof I will briefly demonstrate how this twofold application our analysis relies on can be paraphrased in natural language.

Recall from our discussion of German that we acknowledged a purely anaphoric use of the reflexive pronoun *sich*. Recall also that the decomposition of the reciprocal pronoun can be expressed in natural language by the use of the adverbial *gegenseitig* ('mutually') as shown in (58-b) which is equivalent to (58-a):

(58) a. Die Kritiker bewundern einander
The critics admire one-another

b. Die Kritiker bewundern sich gegenseitig
The critics admire REFL mutually

I take it that the adverbial translates as the non-identity statement (and the purely anaphoric REFL as bound variable). Now, how can we translate the Geach-Kaplan sentence into German, with *nur* as the translation of *only*? As one would expect, (59-a) is fine, but for reasons not intelligible to me, the (59-b) is somehow odd:

(59) a. Die Kritiker bewundern nur einander
The critics admire only one-another

b. Die Kritiker bewundern nur sich gegenseitig
The critics admire only REFL mutually

In fact, a more natural paraphrase of (59-b) would result from rephrasing it as two sentences, namely (60-a) and (60-b):

(60) a. Die Kritiker bewundern nur sich
The critics admire only REFL

b. Die Kritiker bewundern sich nur gegenseitig
The critics admire REFL only mutually

This splitting up into two statements that both involve *nur* is exactly what seems to be going on in our understanding of the Geach-Kaplan sentence. This becomes more apparent as soon as one looks at the right hand side of the material implication in (57-b). Here we find a conjunction of two statements so that (57) becomes equivalent to (61) by conjoining (a) and (b):

(61) $(\exists X)^*\mathbf{critics}(X) \wedge$
a. $(\forall y \in X)(\forall z)(\mathbf{admire}(y, z) \rightarrow z \in X)$
b. $(\forall y \in X)(\forall z \in X)(\mathbf{admire}(y, z) \rightarrow y \neq z)$

In what follows I will show that (61-a) corresponds to (60-a), and (61-b) corresponds to (60-b). Taking these two results together I will show that the conjunction of the above statements is exactly what we get when applying *nur/only* to the reciprocal *einander/one another*. As evidenced by (61) and its German paraphrases this must be done in such a way as to duplicate the operator. Thus, in order to get (57-b) *nur/only* must distribute over both meaning components of the reciprocal, as in the clumsy (and hardly understandable) paraphrase (62):

(62) Die Kritiker bewundern nur sich nur gegenseitig

I see this result as further corroborating evidence in support of our analysis: only if we choose meaning components correctly is it likely, or even possible, to get the correct result when combining them with operators.

3.3.2. A Formal Analysis

As a starting point for the announced proof I describe the meaning on *only* by analyzing a simple sentence like (63):

(63) Some critics admire only John and Mary

Following the standard analysis of *only* I assume that *only* is at LF a two place operator, as shown in (64):

(64) $(\exists X)^*\mathbf{critics}(X) \wedge \mathrm{ONLY}(\{j, m\}, \lambda Y.^{**}\mathbf{admire}(X, Y))$

The semantics of ONLY (ignoring presupposition) is given in (65):

(65) $\mathrm{ONLY}(x, \lambda y\varphi)$ iff $(\forall z)(\lambda y\varphi(z) \rightarrow z \subseteq x)$

Applying this meaning postulate to (64) yields (66):

(66) $(\exists X)^*\mathbf{critics}(X) \wedge (\forall Z)(^{**}\mathbf{admire}(X, Z) \rightarrow Z \subseteq \{j, m\})$

Assuming that *admire* is D-reducible, this is logically equivalent to (67):

(67) $(\exists X)^*\mathbf{critics}(X) \wedge (\forall x \in X)(\forall z)(\mathbf{admire}(x, z) \rightarrow z \in \{j, m\})$

This sentence asserts that if some z is admired, z is either John or Mary. Apart from presupposition, this is obviously the correct truth condition for (63).

By way of analogy and given this equivalence between (64) and (67), it is easy to see that (68-a)translates as (68-b), which in turn is equivalent to (68-c). But this last sentence is in fact (61-a):

(68) a. Einige Kritiker bewundern nur sich (cf. (60-a))

b. $(\exists X)^*\mathbf{critics}(X) \wedge \mathrm{ONLY}(X, \lambda Y.^{**}\mathbf{admire}(X, Y))$

c. $(\exists X)^*\mathbf{critics}(X) \wedge (\forall x \in X)(\forall z)(\mathbf{admire}(x, z) \rightarrow z \in X)$ (cf. (61-a))

To complete the analysis we only have to add the presupposition of (68-a), namely that each of the critics is in fact admired by at least one of them. Adding the presupposition is not essential at this point of the analysis, but I will come back to the issue further below.

Above we applied ONLY to only the first "segment" of the reciprocal. The more interesting task is to apply ONLY to the second "segment," namely $x \neq y$, as shown in the formula corresponding to (60-b):

(69) $(\exists X)^*\mathbf{critics}(X) \wedge \langle X, X\rangle \in {}^{**}\lambda xy[\mathrm{ONLY}(x \neq y, \lambda p[p \wedge \mathbf{admire}(x, y)])]$

Applying meaning postulate (65) to (69) yields:

(70) $(\exists X)^*\mathbf{critics}(X) \wedge \langle X, X\rangle \in {}^{**}\lambda xy[(\forall p)((p \wedge \mathbf{admire}(x, y)) \rightarrow$
$p \subseteq x \neq y)]$

Now observe that $p \subseteq x \neq y$ simply means $p \rightarrow x \neq y$. This follows from the interpretation of truth values as numbers zero and one, which in turn can be represented as sets, so that zero is the empty set and $n + 1$ is $\{n, \{n\}\}$. We now get (71):

(71) $(\exists X)^*\mathbf{critics}(X) \wedge \langle X, X\rangle \in {}^{**}\lambda xy[(\forall p)((p \wedge \mathbf{admire}(x, y)) \rightarrow$
$[p \rightarrow x \neq y])]$

But any formula $(p \wedge A) \rightarrow (p \rightarrow B)$ is logically equivalent to $(A \rightarrow B)$, for any p. Hence, (71) simplifies to (72):

(72) $(\exists X)^*\mathbf{critics}(X) \wedge \langle X, X\rangle \in {}^{**}\lambda xy[\mathbf{admire}(x, y) \rightarrow x \neq y]$

Using again the fact that *admire* is D-reducible, we know that (72) is logically equivalent to (73):

(73) $(\exists X)^*\mathbf{critics}(X) \wedge (\forall x \in X)(\exists y \in X)[\mathbf{admire}(x, y) \to x \neq y] \wedge (\forall y \in X)(\exists x \in X)\ [\text{admire}(x, y) \to x \neq y]$

But a moment's reflection will reveal that that these truth conditions are much too weak. We do get what we want, however, by switching from weak cumulative pluralization to strong distributive pluralization. Thus, starting with $^{\Delta\Delta}$ instead of ** in (69) yields (74-a), which is in turn equivalent to (74-b) and (74-c):

(74) a. $(\exists X)^*\mathbf{critics}(X) \wedge \langle X, X\rangle \in {}^{\Delta\Delta}\lambda xy[\mathbf{admire}(x, y) \to x \neq y]$

b. $(\exists X)^*\mathbf{critics}(X) \wedge X \in {}^*\lambda x.X \in {}^*\lambda y[\mathbf{admire}(x, y) \to x \neq y]$

c. $(\exists X)^*\mathbf{critics}(X) \wedge (\forall x \in X)(\forall y \in X)[\mathbf{admire}(x, y) \to x \neq y]$

Ignoring presupposition again it is clear that (74) correctly captures the meaning of (60-b) and is in fact equivalent with (61-b).

It now remains to show that we get the correct result when combining the two applications of ONLY in one formula. Let us start with (75-a). After a number of transformations whose validity has been proven in the above paragraphs we end up with (75-f):

(75) a. $(\exists X)^*\mathbf{critics}(X) \wedge \text{ONLY}(X, \lambda Y.\langle X, Y\rangle \in {}^{\Delta\Delta}\lambda xy[\text{ONLY}(x \neq y, \lambda p\ [p \wedge \mathbf{admire}(x, y)])])$

b. $(\exists X)^*\mathbf{critics}(X) \wedge \text{ONLY}(X, \lambda Y.\langle X, Y\rangle \in {}^{\Delta\Delta}\lambda xy[\mathbf{admire}(x, y) \to x \neq y])$

c. $(\exists X)^*\mathbf{critics}(X) \wedge (\forall Z)(\lambda Y[\langle X, Y\rangle \in {}^{\Delta\Delta}\lambda xy[\mathbf{admire}(x, y) \to x \neq y]](Z) \to Z \subseteq X)$

d. $(\exists X)^*\mathbf{critics}(X) \wedge (\forall Z)(\langle X, Z\rangle \in {}^{\Delta\Delta}\lambda xy[\mathbf{admire}(x, y) \to x \neq y] \to Z \subseteq X)$

e. $(\exists X)^*\mathbf{critics}(X) \wedge (\forall Z)((\forall x \in X)(\forall y \in Z)[\mathbf{admire}(x, y) \to x \neq y] \to Z \subseteq X)$

f. $(\exists X)^*\mathbf{critics}(X) \wedge (\forall x \in X)(\forall y)[\mathbf{admire}(x, y) \to x \neq y] \to y \in X$

But (75-f) is not yet what we intended to get, because (75) is perfectly compatible with a situation where each critic admires only himself. In order to get the meaning of the Geach-Kaplan sentence come out right it seems necessary to formally analyze its presupposition, namely that no critic admires himself. The presupposition of ONLY$(x, \lambda y\varphi)$ is simply $\lambda y\varphi(x)$, and applying the rule schematically to (76-a) (= (75-b)) yields (76-b), which in turn transforms successively into the remaining clauses of (76), up to (76-e):[30]

(76) a. $(\exists X)$***critics**$(X) \wedge \text{ONLY}(X, \lambda Y.\langle X, Y\rangle \in {}^{\Delta\Delta}\lambda xy[\textbf{admire}(x, y)$
$\rightarrow x \neq y])$

b. $(\exists X)$***critics**$(X) \wedge [(\forall Z)(\lambda Y.\langle X, Y\rangle \in {}^{\Delta\Delta}\lambda xy[\textbf{admire}(x, y)$
$\rightarrow x \neq y](Z) \rightarrow Z \subseteq X)] \wedge \langle X, X\rangle \in {}^{\Delta\Delta}\lambda xy[\textbf{admire}(x, y) \rightarrow x \neq y]$

c. $(\exists X)$***critics**$(X) \wedge [(\forall x \in X)(\forall y)[\textbf{admire}(x, y) \rightarrow x \neq y] \rightarrow y \in X]$
$\wedge \langle X, X\rangle \in {}^{\Delta\Delta}\lambda xy[\textbf{admire}(x, y) \rightarrow x \neq y]$

d. $(\exists X)$***critics**$(X) \wedge [(\forall x \in X)(\forall y)[\textbf{admire}(x, y) \rightarrow x \neq y] \rightarrow y \in X]$
$\wedge (\forall x, y \in X)[\textbf{admire}(x, y) \rightarrow x \neq y]$

e. $(\exists X)$***critics**$(X) \wedge (\forall x \in X)(\forall y)[\textbf{admire}(x, y) \rightarrow [x \neq y \wedge y \in X]]$

But this last formula is exactly what we started with when analyzing (56) as (57-b). We thus have shown that the analysis of (56) must be (75-a) which, as we have seen above, reduces to (76-e). This completes our formal analysis and the proof that the meaning components of the reciprocal must be analyzed in exactly the way we did.

4. Augmented Logical Form

In the preceding sections I have advanced the hypothesis that, given a certain anaphoric context, a reciprocal expression simply translates as an argument X_i plus a non-identity statement $x_i \neq y_j$. I will now try to integrate these findings into a general theory of LF. This theory differs from current theories in that it makes the logical interpretation of syntactic representations more transparent. The additional transparency will be achieved by adding logical machinery to syntactic structures – usually conceived of as LF – which specifies how the structure is to be interpreted semantically. This is the theory of Augmented Logical Form, called ALF. Before giving a brief outline of ALF I will make some general remarks on the relationship between syntax and LF.

4.1. Semantic Glue

Current theories of the syntax-semantics interface differ a great deal in regard to the modes of semantic combination that are permissible as counterparts to syntactic

[30] Applying the rule to the second occurrence of ONLY in (75-a) does not give a meaningful result, because in that case the presupposition would be logically stronger than what is asserted. I must assume, therefore, that such a presupposition is canceled, but have no real explanation for this process.

rules or structures. Oversimplifying a bit, categorial grammar is strictly compositional as it acknowledges only functional application and type shifting as ways to combine meanings. To illustrate, noun modifiers such as relative clauses must be functions from properties to properties, so that the semantic effect of intersection we assumed as the interpretation of relativization must be incorporated into meaning analysis of a single *word*, namely the relative pronoun, whose semantics will become quite complicated.

This approach fails, however, as soon as there is no word into which to pack the relevant semantic information. Partee (1984) in her essay on Frege's principle, compares the following constructions:

(77) a. Being a master of disguise, Bill would fool anyone.
b. Wearing his new outfit, Bill would fool anyone.

It is obvious that different semantic relations hold between the main constituents: in (77-a), the participial clause receives a causal, factual interpretation ("Because he is a master of disguise, Bill would fool anyone"), but in (77-b) it receives a non-factual, conditional interpretation ("If he wore his new outfit, Bill would fool anyone"). But there is arguably no *syntactic* difference between the two constructions, nor is it possible to attach the difference in meaning to a specific lexeme.

On the assumption that there are syntactic constructions that exhibit semantic ambiguity, one can, however, no longer assert that the meaning of the whole depends on syntactic structure and the meaning of its parts alone. What Partee's interpretation of the examples therefore amounts to is a refutation of Frege's principle of compositionality in one of its stricter variants.

A weaker formulation of the principle, however, would admit, besides syntax and word meaning, a third component as a meaning determinant. Besides functional application and conjunction, Universal Grammar provides us with a narrowly defined repertoire of semantic linking elements which need not be expressed by particular morphemes, nor need there be a one-to-one correspondence between these "semantic morphemes" and particular syntactic constructions. In general, however, syntax and semantics will conspire and provide sufficient clues so as to disambiguate and recover the semantic glue that has to be posited in order to make sense out of a string of words.

Part of the evidence in favor of semantic glue has been gained from Scha's and Jackendoff's examples. Given that a complex operator like ** cannot be decomposed into two NP-meanings or parts thereof, and granted that ** cannot be encoded by verbal morphology, strict compositionality has proven untenable. In the augmented conception of LF, then, I describe the conditions that underlie the insertion of semantic material that is not overtly present in surface syntax. In the following sections I will briefly sketch some elements of a general theory of ALF along these lines.

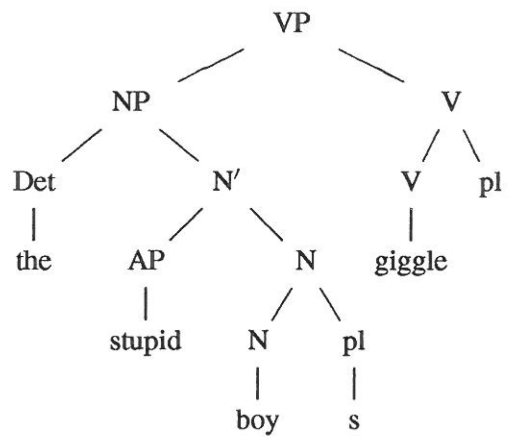
4.2. Elementary Predication

In this section I will try to set up an ALF of the following sentence:

(78) The stupid boys giggle.

I ignore functional projections and assume that the morpho-syntactic analysis basically yields something like this:

(79)

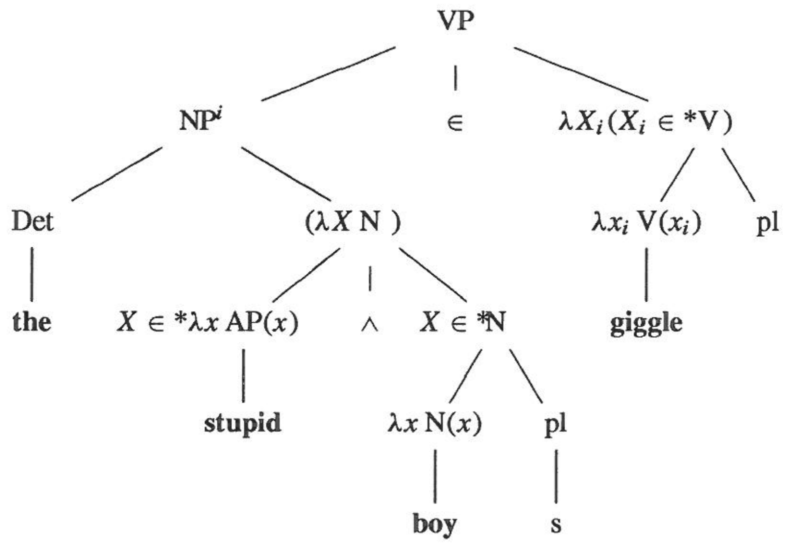


The previous discussion has shown that the meaning of (79) can be represented as

(80) $\iota(\lambda X[X \in {}^*\lambda x.\mathbf{stupid}(x) \wedge X \in {}^*\lambda x.\mathbf{boy}(x)]) \in \lambda X_i[X_i \in {}^*\lambda x_i.\mathbf{giggle}(x_i)]$

It is obvious that the logical formula (80) can be read off directly from an ALF like (81):

(81)



The present task, then, is to specify the rules that transform (79) into (81). Before doing so I would like to point out again that in principle all rule application is optional. However, at some points in a derivation, certain lexical categories may require that the rules have generated certain corresponding semantic categories. For example, it is plausible to say that all IPs should correspond to propositions. Such a statement must then be interpreted as a filter, i.e., a well-formedness condition on ALFs. I will not go into the details of these matching rules, however.

Let me begin with lexical items like verbs and nouns. The first rule will transform these into open formulas (OF):

(OF) If α is an n-place predicate expression of category X, then $\mathrm{X}(x_{i_1}, x_{i_2}, ..., x_{i_n})$ is a formula headed by X.

Our general convention is that the order of arguments $x_{i_1}, x_{i_2}, ..., x_{i_n}$ reflects the hierarchy of positions, so that the first variable represents the subject, the second the object, etc. Note that the resulting open formulas are not yet well-formed in the sense of θ-theory, since the variables introduced by (OF) are not linked to arguments in a "chain," i.e., they still have to be bound.

The next rule of UG is lambda abstraction or set formation:

(LA) If α is a formula headed by X and $x_{i_1}, x_{i_2}, ..., x_{i_n}$ occur free in α, then $\lambda x_{i_1} x_{i_2} ... x_{i_n} \alpha$ is an n-place relation headed by X.

The rule will be used frequently in order to generate n-place relations. As can be seen in the next rule, pluralization presupposes lambda abstraction of exactly those argument places that should be pluralized. Above in this section all predicates were one place, and since $*P = *\lambda x.P(x)$, the operation is somewhat trivial in that case.

The general rule for pluralization (PL) is more complicated, as it has to generalize over n-place relations. Moreover, it applies optionally and freely, i.e., it does not depend on plural morphology. However, we do not want to overgenerate by giving a sentence like *The boy giggles* the same interpretation as *The boys giggle.* We achieve this by preventing (PL) from applying to external arguments of nouns. Noun pluralization will be formulated further below as a separate rule (MorphPL) that is contingent on plural morphology. Thus, the general LF-rule of pluralization is this:

(PL) If $\lambda x_{i_1} x_{i_2} \ldots x_{i_n} \alpha$ is an n-place relation headed by Y and no x_{i_j} is an external argument of a noun then $\langle X_{i_1} X_{i_2} \ldots X_{i_n} \rangle \in {}^{*_1 \cdots *_n} \lambda x_{i_1} x_{i_2} \ldots x_{i_n} \alpha$ is a formula headed by Y (and similarly for ${}^{\Delta}$, ${}^{©}$ and other plural operators mentioned above).

It is important to note that the indices of the upper case variables are identical to those of the lower case variables. Although this might not be logically necessary, it helps to keep track of the relation between argument *expressions* and the occurrence of the argument *positions* introduced by (OF). Thus, e.g., for each variable that is free before applying (PL) there is a corresponding (co-indexed) variable that is free in the resulting formula, so that (PL) does not really affect matters of θ-theory.[31] We will say that by applying (PL) the variable x_j becomes *p-linked* to the corresponding "plural"-variable X_j.

The above rules all turn a single node into yet another node of ALF whose logical interpretation is straightforward: given that the lexemes are interpreted as constants of our formal language $\mathcal{L}$ of semantic representation (e.g. of second order predicate logic plus symbols like $\in$ and operators like lambda abstraction, ${}^{©}$, or $*$), the meaning of a derived expression $\alpha X \beta$ of ALF is computed by concatenating α with the formula of $\mathcal{L}$ dominated by the syntactic category X, concatenated with β.

Basically the same goes for branchings in an LF tree (a recursive definition is straightforward and left to the reader). The following rules define ALFs for branching categories:

(TDI) A branching structure $[_X\ \alpha\ \beta]$ can be turned into

a. $[_X\ \alpha(\beta)]$ (functional application),

[31] It is immaterial, however, whether or not we use the same letters "x" and "X"; we have done so only for mnemonic purposes and could have written "$\langle y_{i_1} y_{i_2} \ldots y_{i_n} \rangle$" in place of "$\langle X_{i_1} X_{i_2} \ldots X_{i_n} \rangle$." Recall also that all variables range over elements in $*D$. Although we will refer to X's as plural variables, this does not mean that their interpretation differs from that of lower case variables. It only means that they are generated outside the scope of (PL).

b. $[_X\, \alpha \in \beta]$ (predication, i.e., functional application in reversed order), or

c. $[_X\, \alpha \wedge \beta]$ (conjunction).

Various syntactic and semantic conditions apply.

Among the obvious semantic conditions is categorial fit, i.e., only formulas can be conjoined, and only functors may apply to arguments. I have therefore dubbed the rule "Type Driven Interpretation" (TDI). Another condition is basically derived from the need to identify certain syntactic positions with certain occurrences of variables. When applying (TDI-b) to a formula built up by lambda abstraction, it is important to properly identify the syntactic position of the argument as subject or object. I will use superscripts for this purpose. The following condition explains the use of indexing for the purpose of θ-theory:

(C1) If α in (TDI-b) is headed by an NP^i, for some index i, then β is of the form $\lambda x_i \varphi$ (and similarly for β in (TDI-a)).

In other words, we require that the index of the NP is identical to the index of the variable that it binds. Thus, the distribution of indices will be as one would expect, so that the index of a variable generated by (OF) will by convention first be associated with a position in the tree, then it will be attached to the argument expression in that position in D-structure, and finally it will be carried along via movement on the route to S-structure and LF. In order to be quite clear about the use of indices, I have decided to use superscripts instead of subscripts, because I do not want to make any claims here about the referential indices of an NP. The above notation merely helps to link syntactic structure with the argument structure introduced by (OF).[32]

[32] Of course, θ-theory and binding theory have to add the usual conditions we will not mention here; and since the above rules are largely unrestricted, they are likely to overgenerate. We will ignore these problems here, noting that the above sketch is only the beginnings of a more elaborate theory we hope to articulate in future work.

On the other hand, the proposal is specific enough to make certain predictions that might be felt incompatible with assumptions concerning type driven QR and much work related to the distinctions between θ-positions and A-positions. One consequence is that there is no need to semantically interprete traces of movement as logical variables. Another one is that there cannot be type driven QR, since quantifiers never occupy the argument slot of the logical variables, whose position does not even appear in syntactic structure. This seems to have welcome implications for the position of negation in languages like German, where arguments always precede sentential negation. Another consequence is that the above mentioned distinctions cannot be semantically motivated; in effect, the notions involved (like θ-position, A-position, etc.) must be purely syntactic, without relying on semantics. Since I have never been convinced of semantic explanations within syntactic arguments for or against a particular underlying syntactic structure, these consequences

Turning now to morphological marking, we simply ignore plural morphology in the same way as we do singular morphology, except for the case of nominal plural morphemes. These morphemes must induce a plural interpretation, otherwise morphology would play no semantic role whatsoever, which is absurd. On the other hand, we explicitly excluded external arguments of nouns from the general rule (PL). It seems necessary, then, to state a second PL-rule that is triggered by the presence of nominal plural morphemes.

(MorphPl) A branching structure $[_{\mathrm{N}}\, \alpha \;\; \mathrm{pl}]$ can be turned into $\langle X_{i_1} X_{i_2} \ldots X_{i_n} \rangle \in {}^{*_1 \cdots *_n}\alpha$ if α is headed by N and is of the form $\lambda x_{i_1} x_{i_2} \ldots x_{i_n} \alpha'$, where x_{i_1} is the external argument of N (and similarly for $^{\Delta}$, $^{©}$ and other plural operators mentioned above).

Although (MorphPL) yields the correct truth conditions, the trouble with this rule is that like all other rules it aplies only optionally. In order to enforce its application, we need a further well-formedness constraint that ensures that pl morphemes are interpreted only if they have to be. This can be expressed as follows:

(C2) Branching structures of the form $[_{\alpha}\; \alpha \;\mathrm{pl}]$ or $[_{\alpha}\; \mathrm{pl}\; \alpha]$ are well-formed only if α is not headed by N.[33]

This reads that pl may have no immediate semantic effect on the interpretation of a branching structure only if the pl-morpheme is not nominal.

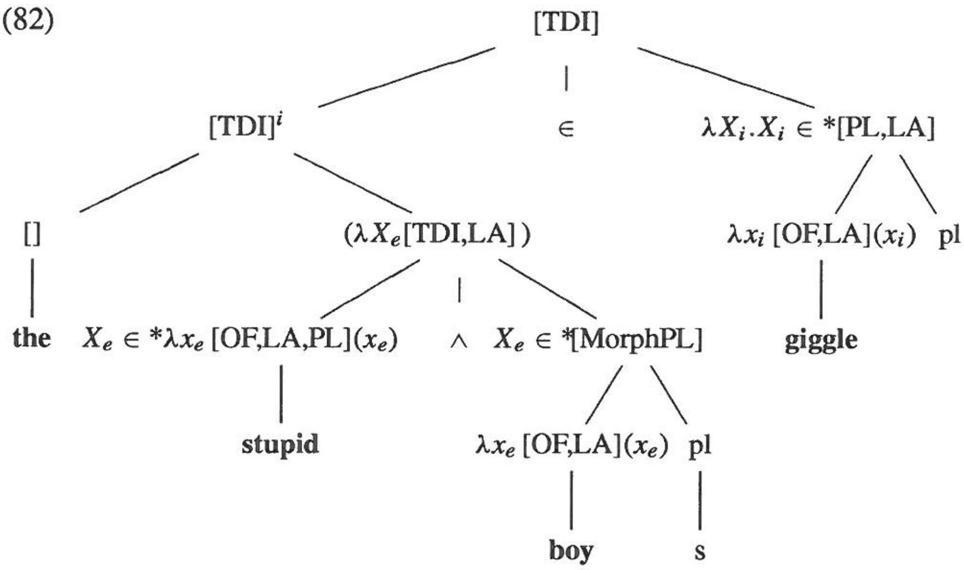
We are now in a position to derive (81) from (79) by applying the above rules. The derivation is shown in (82), and rule application has been encoded by replacing the category labels with the names of the rules that were used to generate the logical material attached to the right and left hand sides of the syntactic categories:

---

seem to be in accord with the idea of autonomy of syntax.

[33] The plural morpheme might occur on a left branch when it is an element of AGR/INFL as the left sister of VP. This configuration might arise if all heads in ALF have to move back to their D-structural position, so that AGR/INFL is no more attached to the right of the verb.

(82)



The above tree reveals that the restriction against pluralization of noun variables cannot be formulated as a restriction on certain indices of variables. This follows since the variables indexed by e in the above example occur both as arguments of nouns and as arguments of the adjective. In the first case we need plural morphology, but in the second we don't. This means, that in the first case, the general rule (PL) cannot apply, but in the second it must. Hence, the restriction embodied in (PL) must be understood as applying to certain occurrences of variables, rather than to types.[34]

4.3. Simple Quantifier Interaction

4.3.1. The Representation of Quantifier Scope

In this section we will apply the theory of ALF to plural quantification. Our aim here is to analyze (83), which can be given a considerable number of logical translations. Only a few of them are shown in (84):

(83) Five men lifted two pianos

(84) a. $(\exists X)\, \mathbf{5}(X) \wedge {}^*\mathbf{man}(X) \wedge (\exists Y)\, \mathbf{2}(Y) \wedge {}^*\mathbf{piano}(Y) \wedge \mathbf{lift}(X, Y)$

[34] Alternatively, one can introduce additional machinery so as to make it a distinction between types, e.g. by assuming that the variables introduced by (OF) range over D only if they are arguments of nouns. We would then have to distinguish between different types of variables, depending on whether they range over D or $^*D$. Up to now, however, this distinction was unnecessary, since the range of variables is automatically reduced to elements in D by the restriction of quantifiers, i.e., by the property they live on. Hence it seems more advantageous to understand the restriction as indicated, i.e., as a condition on occurrences of variables.

b. $(\exists X)\,\mathbf{5}(X) \wedge {}^*\mathbf{man}(X) \wedge X \in {}^*\lambda x[(\exists Y)\,\mathbf{2}(Y) \wedge {}^*\mathbf{piano}(Y) \wedge \mathbf{lift}(x, Y)]$

c. $(\exists X)\,\mathbf{5}(X) \wedge {}^*\mathbf{man}(X) \wedge (\exists Y)\,\mathbf{2}(Y) \wedge {}^*\mathbf{piano}(Y) \wedge Y \in {}^*\lambda y[\mathbf{lift}(X, y)]$

d. $(\exists Y)\,\mathbf{2}(Y) \wedge {}^*\mathbf{piano}(Y) \wedge Y \in {}^*\lambda y[(\exists X)\,\mathbf{5}(X) \wedge {}^*\mathbf{man}(X) \wedge \mathbf{lift}(X, y)]$

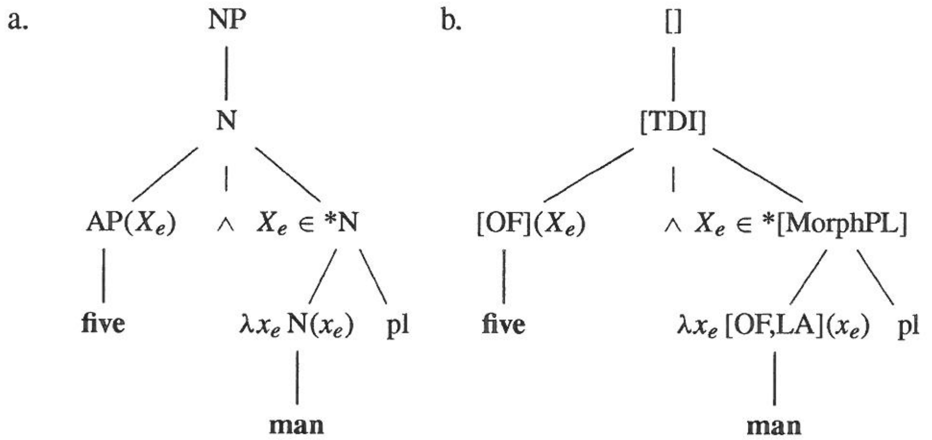
e. $(\exists Y)\,\mathbf{2}(Y) \wedge {}^*\mathbf{piano}(Y) \wedge (\exists X)\,\mathbf{5}(X) \wedge {}^*\mathbf{man}(X) \wedge X \in {}^*\lambda x[\mathbf{lift}(x, Y)]$

f. $(\exists X)\,\mathbf{5}(X) \wedge {}^*\mathbf{man}(X) \wedge (\exists Y)\,\mathbf{2}(Y) \wedge {}^*\mathbf{piano}(Y) \wedge X \in {}^*\lambda x[Y \in {}^*\lambda y[\mathbf{lift}(x, y)]]$

The first one expresses a purely collective reading, where five men jointly lift two pianos, one piled on the other. Since one piano is already heavy enough, this is hard to imagine, but (84-b) is still 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 the other. This and the following analyses, except for (84-f), mix the group reading with the cumulative reading. The analysis in (84-c) 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. (An analysis that enforces this interpretation must, however, be formulated in terms of $^\Delta$ instead of *.) (84-d) reverts the scope of the quantifiers. In this analysis it is possible that each of the pianos has been lifted by five men, which results in a maximal number of ten men. (84-e) is similar to (84-b), but scopeless again. This time the number of pianos cannot depend on the men, hence the men will lift only two pianos at a time. (84-f) is a purely cumulative analysis and hard to express. It means that we can find subsets of men and subsets of pianos such that each subset of men jointly lifted each subset of pianos.

All these logical representations correspond straightforwardly to certain ALFs. It is easily seen that an indefinite NP like *five men* first decomposes into the formula "$\mathbf{five}(X) \wedge X \in {}^*\lambda x[\mathbf{man}(x)]$" in (85-a) with the rules applied as shown in (85-b):

(85)



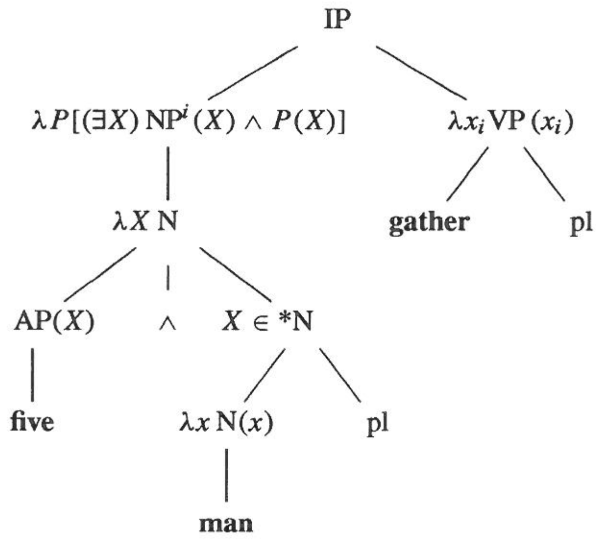
What remains to be done is to add existential quantification. The conventional way to procede is assuming an empty determiner whose meaning is $\lambda Q \lambda P[(\exists x) P(x) \wedge Q(x)]$. There is evidence, however, that such a theory is "too local." Probably, the binding of variables by existential quantification is to be supplemented by various alternative mechanisms, e.g. that of unselective binding by adverbs of quantification as in Lewis (1975), or Heim's (1982) "non-local" binding of free variables in her treatment of donkey anaphora. Binding by an existential quantifier is optional in these theories; basically, existential quantification is inserted wherever it is needed. A more restrictive way to formulate such a theory is to state a rule like this:

(ExQ) Whenever α is a property headed by NP, $\lambda P[(\exists x)\alpha(x) \wedge P(x)]$ is a well-formed ALF.

Note that this rule is optional; if it does not apply some other rules, e.g. rules formulated in the spirit of Heim (1982) or Abusch (1993), must take care of the free variable.

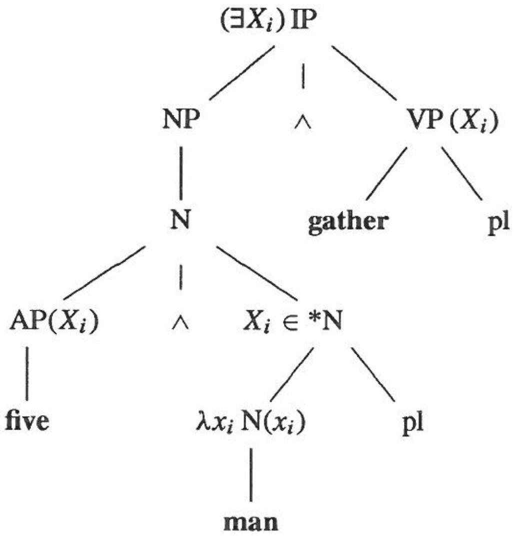
Applying (ExQ) to *Five men gathered* yields (86):

(86)



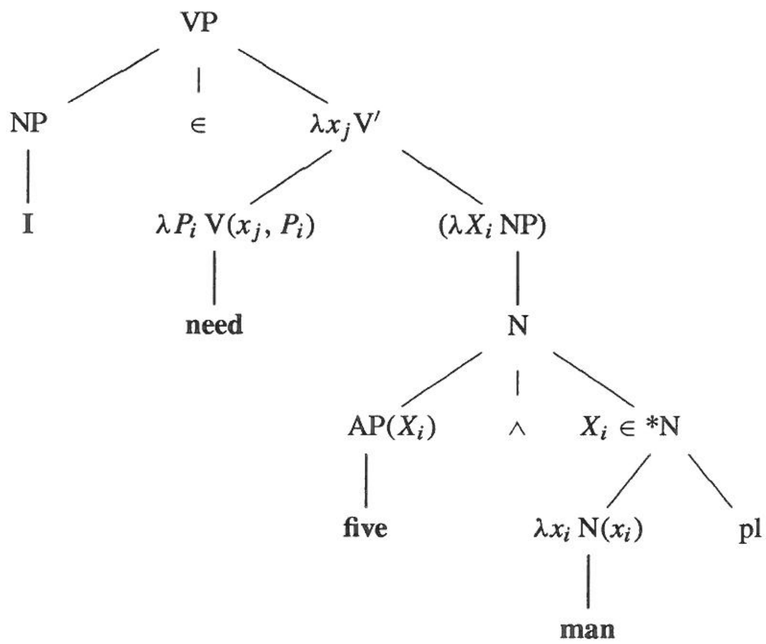
To increase readability, I will do some lambda abstraction and re-indexing already at ALF, so that (86) is transformed into (87) as a more convenient representation of existential quantification:

(87)



Note also that the above theory of existential quantification is neutral with respect to the treatment of intensional contexts. If for example, *need* takes a property as semantic argument, as recently discussed by Zimmermann (1993), the theory straightforwardly yields (88) as a well-formed ALF:

(88)

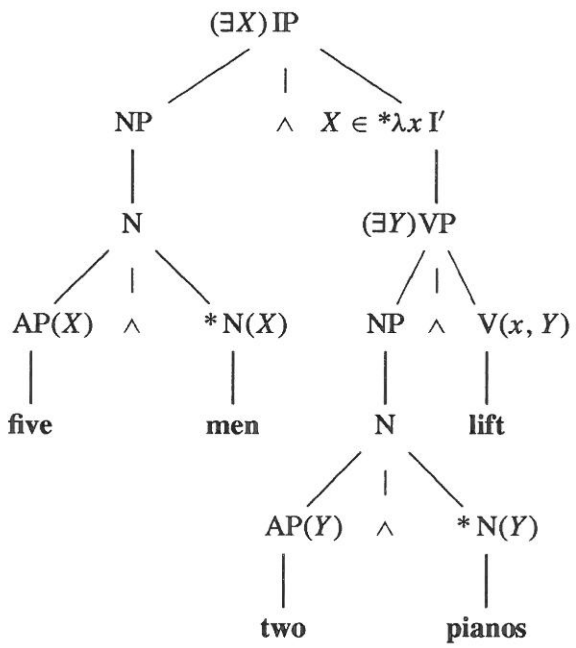


If, however, the semantic type of the complement must be a generalized quantifier, we have to resort to (ExQ), which may optionally apply to the object NP in order to get the kind of representation one is familiar with from Montague grammar.

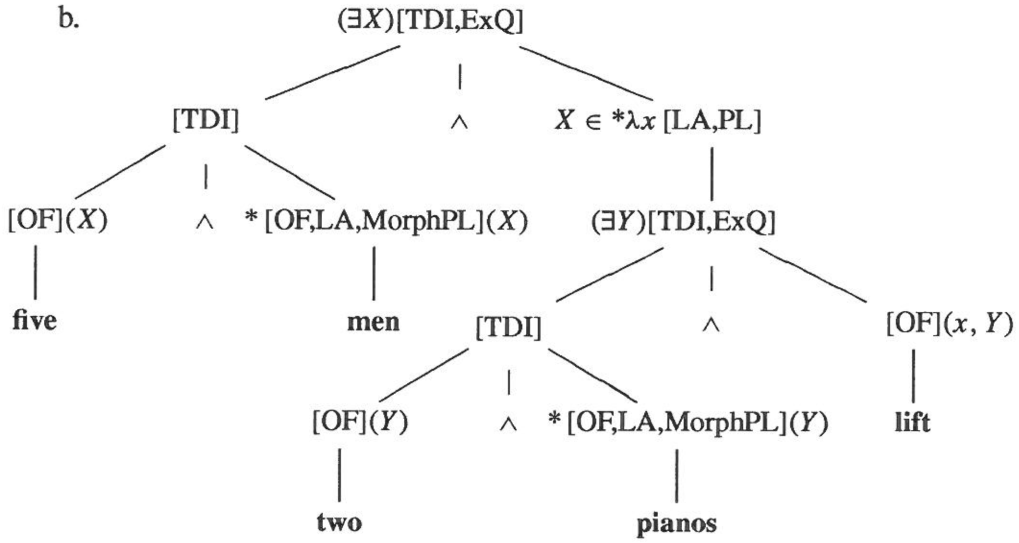
Another issue I will take up only very briefly in the next subsection is "word order" at ALF – a potential problem which is more or less irrelevant for the few sentences we are concerned with here. Of course we assume a rule of scoping, traditionally called QR, that moves one NP across another in order to gain scope over it. Thus, (84-d) and (84-e) will be generated via scoping of the object over the subject. As regards the remaining formulas in (84), observe that the object NP that triggers existential quantification must bind a variable generated as an argument of the verb. But in English, the verb and the variable to be bound precede the quantifier. One way to deal with the problem is QRing of the object. Above, however, we assumed that there is no QR simply for reasons of type theory. Now, in order to maintain that QR is a rule of scoping, we can as well assume that English is head final at D-structure. If so, the verb can simply move back at LF into its base position.

Let us look now at an arbitrarily chosen example from (84), for instance (84-b). In order to keep the trees manageable, I suppress the pl-morpheme and remind the reader of abbreviations like ***men**(X) and ***pianos**(Y) for $X \in {}^*\lambda x[\mathbf{man}(x)]$ and $Y \in {}^*\lambda y[\mathbf{piano}(y)]$, respectively. This much said it is clear that the following tree structure matches with formula (84-b):

(89) a.



b.



As the reader may verify, all remaining examples can easily be generated by varying the order of rule application.

4.3.2. Some Restrictions on Quantifier Scope

The relative scope of quantifiers within a simple clause is largely language specific. It has often been suggested that the simplest languages in this respect are those which "wear their LF on their sleeve" (Pesetsky 1987:117, 1989:51), which means that surface order unambiguously encodes scope relations. We have seen above, however, that this picture might be too simplistic. In fact, an ambiguity between scope dependent and scope independent readings is likely to exist in all languages of the world. Since these readings have different ALFs, there cannot be a one-to-one correspondence between surface structure and any kind of LF that unambiguously encodes scope relations.

Descriptively, the task of defining the simplest types of languages is then reduced to saying that in these languages one quantifier cannot gain scope over another. We have seen above that indefinite plural NPs are "suprasegmental" in corresponding at ALF to an existential quantification and a related (but optional) distributive operator associated with pluralization of the verb. Whereas the former process may be scope dependent, the latter may be scope inducing. I will call the first component of the indefinite NP its scope dependent segment; the second part of such a quantificational structure will be called its scope inducing segment. Note that these "segments" do not have a contiguous representation at ALF. Where as the first (i.e. existential quantification) is part of the indefinite NP in question, the distributional part is introduced by (PL), outside of the NP it depends on.

In consequence, then, the difference between scope dependent and scope independent readings can simply be described as different options in placing the scope inducing segments via (PL). If (PL) applies "immediately," i.e., before interpreting another indefinite NP, the result is scope dependence of that NP. If (PL) applies only later, i.e., after interpreting the second NP, the result is scope independence. Both options exist universally, so that the most simplest languages are precisely those that lack a syntactic rule like QR: If QR were allowed to move one quantifier α across another indefinite NP β, and if the scope inducing segment of α were interpreted "immediately," β would become semantically scope dependent on α. But it is precisely this kind of scope inversion that must be ruled out for the most simple languages.

Although English is not as simple as that and therefore permits genuine scope ambiguities, not all combinations of quantifiers make scope reversal available, and in some cases even the scope independent readings are awkward. This is illustrated in (90):

(90) a. Most students read five books.
b. Most students read fewer than five books.

Clearly, (90-a) has a reading where *five books* is scope dependent on *most students*, and we also get a scope independent reading, presumably of the Jackendoff type only. But *most* is one of the quantifiers of English that block QR, i.e. no other quantifier can move across it: we do not get the reading where the subject depends on the object so that for each of the five books we were allowed to choose a different majority of students, and each member of that majority would read each of the books. Quantifiers that block QR will be called *strongly scope inducing*.

Consider next (90-b). Here it is much harder to get the scope independent reading; in fact, this reading seems impossible to me. As a first hypothesis, we might assume that *fewer than five* never allows for scope independence. It is easily verified, however, that such a condition would incorrectly exclude the possibility of getting a scope independent reading for a sentence like (91):

(91) Five men read fewer than five books.

Since this reading is easily available, the proper conclusion must therefore be that, for a certain class of quantifiers, the restriction against independent readings holds only if it interacts with another class of quantifiers, e.g. when *fewer than five* interacts with *most*.

Following Liu (1990), *fewer than five* belongs to a class of quantifiers that cannot gain scope over any higher quantifier in the same clause. I will call these quantifiers *strongly scope dependent*. Whether or not an expression belongs to one of these classes is an empirical question and may be language specific. There seem to be core cases, however, that do not permit idiosyncratic variation. Following Liu (1990) we assume that quantifiers that correspond to *most* and *a certain* in English universally belong to the class of strongly scope inducing quantifiers. Modified numerals and most downward entailing quantifiers universally belong to the class of strongly scope dependent quantifiers (cf. Liu's "strongly G-specific" and "non-G-specific" classes of quantifiers). In English, the remaining quantifier expressions are free to interact, i.e., they are neither strongly inducing nor strongly dependent.

As a rule of thumb, a scope independent reading is unavailable when a strongly scope dependent quantifier in object position interacts with a strong quantifier in subject position. A quantifier is strong if and only if it is either strongly scope inducing or strongly scope dependent. Unfortunately, the rule for scope independent readings does not follow from the above conditions on QR which are still compatible with the existence of scope independent readings. If the above generalizations are indeed correct, it should be possible to characterize strong quantifiers in such a way that the rule follows automatically. I do not see how this could be achieved, and leave the topic for further research.

The above principles only cover the interaction of two quantifiers and need to be supplemented by principles that govern interaction with negation, adverbials, modals, and other scope inducing contexts. We cannot go into an investigation of these matters here, but one of the scope sensitive contexts is question formation. This topic will be dealt with more extensively in section 5.

4.4. Reciprocals

In this section we take up the issue of reciprocal interpretation. Let us start with reconsidering the ALF rule of lambda abstraction (LA). This rule would allow simultaneous binding of two argument positions, as in $\lambda x P(x, x)$. But for reasons of θ-theory argument reduction of this type will be blocked. This could be formally made explicit by a constraint on the rule that attaches variables to lexical items:

(92) In any application of (OF), all variables have different indices.

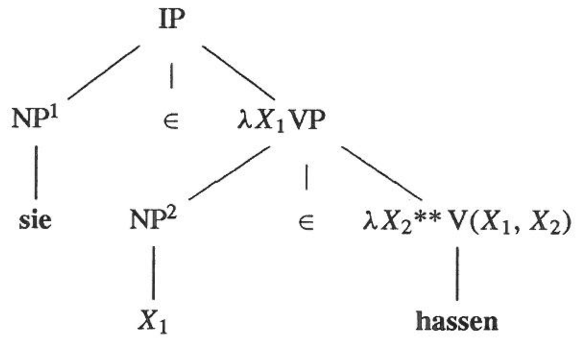
(92) is equivalent to the statement that NPs that function as arguments of the same predicate cannot be coindexed. This, however, already follows in turn from our use of superscripts; these function to identify different grammatical functions. Therefore, if two NPs are arguments of the same head, identity of superscripts would imply identity of grammatical function. Hence, we need not state (92) as a separate rule, since it follows from the fact that the indices are linked to different positions.

Now, in order to express anaphoric binding, we have to distinguish this "syntactic" index of an NP from its anaphoric index. As an illustration, let us look again at a sentence like

(93) weil sie sich hassen
because they REFL/REC hate

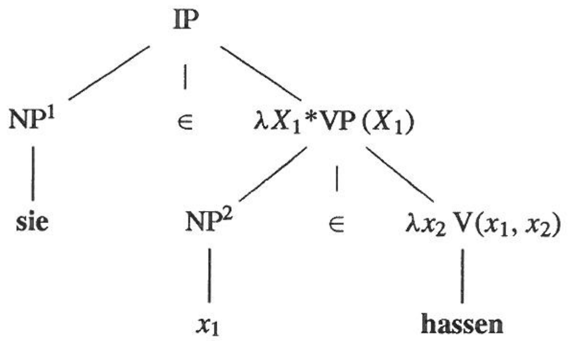
in German, which according to our analysis in section 3.1.1. has three different construals: a reflexive, a reciprocal, and presumably a purely anaphoric one. We will first analyze the neutral one. Recall our notational convention of writing "$*\cdots*\alpha(X_1, ..., X_n)$ as short for $\langle X_1, ..., X_n\rangle \in *\cdots*\lambda x_1, ..., x_n[\alpha(x_1, ..., x_n)]$. It is easy to see then that our rules will generate the following ALF as a representation for the purely anaphoric reading:

(94)



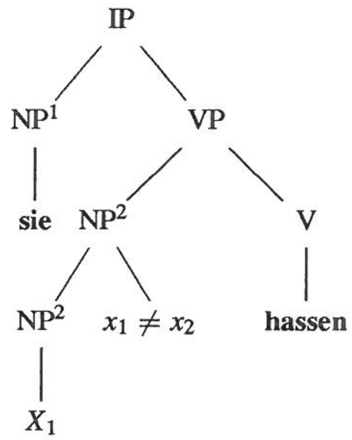
The reflexive reading will be obtained by applying basically the same rules, but in a different order:

(95)



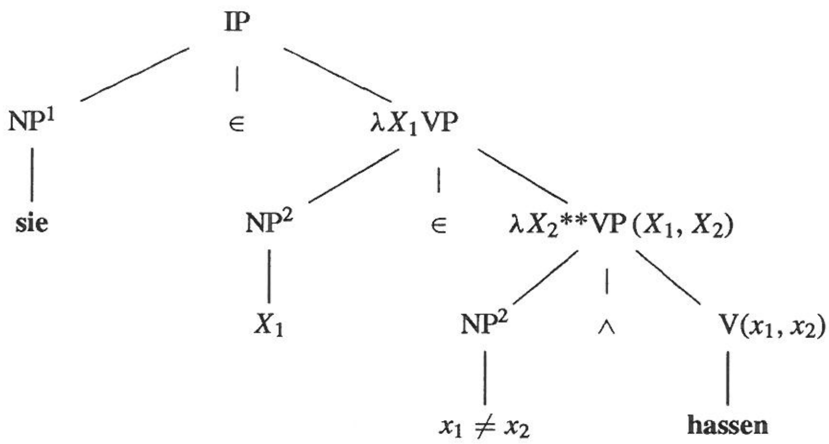
Note that the reflexive translates as X_1 in (94) and as x_1 in (95). As in previous cases, I insist that these are not really different translations as such, but only reflect differences in the scoping of the reflexive. Alternatively, we might assimilate the ALF of reflexive sentences to that of reciprocal sentences. To get the reciprocal reading, a non-identity statement must be involved which is drawn from the lexicon as part of the meaning of the reciprocal pronoun, and therefore must be generated together with the argument expression. I will adjoin this non-identity statement to NP^2, so that the complex NP in (96) represents the meaning of the reciprocal use of *sich*:

(96)



This tree cannot yet be interpreted using our rules; before being in a position to do so, we must move the lower segment of NP² into a position where it can serve as an argument. Adjoining it to VP will lead to an interpretable ALF:

(97)



Of course, the English sentence *they hate each other* will receive essentially the same analysis. The crucial difference from traditional treatments is this: although we follow standard practice in moving part of the reciprocal to the antecedent, the part we move is not a distributive operator (i.e., "each" of *each other*). Rather, we move the anaphoric argument out of the scope of such an operator, i.e., out of *.

It would take us another ten pages or so to run through all the examples we have analyzed in previous sections, but besides getting new combinations of the above rules nothing really new will happen. To close this section, then, I would like to turn to an issue that is often considered the only syntactically relevant one with respect to the anaphoric nature of reciprocals, namely binding theory. In fact, however, I have nothing new to say here about this module, i.e., the mechanism that relates anaphoric indices. I even did not represent these indices in the above representations, or have done so only indirectly when identifying subscripts of variables with superscripts of NPs. Something very similar can be found in Heim (1993), although the mechanism and motivation explicated there is different from mine and shouldn't be confused with the present use of indices. Nonetheless, there is a striking parallel to be observed: both in her paper and in the above treatment of reciprocals we use two different indices on referential NPs. For example, NP^2 in (97) also has the index 1, since it is the syntactic category of X_1. The index 2 serves as the index that binds via lambda abstraction, and the index 1 serves as an index that can be bound. This is exactly as in Heim (1993), though independently developed and motivated. This striking similarity is additional support for Heim's theory, and indeed it makes superfluous for me to discuss binding here; cf. Heim (1993) for everything that can be said about the topic.

5. Questions, Lists, and Scope

In this section I will be discussing some problems that result from trying to unify the semantics of plural operators with some standard version of the semantics of questions. I will try to motivate that plural *wh*-terms seem to call for a modified semantics of questions, the motivation being drawn from problems of scope interaction. More precisely, the problem to be accounted for is that certain quantifiers may and sometimes must have wide scope with respect to the *wh*-operators.

5.1. Simple Questions

The most well-known semantics for questions, namely the one proposed in Karttunen 1977, handles questions via a collection of true answers. Within that theory, which I adopt, (98-a) is translated as (98-b):

(98) a. Which man left?

b. $\lambda p (\exists_D x)\ \mathbf{man}(x) \wedge {}^{\vee}p \wedge p = {}^{\wedge}\mathbf{leave}(x)$

The index on the quantifier means that we quantify over individuals only. In general, true answers are derived from the set of possible answers, each of which is described as an open proposition of the form $P(x)$ (i.e., "$\mathbf{leave}(x)$" in (98-b)) such that $P(x)$ is the Nucleus of the question and a possible value of x is restricted by the property that restricts the *wh*-quantifier (i.e., "**man**" in (98-b)).

Let us now shift from singular to plural *wh*-terms as in (99):

(99) a. Which men left?

b. $\lambda p (\exists X)\ {}^{*}\mathbf{man}(X) \wedge {}^{\vee}p \wedge p = {}^{\wedge}\mathbf{leave}(X)$

Assuming that *leave* is D-reducible, it now follows that (98) and (99) have exactly the same truth conditions. This is unexpected, since there is still an intuitive difference between these sentences which could be described by saying that (99-a), but not (98), can prompt answers of the form "a and b left."

One might try to cope with the problem as follows. Let us say that the denotation of the question is a set Q of propositions. Then the relation between Q and what is considered an appropriate answer can be described as follows:

(100) a. A (possibly partial) answer is a conjunction of any non-empty subset of Q, and

b. the above sentences do not differ in meaning, but only in presupposition.

These assumptions seem fairly standard, although they require an independent calculation of uniqueness presuppositions – a matter that turns out to be far from trivial, particularly in multiple *which*-questions. In what follows I will simply assume that such an analysis is feasible, but leave it open as to how exactly the relation between meaning and presupposition should be spelled out.

The next step is to look at predicates that are not reducible. It is here that one may expect to encounter difficulties, since the above proposal cannot account for the distinction between the different kinds of predication we have associated with the group reading, the cumulative reading, and the distributive reading. Thus, consider a question like "Which composers wrote musicals?" A natural (cumulative) answer would be "Rodgers, Hart, and Hammerstein." However, this answer cannot be described as the conjunction of propositions in (101-a), because the denotation of (101-a) in Gillon's model is the set containing (101-b) and (101-c):

(101) a. $\lambda p(\exists X)\ {}^*\mathbf{composer}(X) \wedge {}^\vee p \wedge p = {}^\wedge\mathbf{write\text{-}musicals}(X)$

b. Rogers and Hart wrote a musical

c. Rogers and Hammerstein wrote a musical

Although the cumulative answer is logically implied by the set of true answers in the given model, it is not equivalent to the conjunction of (101-b) and (101-c). However, the basic intuition underlying the concept defined in (100) seems to be that cumulative answers should count as possible partial answers, although this does not yet follow from the theory. Inspecting the truth conditions embodied in (101-a) rather suggests that we have only formulated the "group reading" here, leaving the more general, perhaps "neutral" construal still unaccounted for. In order to get that reading it seems necessary to add an appropriate pluralization of the verb.

Now, given the general format of the theory, there are only two ways of adding a distributive operator. The *-operator can be posited either inside or outside the nucleus of the question, i.e., it can, in principle, appear either to the right or to the left of the equation. The two possibilities are depicted in (102):

(102) a. $\lambda p(\exists X)$ $^*\mathbf{composer}(X) \wedge {}^{\vee}p \wedge p = {}^{\wedge}(X \in {}^*\lambda x[\mathbf{write\text{-}musicals}(x)])$

b. $\lambda p(\exists X)$ $^*\mathbf{composer}(X) \wedge X \in {}^*\lambda x[{}^{\vee}p \wedge p = {}^{\wedge}\mathbf{write\text{-}musicals}(x)]$

(102-b) is reminiscent of problems with quantifying into questions, because formulas like $\lambda p(\forall x) \ldots\ p = {}^{\wedge}P(x)$ or $\lambda x \ldots\ p = {}^{\wedge}P(x)$ imply that there is only one individual x that satisfies the equation $p = {}^{\wedge}P(x)$. Clearly, this is far from being intended; hence quantifying-in or pluralization cannot be adequately analyzed by putting a distributive operator outside of the nuclear scope.

This leaves us with (102-a) as the only straightforward analysis of the neutral answers. A possible answer is determined by a set X of men chosen outside the nucleus of the question. An answer with respect to that set will assert that X (cumulatively) has the property in question; in consequence, Q will also contain "Rogers, Hart, and Hammerstein wrote musicals."

Since $\lambda x.P(x)$ is as informative as $\lambda X.{}^*P(X)$, there is an obvious relationship between the answers encoded in (102-a) and the ones in (101-a). First, we observe that if different propositions p and q are elements of (102-a) or (101-a), then neither one will logically imply the other, the reason being that the predicate **write-musicals** is not C-reducible. On the other hand, if we know that $p =$ (101-b) and $q =$ (101-c) are elements of (102-a), it is inevitable that "Rogers, Hart, and Hammerstein wrote musicals" also is, so that truely cumulative answers are somehow redundant elements in Q. Thus, it follows that the conjunction of propositions in (102-a) is the same as the conjunction of (101-a), and that (102-a) is the closure of (101-a) with respect to entailment. In consequence, both representations are informationally equivalent.

Intuitively, list answers do not contain redundant information. An optimal list answer will furthermore be exhaustive, i.e., it provides the hearer with non-redundant and complete information. According to the above analysis, a subset of (102-a) can be interpreted as a list answer. And an even smaller subset will qualify as complete information. From the above consideration it thus follows that the most informative subset of (102-a) coincides with (101-a). In consequence, cumulative predication cannot occur in list answers, when the latter are construed as maximally informative, non-redundant answers.

Summarizing so far, we have seen that only by inserting the *-operator can we allow for proper cumulative answers.[35] But according to our conception of answers, these will never appear as elements of an optimal list answer. However, since maxi-

[35] Observe that this is not the case for D-reducible predicates as in

mally informative answers can be defined on the basis of the set of possible partial answers, it follows that the cumulative analysis can in some sense be considered the more basic one, i.e., the more *general* representation that may be taken as a departure for a more refined analysis of questions.

5.2. Scope Interaction

Let us posit the case that three students out of five read both *Syntactic Structures* and *Aspects*. Consider now (103-a) and its canonical formalization (103-b):

(103) a. Which books did most students read?

b. $\lambda p(\exists X)$ $^*\mathbf{book}(X) \wedge {}^{\vee}p \wedge p = {}^{\wedge}(\exists Y)(\mathbf{most\text{-}students}(Y) \wedge Y \in {}^*\lambda y[X \in {}^*\lambda x\ \mathbf{read}(y, x)])$

Inspecting the truth conditions of (103-b) will reveal that there are three answers, namely (A) = "most students read *Aspects*," (B) = "most students read *Syntactic Structures*," and (C) = "most students read both *Aspects* and *Syntactic Structures*." Suppose we are only interested in complete answers. Since (C) logically implies (A) and (B), it follows that (103) describes a trivial list answer, namely, the subset of Q which contains just one proposition, namely (C). This answer is normally worded as "*Aspects* and *Syntactic Structures*." It is important to note here that this answer is not understood as giving a two element list.

But let us now change the scenario. Assume that students a, b, and c read *Syntactic Structures* and c, d, and e read *Aspects*. In such a situation, the above answer is felt to be inadequate, although, according to the above analysis, "*Aspects*" and "*Syntactic Structures*" are both correct complete answers that might count as elements of a list answer generated by conjoining (A) and (B). But nevertheless it would be misleading to just utter "*Aspects* and *Syntactic Structures*": the hearer would automatically infer the stronger construal described as (C) above.

One way to handle the problem is to say that, contrary to what we assumed initially, list answers are not short for a list of propositions, but rather they are elliptical for sentences, so that the above answer must, for linguistic rather than logical reasons, be construed as "Most students read *Aspects* and *Syntactic Structures*." And this sentence does not have the narrow scope reading for *most*.

(i) $\lambda p(\exists X)$ $^*\mathbf{man}(X) \wedge {}^{\vee}p \wedge p = {}^*\mathbf{leave}(X)$

Here, conjunctions of propositions like "a left" and "b left" are equivalent to the one proposition "a and b left" that results from choosing X as $\{a, b\}$. The complex proposition is not properly cumulative, since it logically implies its conjuncts.

Such an account is not really satisfactory, however, since in other contexts we do take list answers as short for listing propositions, particularly in the context of multiple questions. Furthermore, even if it were possible to maintain the above analysis, it seems to me that the full answer, namely the conjunction of (A) and (B), is still felt to be somehow inadequate. I suspect that the conjunction is to be taken as uttering two possible complete answers, rather than what is normally considered a list answer. Hence it appears that the "wide scope" property of *most* has not yet been captured by our conception of list answers:[36] even if *most* always has wide scope *within the nucleus*, this does not formally exclude the possibility of "changing majorities," i.e., it is not excluded by (103-b) that what we have called a list answer above may list books that were not read by the same majority. However, according to intuition (cf. Liu 1990), the question does not allow for such an answer. As pointed out by Liu – who presumably was the first to state the fact without really casting the problem into a formal semantic analysis – *most* still retains its property of not being interpretable as scope dependent, not even as scope dependent on the *wh*-operator.

This much said, the remarkable contrast to be accounted for is this: while *most* cannot be interpreted as scope dependent when occuring in subject position, all other subject quantifier phrases must be so interpreted, i.e., they cannot receive the scope independent interpretation, as illustrated in (104-a). The three students in question can vary from book to book, but it seems impossible to get the scope independent reading – a reading the corresponding declarative clause "at least three students read two books" can easily acquire.

(104) a. Which books did (exactly, at most, at least, etc.) three students read?

b. $\lambda p(\exists X)$ $^*\mathbf{book}(X) \wedge {}^{\vee}p \wedge p = {}^{\wedge}[X \in {}^*\lambda x[(\exists Y)$ $\mathbf{three\text{-}students}(Y)$ $\wedge\, Y \in {}^*\lambda y[\mathbf{read}(y, x)]]]$

c. $\lambda p(\exists X)$ $^*\mathbf{book}(X) \wedge {}^{\vee}p \wedge p = {}^{\wedge}[(\exists Y)$ $\mathbf{three\text{-}students}(Y) \wedge$ $Y \in {}^*\lambda y[X \in {}^*\lambda x[\mathbf{read}(y, x)]]]$

While (104-b) is unproblematic, because "three students" is scope dependent on pluralization, the more interesting possibility is (104-c). In a situation where the same three students have read *Aspects* and *Syntactic Structures*, the most complete and informative list answer is the scope independent one. Intuitively, however, an

[36] This property is to be contrasted with the inadequacy of (i):

(i) $\lambda p(\exists X)$ $^*\mathbf{book}(X) \wedge {}^{\vee}p \wedge p = {}^{\wedge}[X \in {}^*\lambda x[(\exists Y)$ $\mathbf{most\text{-}students}(Y) \wedge Y \in {}^*\lambda y[\mathbf{read}(y, x)]]]$

Here, *most* has been given narrow scope with respect to the plural operator. This analysis has already been ruled out by the rules of scope stated in section 4.3.2.

elliptical list answer will not be understood as having this reading. Scope independence of the question can only be induced by using deictic expressions like *the/these three students*. This is unexpected, since it does not follow from the rules of scope interaction that hold for non-interrogative clauses.

To summarize, we are faced with two problems. One is to enforce a wide scope interpretation for *most* and *certain*. The second is to enforce a narrow scope interpretation for all other quantifiers. We will deal with the first problem in the next section. The second problem can be solved, although perhaps not in a most explanatory way, by alluding to a syntactic condition that blocks representations like (104-c). The obvious way to state the generalization is this:

(105) Within the nucleus of a question, all quantifier expressions must be interpreted as scope dependent on the distributive part of *wh*-quantifiers.

It now follows that *-operators being induced by a *wh*-operator always have wide scope within the nucleus. This gives the correct results, except for one case: (105) together with the general restriction on narrow scope for *most* blocks a well formed representation for (103-a), unless we can find some way of moving *most* out of the nucleus (without treating the quantifier on a par with a *wh*-phrase). In other words, we need to find a way of quantifying into questions, and I will show in the next section how this can be done.

5.3. Quantifying-in

Intuitively, what we want as answers is elements of a list, or a conjunction thereof. Think of these elements as elements in a set of propositions that satisfies a certain condition. Usually, this condition states that for each element $x \in X$ a certain proposition of the form $P(x)$ is in that set. For instance, one would like to say that each proposition of the form "write-musicals(x)" should be an element of the list, if only x is chosen appropriately. Formally, we represent the smallest set satisfying a certain condition C as the intersection of all sets with property C. Thus, what we are interested in is something like this:

(106) $\bigcap\{Q' : (\forall q)[(\exists X)\ {*}\mathbf{composer}(X) \wedge q = {^\wedge}{*}\mathbf{write\text{-}musicals}(X)] \rightarrow q \in Q'\}$

This set describes the set of all possible answers to the question of which composers write musicals.[37] In order to make a list out of it one can put all these answers into a single proposition by simply conjoining the set via intersection. Nor-

[37] The reader should not be dismayed by the fact that (106) looks unduly complicated. In fact, it is easy to see that (106) is equivalent to (i-a), which in turn is the same as (i-b):

mally, however, we are not interested in all possible answers, but only in a subset thereof, e.g. the set of answer that are compatible with presuppositions, the set of true answers, the set of most informative answers, etc. I will put the condition that answers should be true directly into the truth conditions. Matters of informativeness and presupposition, however, will not be formally incorporated into the semantics of questions. We are thus interested in a subset Q of (106) such that the conjunction over that smaller list will be regarded as a true (and perhaps most informative, etc.) answer. Thus, we now look at meaning representations of the form given in (107), but this time each p will already encode a (possibly complete) list answer (rather than an element of a list).

(107) $\lambda p(\exists Q)\,{}^{\vee}p \wedge p = \bigcap Q \wedge Q \subseteq \bigcap\{Q' : (\forall q)[(\exists X)\ {}^{*}\mathbf{composer}(X) \wedge q = {}^{\wedge *}\mathbf{write\text{-}musicals}(X)] \rightarrow q \in Q'\}$

It is important to note that list answers are not conjunctions over the elements of (107); rather we have already incorporated the list into each p by conjoining over Q.

The next task is to formalize quantifying-in. In order to do so we only have to identify the landing sites of quantifiers. In general, there will be two types of quantifying-in. One will be enforced by the rule that *most* cannot have a scope dependent interpretation when in subject position. This is quantifying-in as shown in (108):

(108) $\lambda p(\exists Q)\,{}^{\vee}p \wedge p = {}^{\wedge}[(\exists Y)\ \mathbf{most\text{-}students}(Y) \wedge {}^{\vee}\bigcap Q \wedge Q \subseteq \bigcap\{Q' : (\forall q)[(\exists X)\ {}^{*}\mathbf{books}(X) \wedge q = {}^{\wedge}Y \in {}^{*}\lambda y[X \in {}^{*}\lambda x.\mathbf{read}(y,x)]] \rightarrow q \in Q'\}]$

Here it will pay that we did not let (108) describe elements of the list; since each p already represents the list, we will not get changing majorities, although we might get more than one true and complete answer. This is exactly the intuition I wanted to formalize: although there are two answers in the situation described above, they cannot be combined so as to enter into what has been called a complete list answer. This is why *most* can get a wide scope interpretation.[38]

The second type of quantifying-in is with a universal quantifier as in (109):

(i) a. $\bigcap\{Q' : \lambda q[(\exists X)\ {}^{*}\mathbf{composer}(X) \wedge q = {}^{\wedge *}\mathbf{write\text{-}musicals}(X)] \subseteq Q'\}$

b. $\lambda q[(\exists X)\ {}^{*}\mathbf{composer}(X) \wedge q = {}^{\wedge *}\mathbf{write\text{-}musicals}(X)]$

It will soon become clear that the more complicated reformulation will enable us to formulate quantifying-in of universal quantifiers.

[38] Of course, we have to add that this type of quantifying-in is only a last resort operation; otherwise we would predict wide scope to always be possible. Rather, we want scoping in exactly those cases that result from a conflict between rule (105) and the scope rule for *most*.

(109) a. Which book did every Englishman read?

b. $\lambda p(\exists Q)\check{}p \wedge p = \hat{}[\check{}\bigcap Q \wedge Q \subseteq \bigcap\{Q' : (\forall q)(\forall y \in [\![\text{Englishman}]\!]_M)$
$[(\exists x)\ \mathbf{book}(x) \wedge q = \hat{}\mathbf{read}(y,x)] \rightarrow q \in Q'\}]$

It is easy to see that the representation in (109) is equivalent to the one we get for (110-a):

(110) a. Which Englishman read which book?

b. $\lambda p(\exists Q)\check{}p \wedge p = \hat{}[\check{}\bigcap Q \wedge Q \subseteq \bigcap\{Q' : (\forall q)[(\exists y \in [\![\text{Englishman}]\!]_M)$
$(\exists x \in [\![\text{book}]\!]_M)\ q = \hat{}\mathbf{read}(y,x)] \rightarrow q \in Q'\}]$

And in fact we get identical complete list answers. The difference in meaning, then, will simply result from a difference in presupposition: in (109) it is presupposed that every Englishman read at least one book, but no such presupposition has been made in (110).[39]

Things are likely to become more complicated when turning to a plural version of (110-a). Most probably, an adequate reading of (111-a) is the one in (111-b); but it is not clear to me whether the alternative strong reading in (111-c) should also be permitted:

(111) a. Which Englishmen read which books?

b. $\lambda p(\exists Q)\check{}p \wedge p = \hat{}[\check{}\bigcap Q \wedge Q \subseteq \bigcap\{Q' : (\forall q)[(\exists Y)\ {}^*\mathbf{Englishman}(Y)$
$\wedge\ (\exists X)\ {}^*\mathbf{book}(X) \wedge q = \hat{}{}^{**}\mathbf{read}(y,x)] \rightarrow q \in Q'\}]$

c. $\lambda p(\exists Q)\check{}p \wedge p = \hat{}[\check{}\bigcap Q \wedge Q \subseteq \bigcap\{Q' : (\forall q)[(\exists Y)\ {}^*\mathbf{Englishman}(Y)$
$\wedge\ (\exists X)\ {}^*\mathbf{book}(X) \wedge q = \hat{}[Y \in {}^*\lambda y[X \in {}^*\lambda x.\mathbf{read}(y,x)]]] \rightarrow q \in Q'\}]$

Be this as it may, it is instructive to see that – given the D-reducibility of "read" – (111-c) turns out equivalent, modulo presupposition, with the meaning of "Which books did every Englishman read?". I leave it to the reader to evaluate these consequences. In any case, all of the above sentences induce the same complete list answers, and since, under normal circumstances, we are only interested in these, it might be difficult to judge the differences between the above construals.

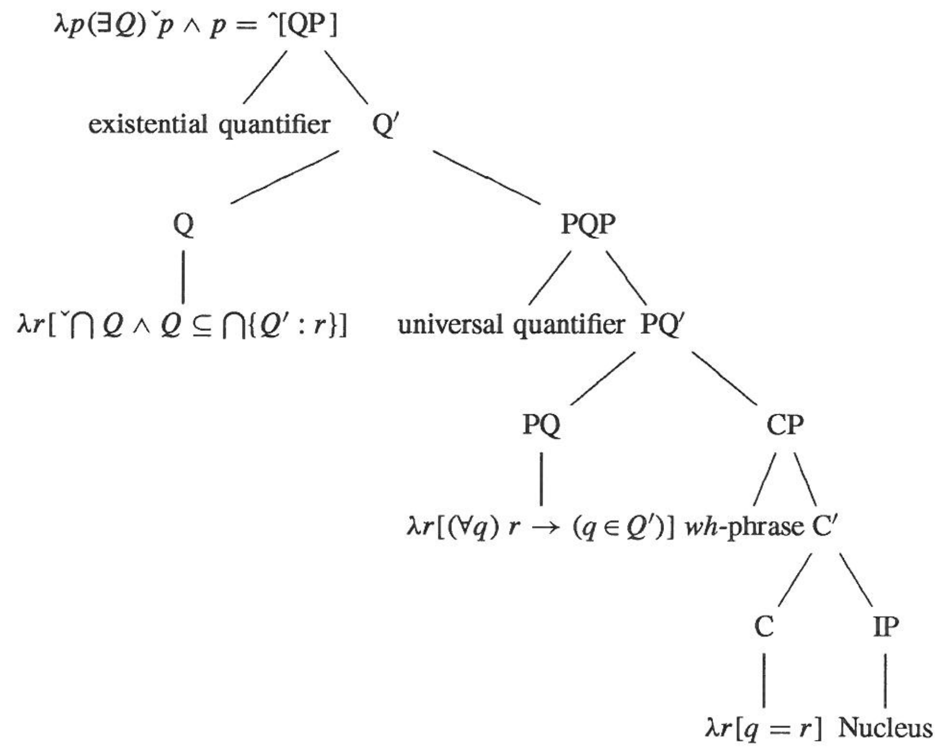
5.4. Augmented Logical Form

In the last section we have identified three locations where wide scope interpretation outside the nucleus of a question is possible: one where *wh*-operator scope is de-

[39] Again, I will simply disregard presupposition formally, but note in passing that (110) and (109) do not have the uniqueness presupposition associated with a simple *which*-question.

termined, one where existential wide scope interpretation is possible, such that the quantifier is outside of the "list scope," and one where universal wide scope interpretation is possible, but the quantifier is still inside of the list scope. This requires a tripartition of the "meaning" representation of the question. We will decompose this meaning according to the following principle: the *wh*-operator(s) will be in the Spec position of CP, the universally quantified material will occupy the Spec position of a small Question Phrase, the Proto-Question-Phrase "PQP," and the wide scope existential quantifier will be in the Spec position of a big Question Phrase "QP." This is schematically depicted in (112):

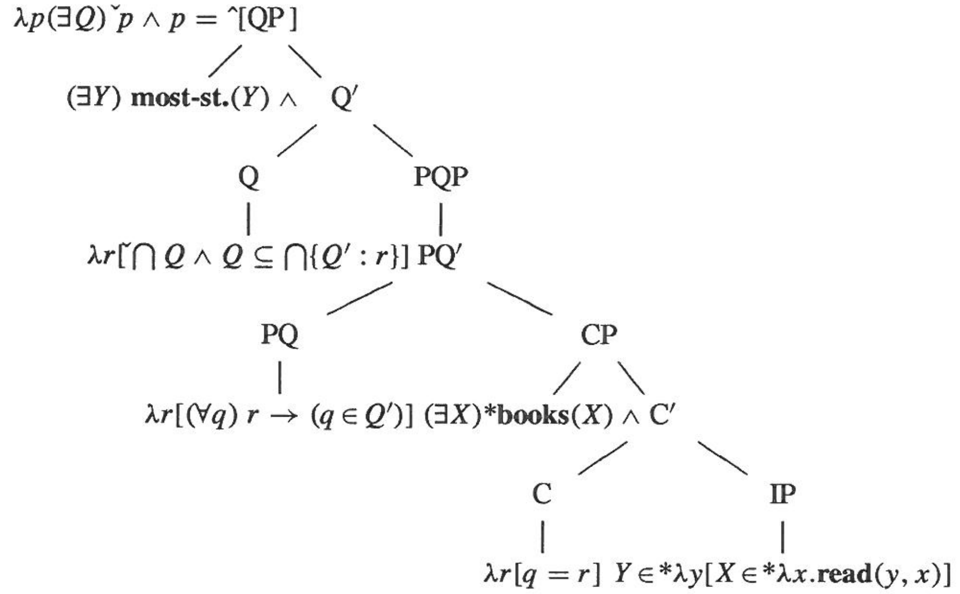
(112)



Semantic interpretation will then proceed in a completely compositional way; the only thing I have added to (112) is the interpretation rule that converts the proposition embedded under QP into the meaning of the question. Apart from that, the representations can be arrived at by simply inserting lexical material and phonologically empty heads C, PQ, and Q into the tree. Insertion of these heads is obligatory. This follows from the fact – perhaps to be encoded more formally by the use of appropriate indices – that each head, beginning with C, introduces a variable that cannot be left free, and therefore must be bound by a quantifier in a higher abstract head.

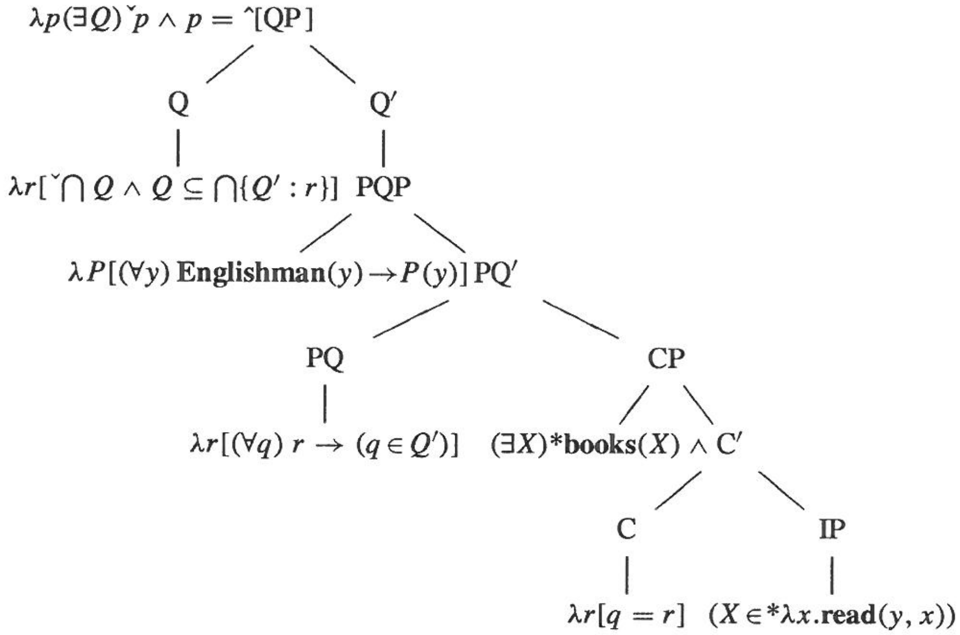
As a concrete example, consider now the position of the existential quantifier in (112) which is spotted by *most* in (113). This gives us our previous analysis in (108):

(113)



(To increase readability, I have omitted brackets that indicate functional application; furthermore I have put the conjunction following existential quantification directly into the formulas, rather than attaching it to symbols of the tree.) Similarly, we get the alternative quantifying-in structure by moving the universal quantifier into its appropriate position:

(114)



This completes our analysis of quantifying in. It might seem that the semantics is very complicated, but recall from footnote 37 that in part we only reformulated the standard theory. Furthermore, alternative approaches, e.g. Chierchia (1992) do not fair much better with respect to complexity, although they are clearly less compositional than the present treatment.[40]

One final remark concerns presupposition. As is easily verified, the scope relations between operators in the above representations exactly correspond to the scope relations that hold in the presuppositions of the respective sentences. Thus, quantifying-in of *every* results in wide scope interpretation with respect to the existential claim made by the *wh*-operator, exactly as one would expect. Likewise, *wh*-operators always have wide scope with respect to quantifiers in the nucleus, again as it should be. But since this holds in general, we have an additional argument for quantifying-in of *most*. Clearly, *most* must also have wide scope in presuppositional representations, and this property can be read off directly from the question representation only if there is quantifying-in of *most*.

References

Abusch, D. (1993): The Scope of Indefinites. DYANA-2, Dynamic Interpretation of Natural Language. ESPRIT Basic Research Project 6852. Deliverable R2.2.A, Part I.

Cartright, H. (1993) On Plural Reference and Elementary Set Theory. *Synthese 96*, 201-254.

Chierchia, G. (1992): Questions with Quantifiers. *Natural Language Semantics 1*, 181-234.

Cresswell, M. (1973): *Logics and Languages*. Methuen, London.

van der Does, J. (1992): *Applied Quantifier Logics*. Academisch Proefschrift, Universiteit van Amsterdam.

Dougherty, R. (1974): The Syntax and Semantics of *each other* Constructions. *Foundations of Language 12*, 1-47.

Engdahl, E. (1986): *Constituent Questions*. Kluver, Dordrecht.

Fanselow, G. (1985): What is a Possible Complex Word? In: J. Toman (ed.): *Topics in German Syntax*. Foris, Dordrecht. 289-318.

Gillon, B. (1987): The Readings of Plural Noun-Phrases in English. *Linguistics and Philosophy 13*, 477-485.

Hamblin, C. L. (1973): Questions in Montague English. *Foundations of Language 10*, 41-53.

[40] Part of the complexity results from integrating into the system relational answers like "his mother" discussed by Engdahl (1986). I didn't try to integrate these types of answers into the present system.

Heim, I. (1982): *The Semantics of Definite and Indefinite NP's*. Ph.D. dissertation, University of Massachusetts, Amherst. Published in 1989 by Garland, New York.

Heim, I. (1993): Anaphora and Semantic Interpretation. A Reinterpretation of Reinhart's Approach. SfS-Report 07-93, Seminar für Sprachwissenschaft, Tübingen.

Heim, I., H. Lasnik, & R. May (1991): Reciprocity and Plurality. *Linguistic Inquiry 22*, 63-102.

Higginbotham, J. & R. May (1981): Questions, Quantifiers, and Crossing. *The Linguistic Review 1*, 41-80.

Horn, L (1972): *On the Semantic Properties of Logical Operators in English*. Ph.D. thesis, University of California, Los Angeles.

Jackendoff, R. (1972): *Semantic Interpretation and Generative Grammar*. MIT-Press, Cambridge.

Karttunen, L. (1977): The Syntax and Semantics of Questions. *Linguistics and Philosophy 1*, 3-44.

Keenan, E. (1987): Unreducible n-ary Quantifiers in Natural Language. In P. Gärdenfors (ed.): *Generalized Quantifiers: Linguistic and Logical Approaches*. Studies in Linguistics and Philosophy 31, Reidel, Boston. 109-150.

Krifka, M. (1986): *Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. Doctoral Dissertation, University of Munich. Published by Wilhelm Finck, Munich, 1989.

Krifka, M. (1990): Boolean and Non-Boolean 'and.' In L. Kálmán & L. Pólos (eds.) *Papers from the Second Symposium on Logic and Language*. Akadémiai Kiadó, Budapest. 161-188.

Langendoen, D. T. (1978): The Logic of Reciprocity. *Linguistic Inquiry 9*, 177-197.

Lewis, D. (1975): "Adverbs of Quantification," in E. Keenan (ed.): *Formal Semantics of Natural Language*, Cambridge UP, Cambridge, 3-15.

Lewis, D. (1991): *Parts of Classes*. Blackwell, Cambridge, Mass.

Link, G. (1983): The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach. In R. Bäuerle, Chr. Schwarze, & A. v. Stechow (eds.) *Meaning, Use, and Interpretation of Language*. De Gruyter, Berlin, 303-323.

Link, G. (1984): Hydras: On the Logic of Relative Clause Constructions with Multiple Heads. In F. Landmann & F. Veltman (eds.): *Varieties of Formal Semantics*. Foris, Dordrecht. 245-257.

Liu, F. (1990): Scope Dependency in English and Chinese. Diss. UCLA.

May, R. (1985): *Logical Form: Its Structure and Derivation*, MIT Press, Cambridge/Mass.

Moltmann, F. (1992): Reciprocals and *Same/Different*: Towards a Semantic Analysis. *Linguistics and Philosophy 15*, 411-462.

Müller, G. & W. Sternefeld (1993): Unambiguous Binding and Improper Movement. *Linguistic Inquiry 24*, 461-507.

Partee, B. (1984): Compositionality. In F. Landman & F. Veltman (eds.): *Varieties of Formal Semantics*. Foris, Dordrecht, 281-311.

Pesetsky, D. (1987): Wh-in-Situ: Movement and Unselective Binding. In E. Reuland and A. ter Meulen (eds.): *The Representation of (In)Definiteness*, Cambridge, Mass.: MIT Press, 98-129.

Pesetsky, D. (1989) Language Particular Processes and the Earliness Principle. Ms. MIT, Cambridge, Mass.

Quine, W. V. O. (1980): *Set Theory and Its Logic*. Revised Edition. Harvard University Press, Cambridge.

Quine, W. V. O. (1982): *Methods of Logic*. Fourth edition. Holt, Reinhart & Winston, New York.

Scha, R. (1981): Distributive, Collective and Cumulative Quantification. In J. Groenendijk et. al. (eds.): *Formal Methods in the Study of Language*. Universiteit Amsterdam, Mathematical Center, Amsterdam, 483-512

Scha, R. & D. Stallard (1988): Multi-Level Plurals and Distributivity. In *Proceedings of the 26th Annual Meeting of the ACL*. Buffalo, NY, 17-24.

Sharvy, R. (1980): A More General Theory of Definite Descriptions. *The Philosophical Review* 89, 607-624.

Schwarzschild, R. (1991): *On the Meaning of Definite Plural Noun Phrases*. PhD. Diss. University of Massachusetts.

Stechow, A. v. (1993): Die Grundlagen der Syntax. In J. Jacobs et. al. (eds.): *Syntax. An International Handbook of Contemporary Research. Volume 1*. De Gruyter, Berlin. 1-89.

Zimmermann, E. (1993): The Proper Treatment of Opacity in Certain Verbs. *Natural Language Semantics 1*, 149-179.