MATHEMATICAL CENTRE TRACTS 136

FORMAL METHODS IN THE STUDY OF LANGUAGE

PART 2

edited by
J.A.G. GROENENDIJK
T.M.V. JANSSEN
M.B.J. STOKHOF

MATHEMATISCH CENTRUM AMSTERDAM 1981

1980 Mathematics subject classification: 03-06,03A05,03B15,03B45,03B50,03B60,03C20
03C80,03E70,03G05,03G70,68F05,68G20,
68F25

ISBN 90 6196 213 7

CONTENTS

PART 1

CONTENTS
PART 2

ADDRESSES OF THE AUTHORS

EMMON BACH, Department of Linguistics, University of Massachusetts, Amherst, Mass. 01003, USA.
RENATE BARTSCH, Department of Philosophy, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.
JOHAN VAN BENTHEM, Department of Philosophy, University of Groningen, Westersingel 19, 9718 CA Groningen, The Netherlands.
HARRY BUNT, Institute for Perception Research, Den Dolech 2, 5612 AZ Eindhoven, The Netherlands.
ÖSTEN DAHL, Department of Linguistics, University of Göteborg, Erik Dahlbergsgatan 11B, 411 26 Göteborg, Sweden.
PETER VAN EMDE BOAS, Department of Mathematics, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.
JOYCE FRIEDMAN, Department of Communication and Computer Science, University of Michigan, Ann Arbor, Michigan 48109, USA.
GERALD GAZDAR, School of Social Sciences, University of Sussex, Brighton BN1 9QN, England.
JEROEN GROENENDIJK, Department of Philosophy, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.
ROLAND R. HAUSSER, Seminar für Deutsche Philologie II, Universität München, Schellingstrasse 3, 8000 München 40, F.R.G.
JAAP HOEPELMAN, Institut für Romanistik, Kepplerstrasse 17, 7 Stuttgart 1, F.R.G.
THEO M.V. JANSSEN, Mathematisch Centrum, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands.
HANS KAMP, Bedford College, University of London, Regent's Park, London NW1 4NS, England.
LAURI KARTTUNEN, Department of Linguistics, University of Texas, Austin, Texas 78712, USA.
EDWARD L. KEENAN, Department of Linguistics, Tel Aviv University, Tel Aviv, Israel
and
Department of Linguistics, U.C.L.A., 405 Hilgard Avenue, Los Angeles, Cal. 90024, USA.
EWAN KLEIN, School of English Language and Literature, The University, Newcastle upon Tyne NE1 7RU, England.
JAN LANDSBERGEN, Philips Research Laboratories, WB3, 5600 MD Eindhoven, The Netherlands.
ALICE TER MEULEN, Max Planck Institute, Berg en Dalseweg 79, 6522 BC Nijmegen, The Netherlands.
BARBARA PARTEE, Department of Linguistics, University of Massachusetts, Amherst, Mass. 01003, USA.
IVAN A. SAG, Department of Linguistics, Stanford University, Stanford, Cal. 94305, USA.
REMKO J.H. SCHA, Philips Research Laboratories, WB3, 5600 MD Eindhoven, The Netherlands.
MARTIN STOKHOF, Department of Philosophy, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.
ANNA SZABOLCSI, Hungarian Academy of Sciences, Szentharomsag U.2, 1250 Budapest 1, Hungary.
FRANK VELTMAN, Department of Philosophy, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.
HENK J. VERKUYL, Department of Linguistics, University of Utrecht, Padualaan 14, 3508 TB Utrecht, The Netherlands.

UNBOUNDED DEPENDENCIES IN PHRASE STRUCTURE GRAMMAR: SLASH CATEGORIES VS. DOTTED LINES

by

Lauri Karttunen

1. INTRODUCTION

From Sussex to Stanford, and in many places between, there is a renewed interest in old-fashioned phrase structure syntax. It is widely believed that it is possible to find adequate, pure phrase structure solutions for many syntactic phenomena that once motivated the introduction of transformational rules. It has also become increasingly clear, especially to those who work on model-theoretic semantics, that semantic interpretation often is facilitated rather than impaired by the absence of transformations. This is true even for constituent questions and relative clauses, which traditionally have been generated by means of unbounded movement and deletion operations. Even for these constructions, there is no compelling reason to assume that syntax has to provide semantics with a sequence of phrase structure trees. One level of syntactic representation, the surface constituent structure, is sufficient for the purposes of semantic interpretation.

A number of approaches are currently being pursued that differ in many other respects but agree on providing only one level of syntactic description. Gerald Gazdar, Ivan Sag, Emmon Bach and Richard Saenz, Stanley Peters, and myself, all assume that syntactic structures are to be interpreted model-theoretically; Joan Bresnan and Ronald Kaplan generate functional structures as an intermediate level of representation. With regard to the treatment of unbounded dependencies, the group splits differently. Gazdar's grammar has a set of basic rules which are designed to generate sentences that do not contain constituent questions or relative clauses. There is a set of conventions which systematically multiply the basic set of rules to produce a set of additional rules. These derived rules contain a special kind of new category symbols, "slash categories", for generating phrases which lack some constituent whose presence is required by the corresponding basic rule. For example, the phrase "I was talking about" in the relative clause "who I was taling about" is an "$S/\overline{NP}$" in Gazdar's terms, that is, a sentence which somewhere lacks a noun phrase. Bresnàn and Kaplan and we at Texas do not make use of "slash categories". As I will demonstrate later on, it is possible to use the same set of basic rules to recognize and generate both complete and incomplete constituents. Instead of adding new categories and rules, as Gazdar does, one can modify traditional notions of how the rules function and thereby achieve the same effect.

Although there may not be any important difference between the two approaches in the cases I will be discussing, it seems to me that the Texas way of doing things can more easily be generalized and extended to cases that presently are beyond any phrase structure description. What I have in mind here are the cases in Scandinavian languages where resumptive pronouns alternate with deletions in relative clauses and constituent questions. Verbal complements in Dutch constitute another difficult case for any type of non-transformational syntax because of overlapping dependencies. (See MALING & ZAENEN.)

2. A SIMPLE PHRASE STRUCTURE GRAMMAR

I will start by outlining a phrase structure grammar for a set of simple English sentences and then discuss the question of what should be done to extend it to cover questions and relative clauses. In order to demonstrate how one can do without Gazdar's derived categories, I will describe in some detail how a parser deals with incomplete constituents, phrases in which some part is not manifested. The work that I am reporting on actually grew out of interest in parsing techniques, and the grammar itself is a fairly standard phrase structure treatment of English noun phrases, auxiliaries, questions, and relative clauses. It has many features in common with the descriptions sketched by Gazdar, Sag, Peters, Saenz, and other friends of phrase structure grammar.

3. FORMAT OF RULES

The rules of my grammar - "PSG" for short - are of the following form. (See BEAR & KARTTUNEN for more details.)

(1) $A \rightarrow B\ C \ldots$ (Conditions).

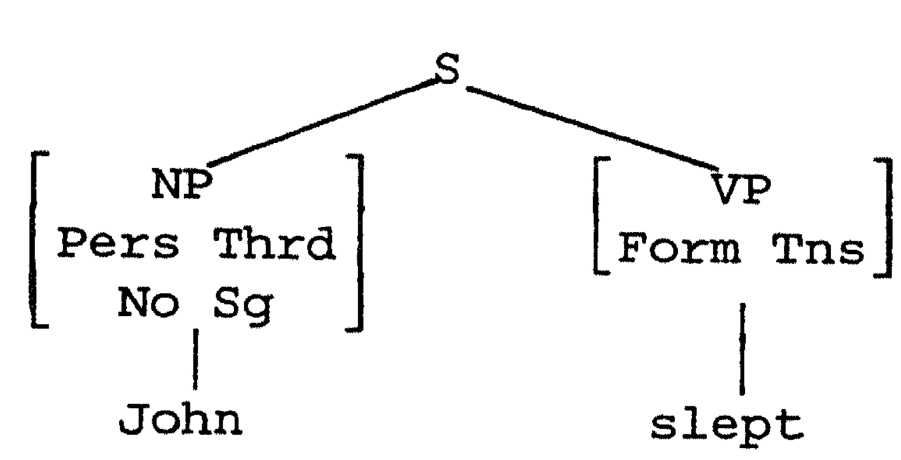
Each rule consists of an ordinary context-free phrase structure rule and a set of conditions for the application of the rule. The rules are intended to be interpreted as tree-building operations or, alternatively, as well-formedness conditions on trees. The rule in (1) says in effect that a constituent of category A may consist of a B, a C, and ... under certain conditions. In PSG, these conditions are always local; that is, they refer only to properties of the structures that are involved, not to the surrounding context.

The conditions are stated in terms of syntactic features. It is assumed that each constituent has a set of features in addition to its syntactic category, features such as number, person, case, form, subcategory, etc. Each feature in turn consists of the name of the feature and its value for the particular constituent, for example, (Number Sg), (Pers Thrd), (Case Gen), (Form Inf), (Subcat Tr), etc. The conditions in rules typically require the presence of particular features, agreement of values, and so on. Here are some examples of (somewhat simplified) PSG rules:

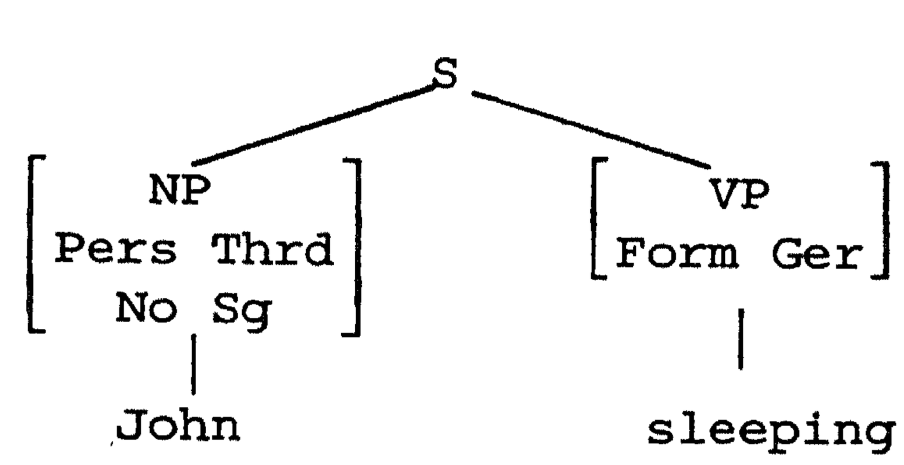
(2) S $\rightarrow$ NP VP (CHECK 3 'Form 'Tns) (AGREE 2 3 '(Pers No)).

The rule in (2) says that an S may consist of an NP and VP provided that the third constituent mentioned by the rule, namely the VP, has the feature (Form Tns). It furthermore requires that the second and the third component, that is, the NP and the VP, agree in person and number. In other words, the rule accepts structures like (3a) and rejects trees like (3b).

(3) (a)



(b)

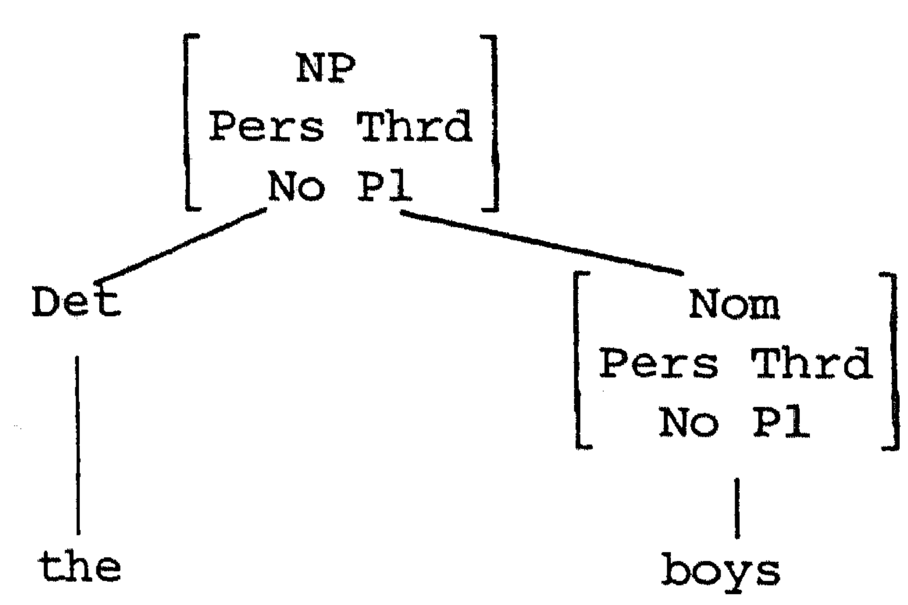


Note that the condition on agreement, (AGREE 2 3 '(Pers No)), is met in this case although the VP in question has no such features. This is because the condition is defined to hold when there is no clash of values. In effect, AGREE requires non distinctness rather than identity of feature values. We will also need a stricter form of agreement, called MATCH, in some of the later rules.

(4) NP $\rightarrow$ Det Nom (AGREE 2 3 'No) (PERK 3 '(Pers No Case)).

Rule (4) says that an NP may consist of a determiner and a Nom provided that the two agree in number. It accepts structures like (5).

(5)



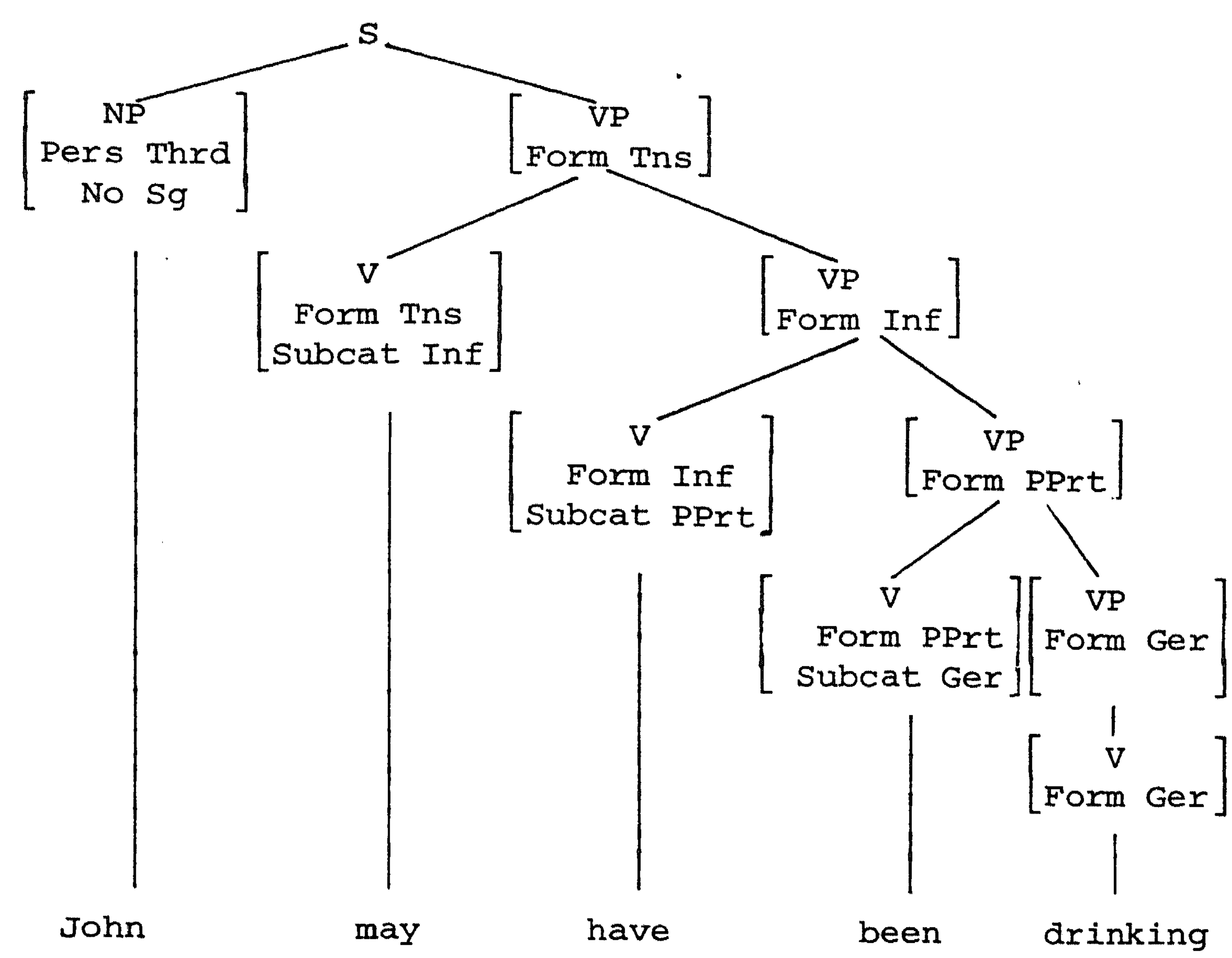
The second condition in (4) states that the number and case features of the Nom constituent are to be "perked up", that is, assigned as features to the resulting NP. Thus the number of the NP is determined by the number of its Nom constituent. In terms of structural configurations this is a case of "vertical agreement", as opposed to "lateral agreement". It is stated as feature percolation only because the PSG rules were intended to be used by a parser that assembles structures from bottom up.

Because of the presence of features, the category labels of PSG are in effect complex symbols. It would of course be possible to eliminate all features from the grammar by simply replacing all composite names by atomic category labels. This would eliminate all the conditions and turn PSG into an equivalent context-free grammar of the standard type. But since it would multiply the number of categories and rules and only obscure linguistic generalizations, carrying out such manipulations would serve no other purpose except to prove context-freeness.

5. AUXILIARIES

For a more complicated example of how such rules work, let us look at the rule that generates/accepts strings of auxiliary verbs. The tree in (6) shows the structure assigned to the sentence "John may have been drinking". (The structure has been simplifed by leaving out features that have no bearing on the issue at hand.)

(6)



The rule that generates the cascade of VPs above is given in (7).

(7) VP → V VP (MATCH 2 'Subcat 3 'Form) (PERK 2 'Form).

It is assumed here that the lexical entries for all verbs have the feature Form. Possible values for this feature are Tns, Inf, PPrt, and Ger (Tensed, Infinitive, Past Participle, and Gerund, respectively). This feature simply indicates what form in the verb paradigm the entry in question has. Besides the form feature, a verb may also have a subcategorization feature, called Subcat, whose values encode the syntactic environments in which the verb can occur. For example, the word "may" and all other modals have the feature (Subcat Inf) because they combine with infinitive phrases; the perfective "have" is marked with (Subcat PPrt) to indicate that it combines with past participles; the progressive "be" has the feature (Subcat Ger). Rule (7) says in effect that a verb and a verb phrase combine to produce a larger verb phrase provided that the form of the daughter VP matches the subcategory of its left sister. Furthermore, it specifies that the resulting VP is of the same form as the verb it dominates. For example, looking at the diagram in (6) we see that "been drinking" constitutes a verb phrase because the form of "drinking" matches the subcategory of "been"; and because "been" has the form PPrt, so does "been drinking". This complex VP in turn combines with "have" to make "have been drinking" since its form matches the subcategory of "have". The resulting VP inherits its form from "have". Finally, "may" and "have been drinking" combine to "may have been drinking". "May" is tensed and, therefore, the whole verb phrase has the feature (Form Tns), so that it can combine with the subject noun phrase to form a sentence (see rule 2 above).

This analysis of auxiliaries is very similar to what Gazdar, Pullum, and Sag have proposed and the basic idea goes back at least to a 1977 article by PULLUM & WILSON. Just as they did, I am assuming here that most of the unwanted sequences of auxiliaries are blocked because of paradigm gaps. English modals do not have infinitive or past participle or gerund forms, hence sequences such as "may must", "have mayed", or "be maying" cannot be generated. The perfective "have" lacks the progressive form; therefore, we do not have "*John is having drunk beer". (There may be additional reasons for its anomaly as well.) The only problem that does have to be solved in some other way is the limited distribution of "do". "Do" is like a modal because it is subcategorized for infinitives but it cannot combine with the perfective "have" or any type of "be". These two verbs must somehow be marked as exceptional with respect to their behavior towards "do" to block phrases like "does have drunk" and "does be tired"; I see no hope of deducing it from other facts about English. The most straight-forward, although blatantly ad-hoc, solution is to create a feature, say, Do?, precisely for this purpose. Only three verbs need to have this feature, namely "do" itself, with the value Yes, and the perfective "have" and any kind of "be" with the value No. Thus the lexical entries for these verbs would look something like (8):

(8) (a) does (Form Tns) (Subcat Inf) (Do? Yes) ...
(b) have (Form Inf) (Subcat PPrt) (Do? No) ...
(c) be (Form Inf) (Subcat (ger Pred Pass..)) (Do? No) ...

By augmenting rule (7) to read as in (9) we can block out precisely the unwanted sequences involving the verb "do". All of the forms in (10a) are now generated by the grammar; all of the ones in (10b) are rejected.

(9) VP → V (not) VP (MATCH 2 'Subcat 4 'Form) (PERK 2 'Form)
(AGREE 2 4 'Do?)

(10) (a) does talk, does not talk, has not talked, is not talking
(b) *does have talked, *does be talking.

Because the syntax of English auxiliaries is traditionally presented in textbooks as a problem that calls for a transformational solution, it is worth observing that in this case at least a phrase structure grammar actually can account for all the facts without any loss of elegance or generality.

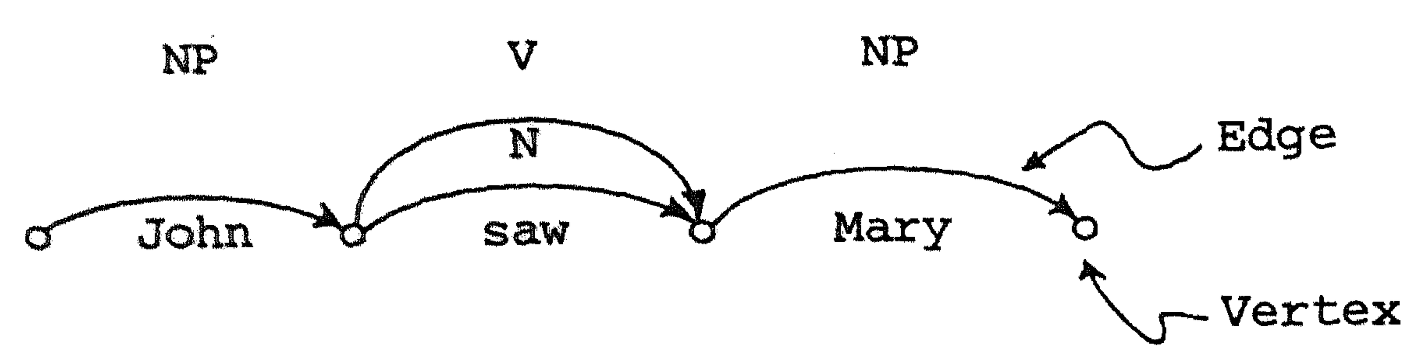
6. PARSING TECHNIQUES

Let us now consider the question of how one implements a grammar of the sort I have just outlined. A phrase structure grammar can be put to use in two ways: one can use it for generation of for recognition. I will pick the latter task because it is more challenging, although I could make exactly the same points by talking about how sentences are generated.

After many years of neglect, parsing is again becoming a popular topic. There are many articles being written about different techniques and approaches: top-down vs. bottom-up, determinism vs. non-determinism. Some people are interested in finding the most efficient parsing method for a particular type of grammar, others are trying to model what goes on in the mind of a hearer when he imposes a syntactic structure on the stream of input words. I will have nothing to say on that score. For my purpose it will suffice to pick any of the great number of proven strategies as long as it can be modified to deal with incomplete constituents in the way I want. My only objective here is to demonstrate how one can implement a phrase structure grammar in a way that makes it possible to dispense with Gazdar's slash categories. The parsing technique that I have chosen is a simplified version of Ronald Kaplan's "General Syntactic Processor", also described in detail by Terry Winograd in his forthcoming book. In order to keep things maximally simple I will assume that the parser works strictly from top down, although Kaplan's framework actually can accommodate any mixture of top-down, bottom-up strategies. I will start by discussing the overall design of the parser, then look at one simple example, and conclude by discussing the parsing of relative clauses and constituent questions.

The parser does its work by creating a data structure called CHART. The chart contains all the objects the parser has constructed in the course of its analysis. The chart consists of VERTICES and EDGES. A vertex is a location in the input sentence. It is convenient to think of vertices as the boundaries that separate the words in the sentence. An edge is a constituent, finished or unfunished, spanning between two vertices. The diagram in (11) represents the chart in an early stage of parsing the sentence "John saw Mary". All of the edges in this chart are finished edges. The round circles are vertices, the lines are edges.

(11)



A **FINISHED EDGE** consists of the following parts:

BEG	Place where the constituent begins (= vertex right before the first word).
END	Place where the constituent ends (= vertex right after the last word).
CAT	Category of the constituent.
STR	The assembled constituent structure tree.

An **UNFINISHED EDGE** has the following parts:

BEG	Place where the constituent begins.
END	Place where the constituent ends. (Originally the same as the starting location; it changes as the constituent is assembled piece by piece.)
CAT	Category of the constituent when finished.
FOUND	List of daughter constituents that have already been discovered. (Initially the list is empty; it will become a constituent structure tree when the edge gets finished.
WANTED	List of category labels of daughter constituents that still need to be discovered. (Initially the right side of some phrase structure rule.)

The work of the parser can be divided into four independent processes: Initial Steps, Basic Procedure, Predicting, and Linking of constituents.

<u>Initial steps</u>

(i) Construct a chart using only lexical information. Make a finished edge for each grammatical category to which the word may belong. (Diagram (11) above shows the resulting chart for "John saw Mary".)

(ii) Using the procedure specified below, make a prediction that the input string constitutes a sentence. (An unfinished S edge will be added to the chart at the initial vertex.)

(iii) Keep adding new edges to the chart following the Basic Procedure below.

Basic Procedure

(i) Whenever an unfinished constituent is added to the chart, PREDICT (see just below) that it will be followed by whatever comes first on its own WANTED list. Check all the adjoining finished constituents immediately to the right of the unfinished edge to see if their category matches the first symbol on the WANTED list. If a match is found, LINK the two edges using the procedure below.

(ii) Whenever a finished constituent is added to the chart, check all the adjoining unfinished constituents immediately to the left of it to see if the first symbol on their WANTED list matches the category of the new edge. If a match is found, link the two constituents using the procedure below.

Predicting

To predict a constituent of category C at vertex V do as follows. Add to the chart an unfinished edge with the following characteristics: the edge begins and ends at V; the category of the edge is C; the FOUND list is empty and the WANTED list is the right hand side of some phrase structure rule for C. Repeat until all the PS rules for C have been used.

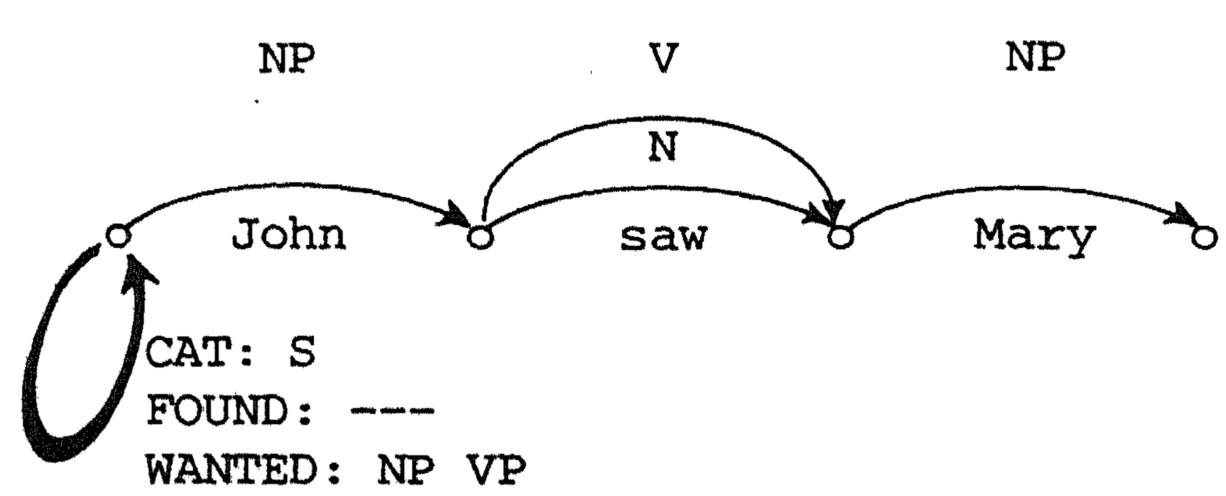
Linking constituents

To link an unfinished edge U with a finished edge F do as follows. Add to the chart a new edge N that has the following characteristics: N begins where U begins and ends where F ends; the category of N is that of U; the FOUND list of N is the same as that of U with the structure of F appended to it; the WANTED list of N is the same as that of U except that the top symbol has been removed. If the list thereby becomes empty, then make N a finished edge by adding a mother node of the appropriate category in front of the daughter nodes in the FOUND list.

In order to illustrate how the Kaplan parser works, let us continue the parse of "John saw Mary" from the stage pictured in diagram (11) above.

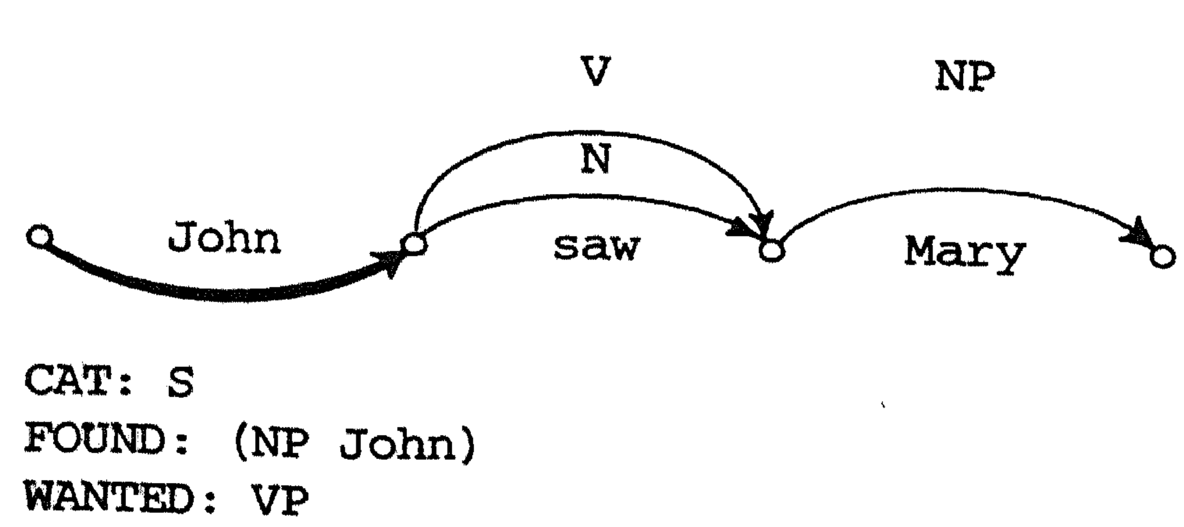
The next step is to predict that it is a sentence, that is, we add to the chart the unfinished edge shown below.

(12)



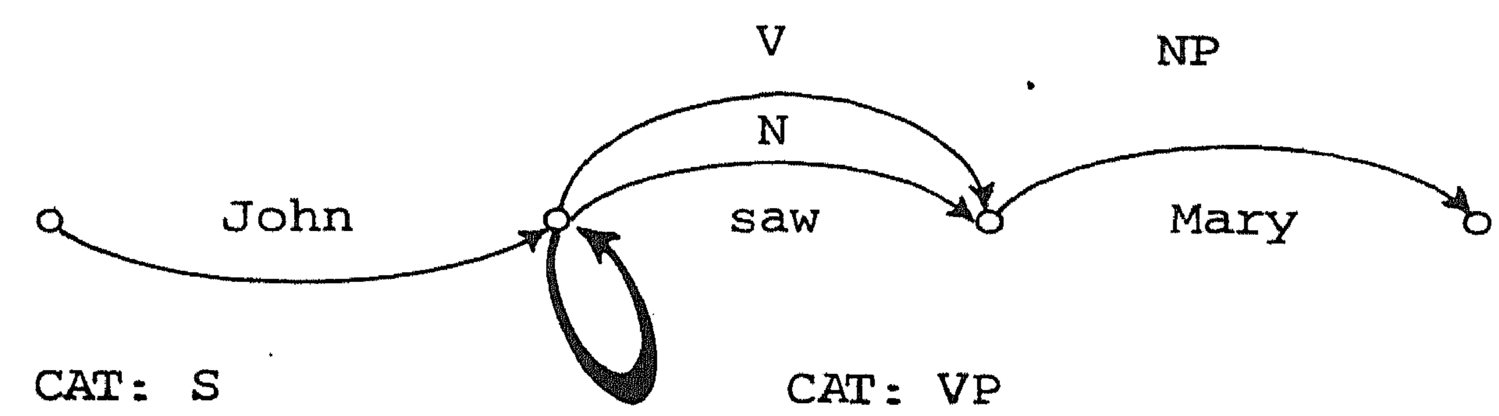
According to Basic Procedure (i), the next step is to predict an NP at the initial vertex because NP is the topmost symbol on the wanted list of the edge we just added. This would result in a number of edges being added to the chart (one for each rule of the form NP ⟶ ...), all of which are similar to the S edge above in that they loop back to the vertex from which they start. Let us overlook this step, however, and move on to the second task specified in Basic Procedure (i). This makes us check if there already is an NP immediately to the right of the unfinished S. Since there is one, the NP "John", we proceed to link the two neighboring edges. This results in the following new edge being put on the chart. (I am leaving out of the picture all edges that are no longer relevant. In the actual chart, however, no edges are ever removed.)

(13)



Since this is an unfinished constituent, Basic Procedure (i) instructs us to predict a VP at ending vertex of the new edge. Actually this would result in as many new edges being put on the chart as there are rules for expanding VPs, but let us here be content on adding just the new edge in (14).

(14)

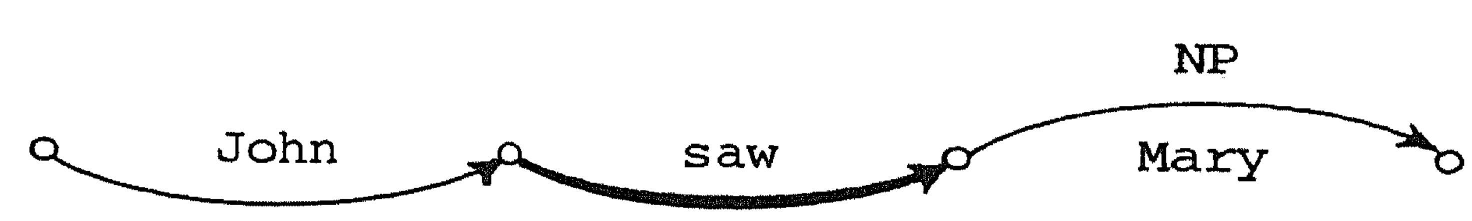


CAT: S
FOUND: (NP John)
WANTED: VP

CAT: VP
FOUND: --
WANTED: V NP

Assuming that there are no rules for expanding V, the first part of Basic Procedure (i) can be dispensed with and we look for a V immediately to the right of the new edge. Since there is one, we proceed to put the following new edge on the chart.

(15)

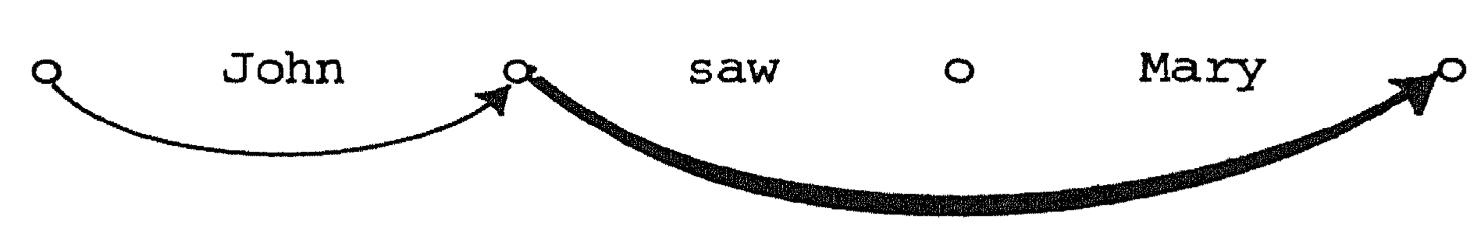


CAT: S
FOUND: (NP John)
WANTED: VP

CAT: VP
FOUND: (V saw)
WANTED: NP

Again, let us forego predicting an NP here because we can simply link the new edge with the NP Mary to the right of it. Unlike all the new edges we have added so far, the resulting new edge is a finished one.

(16)



CAT: S
FOUND: (NP John)
WANTED: VP

CAT: VP
STR: (VP (V saw) (NP Mary))

Part (ii) of the Basic Procedure tells us to check whether there is an unfinished constituent looking for a VP immediately to the left of the new edge. Since there is, we link the unfinished S constituent with the new VP. This completes the parse:

(17)

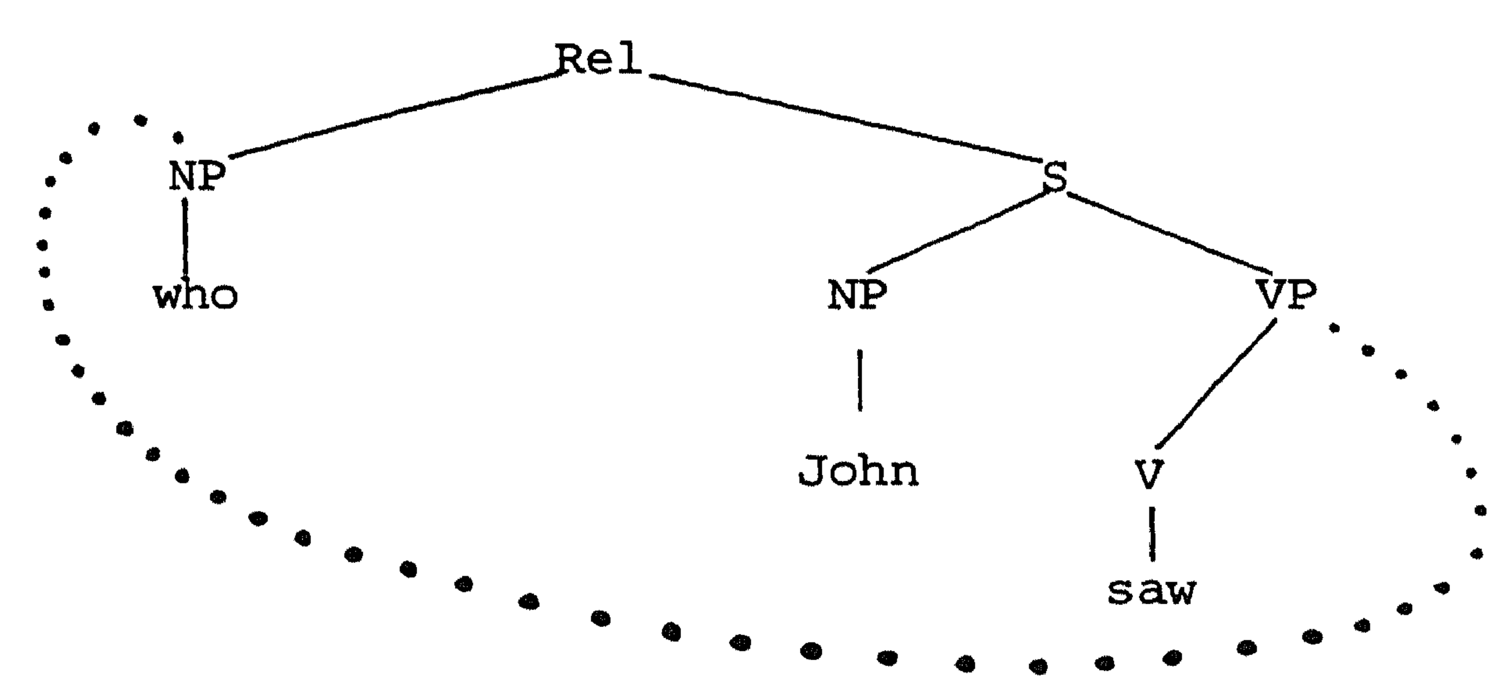
CAT: S
STR: (S (NP John) (VP (V saw) (NP Mary))).

In concentrating on the bare essentials of the parsing strategy I have neglected to mention a number of details that actually do need to be taken care of. Most importantly, at some point we need to pay attention to the fact that my phrase structure rules have a set of conditions in addition to the pure categorial part. Let us simply say that the conditions are applied at the point where a finished constituent is about to be added to the chart. If the conditions fail, the edge is rejected.

7. RELATIVE CLAUSES

Let us now consider what needs to be done to enrich the procedure I have just sketched to parse structures such as the one in (18).

(18)



This diagram represents the Texas view (first conceived by Stanley Peters) of the structure of the relative clause "who John saw". The structure consists of a dislocated element, the NP "who", and an incomplete S with a missing noun phrase. The dotted line that loops from the VP node back to the dislocated element is a way of indicating the location of the gap in the object position under the VP. It is also meant to suggest that there is a certain dependency between the gap and the dislocated element. This dependency is both syntactic and semantic. In general (especially in languages with richer systems of case marking), the form of the dislocated noun phrase in a relative clause depends on where the gap is, and the location of the gap is also the place where the meaning of the dislocated element is taken into account in computing the meaning of the relative clause.

Many theories have been built around the fact that the dislocated element in a shadowy way plays a role in a location other than in which it actually appears. There are two differences between Gazdar's and mine; one trivial, the other more substantive. Under my view there is no invisible NP or trace under the VP where the dotted line starts, but if it should turn out to be useful to have one there, I have nothing against it. The more important difference is that I regard the incomplete VP and the incomplete S that dominates it as belonging to the same syntactic category with ordinary

VPs and Ss. Gazdar does not. In other words, I prefer not to employ slash categories and derived rules. It seems to me that they are unnecessary and that one is better off without them in some respects.

To show that derived categories are not needed, I will modify the above parsing method in such a way that it can deal with incomplete constituents using only the rules that normally generate complete structures. Before going into details, let us first reflect a bit on what the dotted line in (18) is supposed to express. This line, of course, is pure fiction; there are no dotted lines in the output of the parser; but it is helpful in giving us an overview of what it is that has to be accomplished.

Two things need attention. First of all, the S node adjacent to the dislocated element must have a gap somewhere, and no more than one, except in the case of across-the-board extractions. In coordinate constructions, for example, each conjunct must have a gap if any of them do, viz. "the girl that [John loves - and Mary hates -]". There are many restrictions on where the gap cannot be, e.g. deep in another relative clause or in a determiner. Secondly, wherever the gap is, the dislocated element must be of the right sort to appear in that context. Thus the dotted line serves a dual function, one end assures the top level that there in fact is a gap somewhere in the structure and makes it possible to check their number and location; the other end ties the dislocated NP down to the lower level and subjects it to local constraints. The dotted line is, so to speak, a two-way channel on which information is passed in both directions.

As I said, the dotted line is fiction but we must find some way to pass all that information up and down the tree. Gazdar's system accomplishes this task by means of slash categories. The position behind the slash can be viewed as a memory cell. It must hold, not just a category label, as Gazdar's notation suggests, but all the other relevant information as well about the dislocated NP, such as its case and gender. What I propose to do is to dissociate this kind of information from the category labels and let the parser (or generator) store it elsewhere.

8. MODIFIED PARSER

The following changes in the parser will bring about the intended effect.

The representation of a finished constituent contains one additional part:

USED A list of dislocated constituents that were used to plug gaps in the course of constructing the edge.

Two new parts are added to the representation of an unfinished constituent:

USED A list of dislocated constituents that were used during the construction of the edge. (This list is initially empty. Contains all the elements from the USED lists of the daughter constituents.)

SHELF A list of dislocated constituents that are still available to fill gaps. (Whenever one of them gets used by a daughter constituent it becomes unavailable to others -- except in across-the-board cases.)

The procedures are modified as follows:

Initial steps -- No change

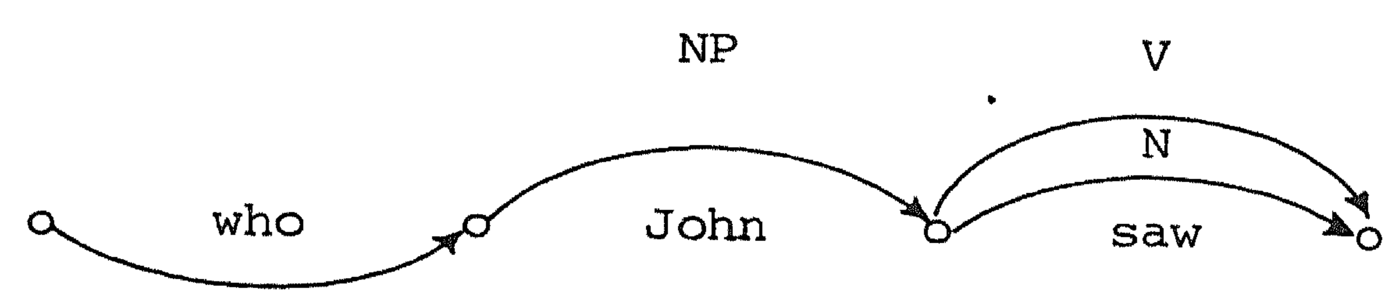
Basic procedure -- Add to the end of part (i): If the first symbol on the WANTED list matches the category of the first item on the SHELF do as follows. Construct a "phantom" edge using the item from the SHELF and put it in the USED list of the new finished edge. Let the edge begin and end where the original edge ends and let it have the category of the used constituent. Then behave as if the phantom edge really existed on the chart; that is, link it with the original edge. Part (ii) -- no change.

Predicting -- The new unfinished edge inherits its SHELF from the edge which causes the procedure to be called. The USED list is empty.

Linking constituents -- The SHELF of the new edge N is the same as that of the unfinished edge U minus any that appear on the USED list of the finished edge F. The USED list of N is obtained by appending the USED list of F to that of U.

The effect of these modifications can best be explained by means of an example. Let us suppose that the parser is working on the relative clause "who John saw" and has reached the stage shown in (19). That is, it has seen the dislocated NP "who" and has predicted that it will find next an S which lacks a noun phrase. Note that the SHELF of the unfinished S edge contains the dislocated NP.

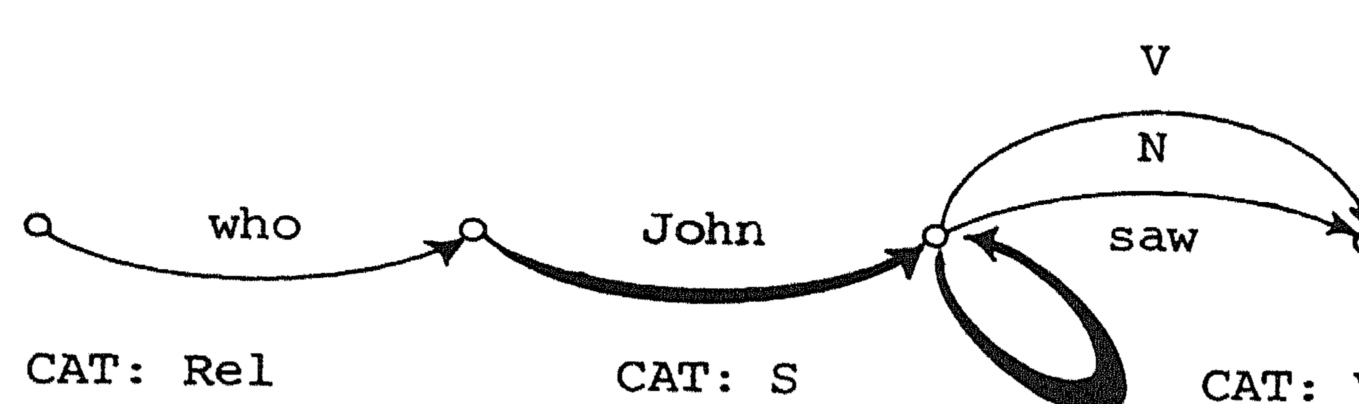
(19)



CAT: Rel	CAT: S
FOUND: (NP who)	FOUND: --
WANTED: S	WANTED: NP VP
USED: --	USED: --
SHELF: --	SHELF: (NP who)

The parser proceeds from this point on in the same manner as in the previous example. Finding the NP "John", the parser advances the unfinished S edge in the manner shown in (20) and predicts that it will next find a VP. Note that the SHELF list of the unfinished VP edge is inherited from the S edge that led to this prediction.

(20)



CAT: Rel	CAT: S	CAT: VP
FOUND: (NP who)	FOUND (NP John)	FOUND: --
WANTED: S	WANTED: VP	WANTED: V NP
USED: --	USED: --	USED: --
SHELF: --	SHELF: (NP who)	SHELF: (NP who)

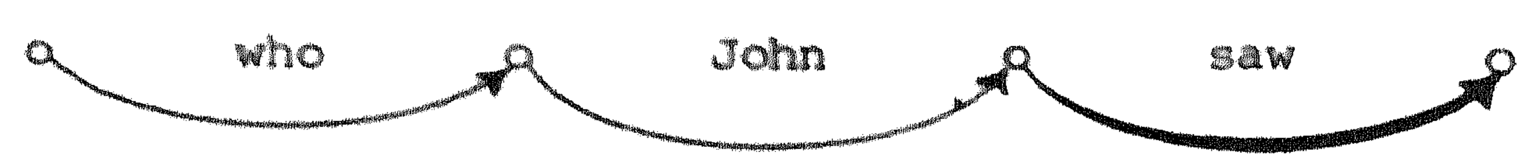
The VP edge in (20) can still be advanced but when the parser reaches the stage pictured in (21) it cannot continue any further without making use of the dislocated constituent that has been passed on from edge to edge on the SHELF list.

(21)

CAT: Rel	CAT: S	CAT: VP
FOUND: (NP who)	FOUND: (NP John)	FOUND: (V saw)
WANTED: S	WANTED: VP	WANTED: NP
USED: --	USED: --	USED: --
SHELF: --	SHELF: (NP who)	SHELF: (NP who)

The revised form of the Basic Procedure in effect enables us to trade in the item off the SHELF for the wanted NP of the unfinished VP edge. This is done by generating a phantom NP edge and linking it with the unfinished VP edge. This results in a finished VP, as shown in (22).

(22)



CAT: Rel	CAT: S	CAT: VP
FOUND: (NP who)	FOUND: (NP John)	STR: (VP (V saw))
WANTED: S	WANTED: VP	USED: (NP who)
USED: --	USED: --	
SHELF: --	SHELF: (NP who)	

The unfinished VP is now used to complete the unfinished S, as shown in (23).

(23)

CAT: Rel	CAT: S
FOUND: (NP who)	STR: (S (NP John) (VP (V saw)))
WANTED: S	USED: (NP who)
USED: --	
SHELF: --	

At this point we can see that two components can be joined to form a relative clause. The finished S does have a gap which was filled by the dislocated constituent. The method by which edges are constructed guarantees that no constituent on the SHELF can be used more than once as edges are extended. Thus we can be sure that there are no more gaps than there are dislocated phrases. (Special measures must be taken to handle coordination and other cases of across-the-board extractions.) The importance of this result lies in the fact that the incomplete sentence was parsed without using any special slash categories and extra rules. The inevitable bookkeeping was done by the processing component without cluttering the grammar.

9. CONCLUSION

I have left a number of loose ends which I will not have time to tie together. For example, I have not said anything about how to modify the phrase structure rule

(24) Rel → NP S

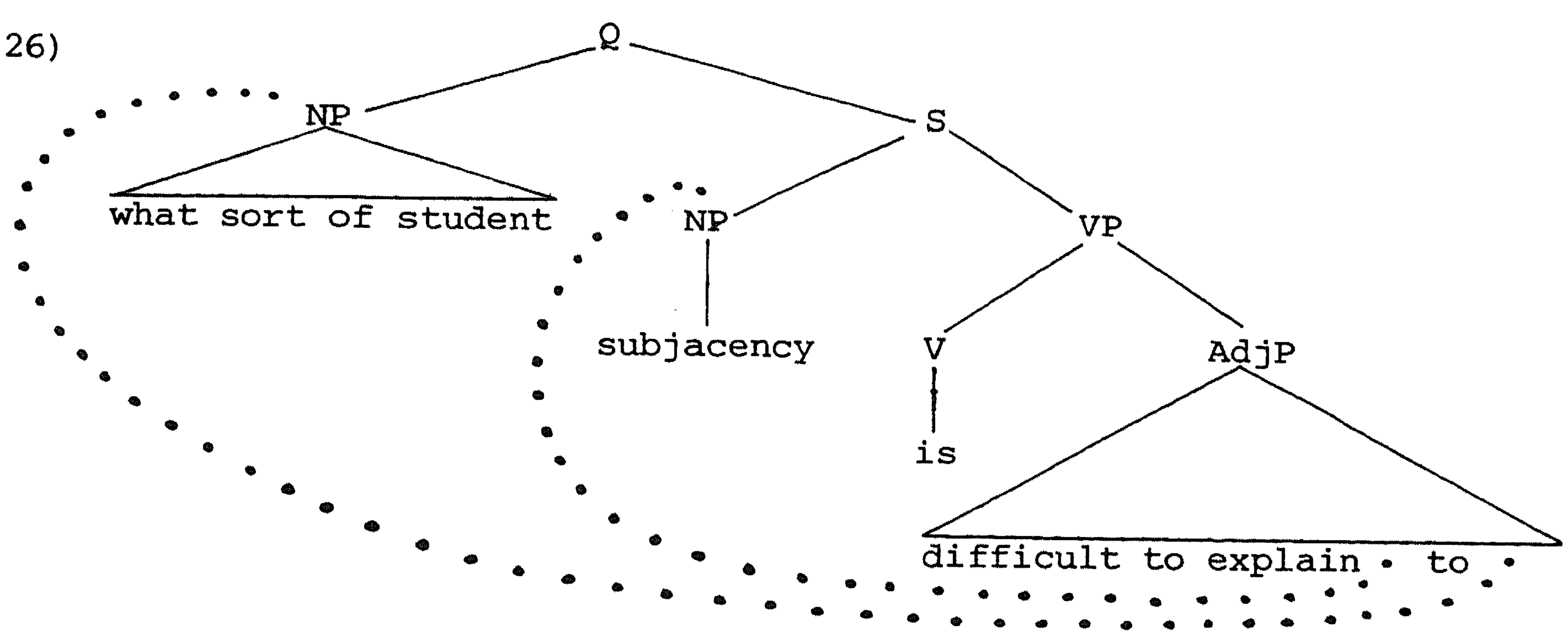
to tell the parser that the S must have a gap in it. Any number of conventions and principles obviously could do the job, but I will not pursue them here. I will instead conclude with a few observations about the advantages that this method of dealing with gaps has when compared to Gazdar's approach.

First of all, it is more flexible. By allowing only one symbol to appear behind the slash, Gazdar's system incorporates a constraint that in effect says that a node cannot have more than one missing daughter constituent. There is no such constraint in my framework, although it could easily be built in, if needed. This added flexibility, or "descriptive power", of my system of course is a vice rather than a virtue unless it can be shown that one cannot describe natural languages without it. As many people have pointed out, there are in fact a number of constructions in English where the added power appears necessary. For example, one can question out of tough-movement complements:

(25) What sort of student
is subjacency difficult to explain __ to __ ?

The structure of the embedded question is shown in (26).

(26)



One could of course easily modify Gazdar's system to allow more than one symbol to appear behind the slash, e.g. in place of the VP node in (26) one would have "VP/NP NP", but this solution raises a number of further problems. In cases like (26) English seems to require that the dependencies be nested. In my terms, this means that the two dotted lines are not supposed to cross one another. In the parsing procedure that I just described, this geometric property translates to the requirement that the items on the shelf should be taken off in the opposite order from which they were put in. In other words, the shelf should be treated as a pushdown stack. In Scandinavian languages it is not only possible to have double or even triple dependencies but they can be either nesting or intersecting. In the case of crossing dependencies, a resumptive pronoun rather than a gap typically appears. The example in (27) comes from Engdahl.

(27) Det här problemet
minns jag inte vilka metoder läraren löste det med __.

"This problem I don't remember which methods the teacher
solved it with __."

In this sentence there are two dislocated elements, "this problem" and "which methods", but only one gap. The gap marks the place of the most recent dislocated element, i.e. "which methods". The place of the more distant element, "the problem", must be indicated by the pronoun "it". The rules governing the use of resumptive pronouns in Swedish are more complicated than this single example shows; nevertheless, it is easy to see how the parsing strategy that I outlined can be modified to deal with this type of problem. For example, the parser could keep a list of potential resumptive pronouns. This list would be consulted when the parser encounters a situation where a clause next to a dislocated element has been successfully parsed and the element itself is still on top of the stack. As a last resort, the stack could be popped at this point on the condition that a matching resumptive pronoun is available.

Although we don't presently have a real solution for problems of this kind, I think they bring into focus the difference between Gazdar's system and the Texas approach to unbounded dependencies. It is undoubtedly possible for Gazdar to duplicate the effect of my suggestion by allowing strings of symbols to appear in category labels behind the slash and by writing rules that generate either a gap or a resumptive pronoun depending on whether the new element corresponds to the first or some subsequent element behind the slash in the category label of its mother node. While this is possible, I think it would be misguided.

First of all, as the number of categories and rules increases the system becomes more and more ineffective from a processing point of view without any compensating descriptive benefit. Secondly, the problems posed by gaps and resumptive pronouns are a part of a larger puzzle involving other kinds of anaphora. It is unlikely that they will be solved without a comprehensive theory of how pronouns are associated with discourse referents. If Gazdar's system is extended in the proposed way, the category labels of a phrase structure tree are in effect used as memory cells to keep track of discourse referents. My guess is that this task is best accomplished by storing the information elsewhere.

Of course, at this point this is merely a conjecture. As things stand now, there is little to distinguish between the two alternatives. Gazdar's system has nice consequences with respect to across-the-board extractions (GAZDAR 1979c). For example, it correctly predicts that, in a coordinate structure, no conjunct can be incomplete unless the others are incomplete in just the same way. This result directly follows from the fact that A, A/B, and A/C are all regarded as distinct grammatical categories. It takes an extra statement to achieve the same effect in the Texas system. On the other hand, Gazdar's system runs into semantic difficulties in cases whenever the meaning of a dislocated element cannot legally be entered on a lower level by means of lambda conversion. Topicalized reflexives, such as "Himself John is mad at __" are a case in point. For the Texas approach sentences of this kind pose no problem because the meaning of "himself" can be taken into account at the point where a dislocated element is used to fill the gap. This insures that the reflexive becomes properly bound. It seems likely that other differences will emerge as the two approaches are explored further.

REFERENCES

BEAR, JOHN & LAURI KARTTUNEN, 1979, 'PSG: A simple phrase structure parser', *Texas Linguistic Forum* 15, 1-46.

BRESNAN, JOAN & RONALD KAPLAN, 1979, *A Formal System for Grammatical Representation,* manuscript.

ENGDAHL, ELISABET, 1980, *Unbounded Dependencies in Swedish,* unpublished paper, University of Massachusetts.

GAZDAR, GERALD, 1979a, *Constituent Structure,* unpublished paper, University of Sussex.

GAZDAR, GERALD, 1979b, *English as a Context-Free Language,* unpublished paper, University of Sussex.

GAZDAR, GERALD, 1979c, *Unbounded Dependencies and Coordinate Structure,* unpublished paper, University of Sussex.

KAPLAN, RONALD, 1973, 'A General Syntactic Processor', in: R. Rustin (ed.) *Natural Language Processing,* Algorithmics Press, New York, 1973.

MALING, JOAN & ANNIE ZAENEN, 1980, *Notes on Base-Generation and Unbounded Dependencies,* unpublished paper.

PULLUM, G.K. & D. WILSON, 1977, 'Autonomous Syntax and the Analysis of Auxiliaries', *Language* 53, 1-37.

SAENZ, RICHARD, 1980, *Local Grammar,* unpublished paper.

SAG, IVAN, 1980, *A Semantic Theory of 'NP-Movement' Dependencies,* unpublished paper, Stanford University.

A BOOLEAN APPROACH TO SEMANTICS

by

Edward L. Keenan

0. INTRODUCTION

By a boolean approach to model theoretic semantics I intend one in which for each model M and each category C of expression in the language, the set of possible denotations of expressions in C (relative to M) is not merely some set T_C defined in terms of M, but is rather a set on which are defined boolean operations and a boolean relation. That is, T_C, *the type for* C (relative to M), is a boolean algebra. Such an approach is compatible with all model theoretic approaches, such as Montague Grammar, and is exemplified in *Logical Types for Natural Language* (KEENAN & FALTZ, 1978/81), henceforth LT.

The purpose of this paper is to present some of the advantages of formulating natural language semantics in this way, irrespective of what other model theoretic apparatus is used. Section 1 below presents some basic concepts of boolean algebra, and Sections 2 - 4 the advantages: 2: simplifying the ontology implicit in the model, and a suggestion for a new approach to intensional properties; 3: extending the class of expressions within a category which can be *directly* interpreted, and a consequent new approach to presupposition; and 4: enriching the class of categories which are treated in the logic, and the consequent possibility of stating universal constraints on the logical form of natural languages which are not apparent (though not necessarily unstatable) on non-boolean approaches.

Finally, this paper supports a further very general claim. First, I note without argument that compared to many commonly studied algebras such as groups, lattices, and rings, boolean algebras possess a particularly rich structure, sufficiently much that it is surprising that any category of natural language is semantically boolean. Second, this paper, and in much more detail, LT(78/81) show that very many categories of natural language are semantically boolean, so the boolean nature of natural language is not category specific. And this suggests, as BOOLE (1847) felt, that the boolean operations represent "laws of thought", properties of the way we understand the world.

1. BOOLEAN ALGEBRA

In general, an algebraic structure is a non-empty set, the *domain* of the algebra, on which are defined various operations (functions) and relations satisfying certain conditions (axioms). I shall first present a familiar example of a boolean algebra and then give the general definition.

Consider as the domain of a boolean algebra the power set of a non-empty set X, that is the set whose members are just the subsets of X. Denote this set P(X). It has at least two members, $\emptyset$ (the empty set) and X itself (taken to be non-empty). And for any $A \in P(X)$ we have that $\emptyset$ is a subset ($\subseteq$) of A and $A \subseteq X$. In this sense $\emptyset$ is the least or zero (0) element of the domain and X itself the greatest or unit (1) element. Further, for all A and B in P(X) we have that $A \cap B$ and $A \cup B$ are subsets of X and thus in P(X). So intersection and union are binary operations defined on P(X). And they have certain characteristic properties, e.g. they are *commutative,* $A \cap B = B \cap A$, and ditto for unions; and each *distributes* with respect to the other, e.g. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, and ditto interchanging the intersection and union symbols. Moreover, $A \cap B$ is a *lower bound* for the set {A,B} in that it is a subset of each member of {A,B}. In fact, it is the *greatest lower bound* in that for all $Y \in P(X)$, if Y is a subset of each of A and B then Y is also a subset of $A \cap B$. Similarly, $A \cup B$ is the least upper bound for {A,B}. (Each of A and B is a subset of $A \cup B$ and if that also holds for Y then $A \cup B$ is a subset of Y.) Note further that $A \subseteq B$ iff $A \cap B = A$. So, perhaps perversely, we could actually define the subset relation in this way.

Finally, we define on P(X) a one place operation called (absolute) complement and denoted ' as follows: For all $A \in P(X)$, A' is the set of those objects in X which are not in A. Clearly $A' \subseteq X$ so $A' \in P(X)$. And note the following obvious truths, for all $A \in P(X)$: $A \cap A' = \emptyset$, $A \cup A' = X$, $A \cap \emptyset = \emptyset$, and $A \cup X = X$.

More generally we define a boolean algebra $\mathcal{B}$ to be a (horrendous) 7-tuple consisting of a non-empty set B (the domain), two elements $0_{\mathcal{B}}$ and $1_{\mathcal{B}}$ of B called the zero and unit elements, two binary operations on B,

$\wedge_{\mathcal{B}}$ and $\vee_{\mathcal{B}}$ called *meet* and *join*, respectively, one unary operation $'_{\mathcal{B}}$ called *complement*, and a binary relation $\leq_{\mathcal{B}}$ called *less than* or *equals*. $\mathcal{B}$ is required to satisfy the following axioms (omitting here and elsewhere the subscripts) for all x,y,z in B:

(a) Commutativity

$(x \wedge y) = (y \wedge x)$ and $(x \vee y) = (y \vee x)$.

(b) Distributivity

$(x \wedge (y \vee z)) = (x \wedge y) \vee (x \wedge z)$ and $(x \vee (y \wedge z)) = (x \vee y) \wedge (x \vee z))$.

(c) Laws of Complements

$(x \wedge x') = 0$ and $(x \vee x') = 1$.

(d) Laws of Zero and Unit

$(x \wedge 0) = 0$ and $(x \vee 1) = 1$.

(e) $0 \neq 1$.

(f) $x \leq y$ iff $(x \wedge y) = x$.

This axiomatization (adapted from MENDELSON, 1970) is quite redundant. (f) is just a definition of $\leq$ and either $\wedge$ or $\vee$ could be defined in terms of the other plus complement since the De Morgan laws (e.g. $(x \wedge y)' = (x' \vee y'))$ follow from the axioms. It also follows that $\leq$ behaves like the subset relation, e.g. $x \leq x$, if $x \leq y$ and $y \leq z$ then $x \leq z$, and if $x \leq y$ and $y \leq x$ then $x = y$. And it follows that $x \wedge y$ is the greatest lower bound for $\{x,y\}$ and $x \vee y$ is its least upper bound.

Given the definition we could now prove that for any non-empty set X, $\langle P(X), \cap, \cup, ', \emptyset, X, \subseteq \rangle$ is a boolean algebra (called a power set algebra). Another useful boolean algebra is $\langle 2, \wedge, \vee, ', f, t, \leq \rangle$ where 2 is the set $\{t,f\}$ of truth values and $\wedge$, $\vee$, and $'$ are defined by the standard truth tables for conjunction, disjunction, and negation, respectively. It follows that $t \leq t$, $f \leq t$, $f \leq f$, and $t \not\leq f$. A further example, useful in intensional logic, is the entire set of functions from a non-empty set J (say a set of possible worlds) into B, the domain of some boolean algebra. Denote this set $F_{B/J}$. The boolean operations are *defined pointwise* on J, by which is meant that for f and g any functions in the set, $(f \wedge g)$ is that function which maps each j in J onto $f(j) \wedge g(j)$, where the meet sign on the right refers to meets in B, since f(j) and g(j) are elements of B.

Similarly $(f')(j) =_{df} (f(j))'$. The zero function maps each j onto the 0 element of B (we write, omitting subscripts, $0(j) = 0$) and similarly $1(j) = 1$. It follows that $f \leq g$ iff for all j in J, $f(j) \leq g(j)$. Note that J can be any set, in particular a set of n-tuples of places, times, etc.

For purposes of Section 2 (which is self contained and may be omitted on first reading as it is somewhat more algebraic than the other sections) two further concepts of boolean algebra will be needed. First, an algebra is said to be *complete* if for every subset K of its domain B there is an element x in B which is the greatest lower bound for K (i.e. $x \leq$ everything in K, and if $y \in B$ meets this condition then $y \leq x$). For a complete algebra the element x referred to above is denoted $\bigwedge K$. Analogously for $\bigvee K$. Note that power set algebras are complete. For $K \subseteq P(X)$, $\bigwedge K$ is just $\bigcap K$, the set whose members are just those elements in all of the k in K.

Second, an *atom* of an algebra is an element of the domain which is not the 0 element but which no other element is strictly less than. That is, b is an *atom* in B iff for all x in B, if $x \leq b$ then $x = 0$ or $x = b$. And an algebra is said to be *atomic* if for all non-zero elements y in the domain there is at least one atom b such that $b \leq y$. Any power set algebra is atomic, the unit sets being the atoms. (Similarly the algebra 2 is complete and atomic, and the pointwise algebra $F_{B/J}$ is complete and atomic whenever B is the domain of a complete and atomic algebra.)

Finally, I note from standard boolean algebra that any complete and atomic algebra is isomorphic to a power set algebra (the function mapping an element onto the set of atoms $\leq$ to it is an isomorphism of the algebra onto the power set of its atoms). So up to isomorphism the complete and atomic algebras are the power set algebras.

2. SIMPLIFYING THE ONTOLOGY

2.1. One of MONTAGUE's (e.g. 1972) linguistically most useful insights was his definition of the types (relative to a model) which enabled us to treat singular terms - *John, the king of France*, etc. - as taking their denotations in the same type as quantified NPs like *every man*. Thus the logical forms for *John walks* and *every man walks* are identical up to the difference in internal structure of the logical forms for *John* and *every man*. Thus we have a better explanation than in standard logic (SL) for how we interpret English sentences as a function of their form; in SL *every man* is not even assigned a logical form at all.

Let us review briefly the essence of Montague's innovation here. First, in SL a model (the ontological primitives) is a pair <2,U>, where 2 is the set of truth values and U is the set of "things that exist". In both cases we have a reasonable pretheoretical idea what the elements of these sets are intended to be. In particular (ignoring 2 as it is "standard") U is the type for the individual constants (≈ proper nouns) and the range of the individual variables, so it may contain things like you, and me, and John, etc. While we might debate whether U and 2 should be taken as *primitives* it is clear that these sets do comprise parts of our ontology and it is not surprising that other categories should have their types given as a function of these. For example, the type for the 1-place predicates in SL will just be the subsets of U, i.e. P(U), or equivalently the set of functions from U into 2. Let us call these sets (or functions) *extensional properties*.

Montague's innovation here was to treat the denotations of proper nouns (PNs) not as elements of U, but rather as sets of extensional properties, called for the nonce *individuals*. More formally, for all $x \in U$ we define the *individual determined by* x to be $\{K \subseteq U: x \in K\}$. So an individual is defined in terms of an element of U. And being a collection of sets of properties (I drop "extensional"), an individual is a subset of P(U), that is a member of P(P(U)). So the type for full NPs like *John* and *every man* is P(P(U)) and that for common noun phrases (CNPs) is P(U). And *every man* will be interpreted as the intersection of the individuals which have the *man* property; *a man* as their union, etc. and *John, every man, a man,* etc. may be treated as expressions in the same gross category. (I say 'gross' because PNs will not be interpreted as arbitrary members of T_{NP} but may only have individuals as their possible denotations, so they are in effect a distinguished subcategory of NP. Note that an intersection (union) of distinct individuals is never an individual.)

But the nice ontological character of SL has been lost, for a model of this system is still a pair <2,U> and while 2 is still the truth values, the elements of U are now not possible denotations for any expressions in English. Hence we have no pretheoretical idea what the elements of U are, and it is mysterious why other expressions, e.g. *John, every man,* and indeed most other expressions, should have their denotations given as a function of elements of U. Notice that the "things that exist", e.g. denotations of PNs, are the individuals, not elements of U. So U seems to be a kind of noumenal world underlying the phenomenological world of individuals.

And our ontological qualms are not assuaged in the slightest by noting that U and the set of individuals are in a natural one-to-one correspondence (so are the even numbers and the odd numbers, but they have very different properties). Such a correspondence (onto) just says the two sets have the same size. But they crucially fail to have the same structure. For individuals, being sets, are the kinds of things we can take boolean combinations of, e.g. intersections, unions, etc. And this we must do (regardless of how the statement is actually formulated) in order to get denotations for *every man, a man,* etc. Moreover, there is no way to assign a boolean structure to an abitrarily chosen set U. For example, all finite boolean algebras have 2^n elements for some finite n. So there are no boolean algebras with 3, or 6, or 7 elements. In fact, closing the individuals under arbitrary intersections and unions gives us P(P(U)), the set of all the sets of properties, much larger than U.

From this point of view part of Montague's innovation lay in trading in the elements of U for things which we can treat in a boolean way. And once this is recognized there is a very easy way to extend his insight so as to eliminate the ontological qualms above. This extension not only yields a new ontology, but it generalizes in ways that permit a new and potentially more adequate approach to the treatment of "hyper-intensional" CNPs, e.g. *imaginary horse, book that John intended to write but never wrote,* etc.

As a first step in the extension notice that if we take T_{CNP} as P(U) we may automatically regard it as a complete and atomic (ca) algebra its atoms being the unit sets of elements of U. And the individual determined by x, namely $\{K \subseteq U: x \in K\}$ is $\{K \in P(U): \{x\} \subseteq K\}$, as trivially $\{x\} \subseteq K$ iff $x \in K$. But $\{x\}$ is an atom of P(U) and $\subseteq$ is just the boolean relation on P(U), so this last set basically defines an individual in terms of the boolean structure of $P(U) = T_{CNP}$. So our first step is the following preliminary definitions:

PRELIMINARY DEFINITION 1. A *model for* L is a pair <2,P>, where 2 is as before and P is any complete and atomic boolean algebra.

PRELIMINARY DEFINITION 2. For each atom b in P, I_b, *the individual determined by* b, $=_{df}$ $\{p \in P: b \leq p\}$.

REMARKS on the preliminary definitions:

1) They do appear to constitute a *new* ontology, for now (extensional) properties and truth values are the ontological primitives, not entities or individuals and truth values.
2) There appear to be no mysteries in the ontology since each primitive is the type for some category of expression. In particular, the elements of P are the kinds of things that expressions like *man, tall man, man who Mary loves,* etc. can denote. And while we shall want to query further what their exact nature is, we will know how to reason about them since we know what ordinary expressions they are the intended interpretations of. And
3) up to isomorphism, the class of possible types for CNP and hence of individuals is the same as on Montague's earlier view. For obviously if T_{CNP} is P(U) for some U then it is a ca-algebra and the above definition picks out the same sets as individuals as the earlier definition. And if P is not specifically a power set algebra it is, by the remark at the end of Section 1 isomorphic to the power set of its atoms, so taking the set of atoms as U we have a type for CNP in the old sense, one that is isomorphic to the given one. Thus any arguments which would be shown valid on the new approach are valid on the old and vice versa. So the two approaches are descriptively adequate to the same extent.

Let us consider some objections to this approach, ones that will revise somewhat our preliminary definitions. First it has been objected that this approach is just a "mathematical trick". But that is silly. It is neither more nor less mathematical or tricky than Montague's observation that the elements of U are in a one-to-one correspondence to the individuals as defined on that approach.

More seriously however one can query whether this approach really constitutes a *new* ontology or whether it just gives us the same one in different mathematical garb. To be more precise: What motivation do we have for taking T_{CNP} (= P) as a ca-algebra other than our desire to treat it as we always did, namely as a power set algebra? And second, while P itself may not generally be mysterious, individuals are defined above in terms of atoms, and are not these properties every bit as mysterious as the elements of U on the old approach?

I will answer these objections as follows. First I will show that we have independent motivation, in terms of correctly representing our judgments of logical truth and entailment on English, for taking P as a complete boolean algebra. Similarly I will show that, taking NP denotations as subsets of P, we have direct motivation for requiring that PN denotations be subsets of P which meet certain conditions, and that when we define individuals as the subsets of P which meet those conditions we get the individuals in the old sense. Neither the judgments of validity and entailment nor the formal conditions mention or in any way presuppose the notion of an atom or that P is atomic. Thus the notion of individual is conceptually and formally independent of that of an atom. And third I will show that there is independent motivation for requiring that P have atoms; specifically that to correctly represent the valid arguments on English there are property denoting expressions which should be and are intended to be interpreted as atoms. So atoms are not particularly mysterious.

But this is as far as the independent motivation for the boolean nature of P will go. And if we merely require that P be complete and have atoms but not be atomic (which would require in addition that for every non-zero q in P there is an atom $b \leq q$) we obtain a properly larger class of T_{CNP}'s than in the extensional systems of Montague or LT, and this larger class is rich enough to provide denotations for the hyperintensional CNPs mentioned above. So in fact what appears to be the descriptively most adequate approach here does not exactly reconstruct the systems of PTQ or LT. If of course we impose the additional requirement that P be atomic we do obtain the earlier systems.

2.2. T_{CNP} should be a boolean algebra

Our intent is that P (= T_{CNP}) provide denotations for expressions like *man, socialist, vegetarian,* etc. among others. We assume that full NPs like *John* and *every man* will be interpreted as subsets of P. So we want our semantics to guarantee that sentences like *John is a socialist* are true in M = <2,P> iff the subset of P which interprets *John* has the property which interprets *socialist* as a member. Now consider sentences like (1):

(1) John is both a socialist and a vegetarian.

We want (1) to be true in <2,P> iff the John set of properties has the property of "being both a socialist and a vegetarian". But which property is that? Clearly it is not arbitrary relative to which elements of P interpret *socialist* and *vegetarian* (call them p and q, respectively). Arguably the property should be $(p \wedge q)$; What is the argument? Well, one argument is that since meet is commutative and thus $(p \wedge q)$ is the same element of P as $(q \wedge p)$, this analysis predicts that (1) should be logically equivalent to *John is both a vegetarian and a socialist* since our semantics will say in each case that the same property is in the John set. And this is pretheoretically judged correct. A similar argument shows that $(p \vee q)$ should be the property of "being either a socialist or a vegetarian".

These two claims jointly make more (correct) predictions. Since e.g. meets distribute over joins we have that $(p \wedge (q \vee r))$ is the same element of P as $(p \wedge q) \vee (p \wedge r)$. Thus (2a) and (2b) should be judged logically equivalent, and they are.

(2) a. John is both a socialist and either a vegetarian or a cannibal
b. John is either a socialist and a vegetarian or a socialist and a cannibal.

Further, judgments on similar sentences force constraints on what sets of properties are admissible interpretations for PNs like *John*. Thus (3a) and (3b) are judged logically equivalent.

(3) a. John is both a socialist and a vegetarian
b. John is a socialist and John is a vegetarian.

Thus John must be a set of properties such that for all p,q in P, both p and q are in John iff $(p \wedge q)$ is. Note that many reasonable sets fail this condition (which we call *strongly closed under meets*). E.g. replacing *John* in (3) everywhere by *no student*, it is obvious that (3a) does not entail (3b). So $(p \wedge q)$ can be in no student without it necessarily being so that both p is in it and q is in it. So John cannot be interpreted by the kinds of property sets which interpret no student.

Similarly replacing *and* in (3) everywhere by *or* we can infer that whenever $(p \vee q)$ is in John then either p is or q is, and conversely, since the a- and b-sentences are again judged logically equivalent. And again many reasonable sets of properties, such as those denoted by *every student*, need not meet these conditions, since logical equivalence fails in (3) if *John* is everywhere replaced by *every student* (and *and* by *or*).

Now consider the trickier case of complements. Our pretheoretical judgments (accepting the two valued nature of the system, something which is easily modified on a boolean approach but not something I am modifying here) tell us that in any state of affairs exactly one of (4a) and (4b) below are true:

(4) a. John is a man

b. John is not a man.

So the conjunction of these two sentences must be logically false and their disjunction logically true. All these judgments are correctly predicted if the property of "being not a man" is taken as the complement of the property "being a man" and we require of any possible PN denotation I that for any property p, I contains exactly one of $\{p, p'\}$. I shall then impose these conditions in order to correctly represent these judgments.

There is also direct motivation of a different sort for taking T_{CNP} as a boolean algebra. Consider the logical properties of extensional adjectives (APs) like *female* and *tall*. Such APs will be interpreted by functions from P into P, and if *female* is such a function f and *socialist* is a property p, then *female socialist* will be interpreted by f(p), the value of f at p. But the property f(p) is not arbitrarily related to p. E.g. *Mary is a female socialist* entails *Mary is a socialist*. So we will want to require that the functions f which can interpret extensional APs meet the condition that for all p in P, $f(p) \leq p$. That is, $f(p) \wedge p = f(p)$ so "being both a female socialist and a socialist" is not different from "being a female socialist". And the requirement that PN denotations be strongly closed under meets guarantees that whenever Mary has the property f(p) then she has the property p. For f(p) is $f(p) \wedge p$ and by strong closure we infer that Mary has p.

So another reason for taking T_{CNP} as boolean is that we want to use the boolean $\leq$ relation on T_{CNP} to correctly characterize certain valid arguments involving extensional APs.

<u>INTERIM CONCLUSION</u>: If we take T_{CNP} as a boolean algebra and constrain the interpretations of PNs in the ways indicated we correctly represent many valid arguments and logical truths (assuming the basic two valued nature of the system). So we have independent motivation for taking T_{CNP} as boolean; and we have not covertly relied on any notion of an atom nor have we in any way assumed that T_{CNP} will be isomorphic to a power set algebra.

2.3. T_{CNP} (=P) is a complete algebra

Consider the following valid argument: *Mary is taller than every man, John is a man*; therefore *Mary is taller than John*. Assuming the analysis so far, what property must Mary have above for the argument to be valid? Well, letting M be the set of PN denotations with the *man* property, and for each m_i in M letting tm_i be the property of "being taller than m_i", we want Mary's property to be: $tm_1 \wedge tm_2 \wedge \ldots$ for each m_i in M. But this is just what is meant by $\bigwedge\{tm_i : m_i \in M\}$. So if we take P as complete we will have denotations for property denoting expressions like *taller than every man*. And if we require of PN denotations that they be strongly closed under arbitrary meets, not just the binary ones mentioned earlier, then the above argument is shown valid. So I shall take P as complete, and define individuals (PN denotations) by:

DEFINITION 1. For P any complete boolean algebra, I is an *individual*[1] on P iff I is a subset of P satisfying (i) - (iii) below:

(i) *Completeness*: for all p in P, either $p \in I$ or $p' \in I$.

(ii) *Consistency* : for all $p \in P$, not both $p \in I$ and $p' \in I$.

(iii) *Meets* : for all $K \subseteq P$, $K \subseteq I$ iff $\bigwedge K \in I$.

As Definition 1 does not mention the notion of an atom, or even require that P have atoms, it is clear that individuals are conceptually and formally independent of that of atoms. Theorem 1 below may then seem surprising:

THEOREM 1. I is an individual on P iff for some atom $b \in P$, $I = \{q \in P: b \leq q\}$.

PROOF. Suppose that I is an individual on P. We show that $\bigwedge I$ is an atom and that $I = \{q: \bigwedge I \leq q\}$, thus proving the first half of the theorem.

(a) 0 (the zero element of P) is not in I. Otherwise, since $0 = 0 \wedge 1$ we have that $0 \wedge 1 \in I$, so from (iii) $1 \in I$. But $1 = 0'$, so both 0 and 0' are in I, contradicting (ii). So $0 \notin I$ (and by (i), 1 is in I, so I is not empty).

(b) If $p \in I$ and $p \leq q$ then $q \in I$. By assumption $p = (p \wedge q)$, so $(p \wedge q) \in I$, so by (iii) $q \in I$. (More exactly: $\{p,q\} \subseteq I$, whence $q \in I$.)

(c) $\{q: \bigwedge I \leq q\} \subseteq I$. By (iii) $\bigwedge I \in I$ (since $I \subseteq I$), so from (b) if $\bigwedge I \leq q$ then $q \in I$.

(d) $I \subseteq \{q: \bigwedge I \leq q\}$. Let $p \in I$. By definition of $\bigwedge$, $\bigwedge I \leq p$.

(e) $I = \{q: \bigwedge I \in q\}$. Immediate from (c) and (d).

(f) $\bigwedge I$ is an atom. Suppose otherwise. Then from definition of atom, either $\bigwedge I$ is 0 or there is a non-zero $p < \bigwedge I$. $\bigwedge I$ is not 0 since $0 \notin I$ and from (iii) $\bigwedge I \in I$. So let p such that $0 < p < \bigwedge I$. Then $\bigwedge I \nleq p$ so from (d) $p \notin I$. But also p' is not in I. For otherwise from (d) $\bigwedge I \leq p'$, whence by transitivity of $\leq$, $p \leq p'$, thus $p \wedge p' = p$. But $p \wedge p' = 0$, contradicting that $p \neq 0$. So p' is not in I. So neither p nor p' are in I, contradicting the assumption that I is an individual. Thus $\bigwedge I$ is an atom.[3]

The other half of the theorem is straightforward. Thus, assume that b is an atom of P and let $M = \{q: b \leq q\}$. We show M is an individual:

(a) *Meets*: first, let $K \subseteq M$. So each k in K is in M, so $b \leq$ each such K. So b is a lower bound for K. But $\bigwedge K$ is the greatest lower bound, so $b \leq \bigwedge K$, thus $\bigwedge K$ is in M.
Second, suppose $\bigwedge K \in M$. So $b \leq \bigwedge K$, and since $\bigwedge K \leq k$, all k in K, we have by transitivity of $\leq$ that $b \leq k$ all k in K, so all k in K are in M, so $K \subseteq M$.
Thus $K \subseteq M$ iff $\bigwedge K \in M$.

(b) *Consistency*: suppose both p and p' in M. Then $b \leq p$ and $b \leq p'$, so $b \leq (p \wedge p') = 0$, contradicting that b is an atom.

(c) *Completeness*: suppose $b \nleq p$. Then $(b \wedge p') \neq 0$. Since b is an atom, then $(b \wedge p') = b$, so $b \leq p'$. So for any p, $b \leq p$ or $b \leq p'$, so either p is in M or p' is. □

Theorem 1 together with Definition 1 tell us that *if* we take P as atomic (and complete of course) then the individuals will be just the sets they were on the earlier definition. So if $P = P(U)$ for some U, the individuals are just the subsets K of U which contain a fixed element of U. However, neither Definition 1 nor Theorem 1 presuppose the existence of atoms, much less that P is atomic and thus (isomorphic to) a power set algebra. If P in Theorem 1 were selected as complete and atomless (there are such algebras) then there would be no individuals on P. Moreover, if P were selected to have atoms but still not be atomic then the individuals on P would still be just the subsets of P which dominate a fixed atom, and thus still have all the properties guaranteed by the definition of individual. Note further (see KEENAN (to appear b) for a proof) that from standard boolean algebra we have that for any cardinal n there are non-atomic complete algebras with exactly n atoms. So we may have as many ordinary individuals as we like without requiring that P be atomic. So the atomicity of P, which is forced in e.g. PTQ and LT, remains an open question.

We do however want to require that P have at least some atoms. There are at least two reasons for this. First, if P has no atoms then by Theorem 1 it has no subsets which meet the condition for being an individual. But we want such subsets in order to provide interpretations for PNs like *John* so that the logical truths and valid arguments mentioned earlier can be shown to be valid.

And second, it follows from Definition 1 and Theorem 1 that there is a one-to-one correspondence between the atoms of P and the individuals on P. (No individual can contain two different atoms, for then it contains their meet, which is 0, and no individual contains 0 by (a) above). So an atom is an extensional property[2] which exactly one individual has. And there are many property denoting expressions in English intended to be interpreted by such properties: *tallest student, first* (*third,* etc.) *man to set foot on the moon, student who stood exactly here at exactly noon yesterday, man who is the only man that Mary loves,* and even *doctor who is John,* etc. Of course, such expressions might fail to denote atoms (if e.g. the two tallest students had the same height then none would have the tallest student property). But clearly such expressions cannot denote properties that more than one individual has. So if P had no atoms these expressions would *have to* denote properties which no individual has, and that is clearly wrong.

CONCLUSION. We have not taken P as complete and atomic in order to, in effect, treat it as the power set of some set (the universe of discourse). In fact, we have not taken P as atomic, but only required that it have some atoms. Moreover, the notion of an individual is conceptually and formally independent from that of an atom, and atoms are not mysterious. They are the intended denotations for many common expressions. We refer the reader to KEENAN (to appear a) for a more detailed discussion of this argument.

2.4. A new approach to hyperintensional CNPs

What are the principle differences between atomic and non-atomic P's (assumed complete without further statement) and what then is the evidence for or against taking T_{CNP} as atomic? Here we note just one such property and refer the reader to KEENAN (to appear b) for proofs of the claims made below and a much more thorough discussion. The principle diffference is this: If P is atomic (and complete), then p and q are the same elements of P iff they are members of exactly the same individuals. So if p is different from q then there is an individual which contains one but not the other. This condition fails however for non-atomic algebras. We may have distinct properties in exactly the same individuals. Query: Do we want this?

The prima facie case is overwhelmingly yes. Intuitively we do not want to say that *doctor* and *lawyer* are the same property even if the individuals with one are just those with the other. But of course "standard" intensional logic (SIL) has addressed this problem as follows. Let us think of T_{CNP} (in an intensional logic) as the set of functions from possible worlds J into extensional properties, i.e. as $F_{P/J}$. And if we interpret *doctor* and *lawyer* by functions in this set then obviously they may have the same values (extensions) at some of the j in J but still be different functions as long as they do not have the same value at all j in J. And this answers the prima facie problem, though it does seem funny that in some models, now taken as triples <2,P,J>, *doctor* and *lawyer* are interpreted as the same elements of $F_{P/J}$. That is, in some models they have the same intension and in others they do not. We rather think of the intension of a CNP as constant, not varying with how the world is.

But the standard approach, as has been recognized, is not sufficiently general. Thus if P is complete and atomic, any two CNPs which of necessity have the same extension (same value at j) in every possible world must have the same intension, that is be the same function from J into P, regardless of what J and P are chosen. But there are many examples of such CNPs which still nonetheless should be interpreted as different properties, e.g. *imaginary horse, imaginary lion, book that John intended to write but never wrote and never will write,* etc. And clearly no individual, such as me or you or my horse, can have the property expressed by *imaginary horse*. And since, in a complete and atomic P, there is only one element that no individual has, namely the 0 property, the extension of *imaginary horse* in each j in J must be the 0 property. Ditto for *imaginary lion,* etc. Hence on the standard view the hyperintensional CNPs must always be interpreted by the same element of $F_{P/J}$, that is have the same intension. So sentences like *an imaginary horse is an imaginary lion,* etc. will be valid, which is obviously wrong.

But if we do not require that P be atomic we may correctly represent different hyperintensional CNPs by different properties with the same extension (in particular the 0 extension) without recourse to a possible world semantics at all. Thus not requiring P to be atomic gives us the potential for correctly representing valid arguments which are incorrectly represented on standard approaches. Notice of course that taking P as complete, with atoms, but non-atomic is a move completely independent of possible worlds representations for CNPs. As that approach does seem to correctly represent at least certain logical notions of necessity and possibility we could on the suggested approach still take T_{CNP} in the intensional logic as $F_{P/J}$, we are merely requiring that P be non-atomic. And we do not need to use the j's to distinguish *imaginary lion* from *imaginary horse* but we can still use them to distinguish say a *possibly Albanian diplomat* from a *necessarily Albanian diplomat*.

If this approach to such irrealis APs as *imaginary, unreal, pretend, make-believe* and perhaps *mythological* and *fictional* proves viable it will constitute strong motivation for a boolean approach to semantics, as it would have been inconceivable had we not been taking T_{CNP} as a boolean algebra in the first place.

3. EXTENDING THE CLASS OF DIRECTLY INTERPRETABLE EXPRESSIONS

3.1. The most obvious advantage of our boolean approach is that we have a general - but not infallible! - way to interpret conjunctions, disjunctions, and negations of expressions in a given category; namely as the meets, joins, and complements respectively of the interpretations of the conjuncts, disjuncts, and "negatees". Thus we need not pretend that the boolean connectives (*and, or, not*) "really" only apply to sentences and "translate" sentences containing boolean combinations of non-sentences into ones where all boolean combinations are sentences. Notice that this is the same type (though in a sense lesser in magnitude) of advantage as Montague's original proposal. There is no particular difficulty in translating *every man walks* into $(\forall x)(\mathrm{man}(x) \rightarrow \mathrm{walk}(x))$ but if we do we are saying that the obvious syntactic structure of the English sentence is not the one we use to assign it a meaning, and we are left with the problem of explaining how a "right" logical form is learned given all the possible ones which differ from the "right" one only by logical operators.

Similarly on the boolean approach taken in LT, sentences like *John read Ulysses, every student read a book,* and *every teacher both read and criticized some book* have identical logical forms up to the difference in internal structure of the NPs and the TVP. So we have a better account on a boolean approach of how sentences are assigned a meaning as a function of their form.

Perhaps more important, we have a better account, at least an account, for why *and, or,* and *not* should be usable so freely in forming complex expressions in most categories (*sing and dance, some but not all, dishonest or careless, can and should, both in and behind,* etc. Namely *and, or,* and *not* are always interpreted as *meets, joins,* and *complements*. That they are meets, etc. of course depends on what their arguments are. So one can imagine that on the basis of a few simple examples, *sing and dance, John and Mary,* etc. one learns the basic boolean properties of the boolean operators, and then extends them naturally to other categories, even in fact to collocations which are not natural categories, as in *every diligent but not necessarily every intelligent student will pass.*

Note of course that interpreting the boolean connectives as the appropriate boolean functions makes very strong predictions concerning the logical behaviour of the expressions in question, ones that are often but not always borne out. For example, not all uses of *and* in English appear to be commutative (cf. *Mary got pregnant and (then) married* vs. *Mary got married and (then) pregnant*). (For more interesting cases see Section 4.)

A second advantage here is that we can directly interpret negation in all categories (*is bald/isn't bald, a solid but not very pretty house, near but not on the table, some but not all,* etc.). So in particular we have a distinction between VP negation and sentence negation, and can thus handle the basic cases of presupposition without recourse to multivalued logics, supervaluations, etc.

3.2. A new approach to presupposition

Using the boolean representations of LT, all the sentences below except (5d) entail (5e).

(5) a. (The king of France)(be bald)

b. (The king of France) [not (be bald)]

c. (The king of France) [be (not bald)]

d. (not [(the king of France)(be bald)])

e. (The king of France)(exist).

To see that the relevant entailments hold, consider first that extensional (transparent) VPs like *be bald* are booleanly speaking structure preserving functions, that is, homomorphisms. More explicitly, we say that a function f from a (boolean) algebra B into an algebra D *preserves meets* iff for all x,y in B, $f(x \wedge y) = f(x) \wedge f(y)$, where the meet on the right of course refers to meets in D since f(x) and f(y) are elements of D. And to see that VPs like *be bald*, constructed as function from T_{NP} (= $P(T_{CNP})$ and thus a power set algebra) into T_S (= 2, the algebra referred to in Section 1) should be constrained to preserve meets, note e.g. that *John and some teacher are bald* must have the same truth value as *John is bald and some teacher is bald*. Similarly we say for f as above that f *preserves complements* iff for all x in B, $f(x') = (f(x))'$. And to see that semantically *be bald* preserves complements notice that *(not(every student)) is bald* must have the same truth value as *it is not the case that every student is bald*. We now define:

<u>DEFINITION 2</u>. h from B into D is a *homomorphism* iff h preserves meets and h preserves complements.

It follows from the (standard) definition that h preserves joins, since $(x \vee y) = (x' \wedge y')'$. Similarly, if $x \leq y$ then $h(x) \leq h(y)$. Further h maps the unit in B onto the unit in D, and ditto for the zero in B (onto the zero in D). To see the last point note that $h(0_{\mathcal{B}}) = h(0_{\mathcal{B}} \wedge 0'_{\mathcal{B}}\mathcal{B}) =$ $= h(0_B) \wedge h(0'_{\mathcal{B}B})$, since h preserves meets, $= h(0_B) \wedge h(0_B)'_D$, since h preserves complements, $= 0_D$, since the meet of any element, even $h(0_B)$, with its complement is the zero element of the algebra.

Second, consider the natural semantics for *the*. It maps properties onto sets of properties as follows: *the*(p) is the unique individual which has p if there is one, and otherwise it is the zero element of T_{NP}, that is the empty set. More exactly, *the*(p) is the individual determined by p if p is an atom, and the empty set otherwise. (So we have another motivation for wanting P to have atoms, if it did not *the*(p) would always be $\emptyset$.)

Third, the (transparent) homomorphisms, e.g. *be bald*, etc. themselves form a natural algebra (we want to interpret expressions like *be bald and not be old,* etc. as the relevant meets and complements of VP denotation. And it turns out that a VP homomorphism of this sort is defined by stating its values on the individuals! See LT for a proof. That is, for any function from the individuals into 2 there is a unique (complete) homomorphism from $T_{NP} = P(T_{CNP})$ into 2 having just those values on the individuals. And in particular if h is a VP homomorphism then (h') is that VP homomorphism which assigns to each individual I the opposite value from what h assigns it. Thus *(not(be bald))* will be true of John iff *(be bald)* is false of John, which is intuitively correct.

Now consider the entailments mentioned above. If either *be bald* or *not(be bald)* hold of the denotation of *the king of France* then that denotation is not the zero element since *be bald* etc. are homomorphisms and map zero elements onto the zero element (f) in 2. And if *the king of France* denotation is not the zero element it is an individual and thus has the existence property. So both (5a) and (5b) entail (5e). Notice that (5d), sentence negation, will not entail (5e) since (5d) will be true just in case (5a) is false, and if France has no king then (5a) is false. So sentence and VP negation are non-trivially different at this point.

And this suggests the following definition of presupposition, using $H_{D/B}$ to denote the set of homomorphisms from B into D:

DEFINITION 3. For all b in B, d in D, h in $H_{D/B}$, if $H_{D/B}$ is a boolean algebra then the pair <h,b> logically *presupposes* d iff $h(b) \le d$ and $(h')(b) \le d$.

And by extension we may define:

DEFINITION 4. For S and T sentences of L, S *presupposes* T iff S is of the form (np,vp) and for all interpretations m of L, <m(vp),m(np)> logically presupposes m(T).

Definition 4 may seem insufficiently general in that it only applies to S's of the subject-predicate form, and not say to ones that are conjunctions of other S's, etc. But the definition does appear to capture the clearest case of presupposition in the literature and it nowise precludes any extensions to larger classes of S's.

On the other hand, one may doubt whether Definition 4 will apply to the other clearest case in the literature, namely that factive sentential predicates presuppose their sentential subjects. It would *appear* that such predicates are not homomorphisms as assumed in Definition 4. E.g. (6a) does not even entail (6b), much less preserve meets.

(6) a. That John passed and Mary failed is strange
b. That John passed is strange and that Mary failed is strange.

In fact, however, there is an independently motivated analysis of factive predicates which does treat them as homomorphisms and which accounts for an ambiguity and some entailments not, to my knowledge, previously noticed. I sketch the analysis here and refer the reader to LT (to appear) for details.[1]

Take (standardly) the intensional type for S to be the set of functions from J, the set of possible worlds, into 2. By pointwise definition on J it is a boolean algebra, isomorphic to the power set of J. Denote this set Pr (for proposition). A *property of a proposition* will be a function assigning each proposition, i.e. each element of Pr, a truth value. And for each p in Pr, define I_p to be the set of properties assigning p value t. Interpret *that* (complementizer) in English as a function mapping each p in Pr onto I_p, and note that *that* is one-to-one from Pr onto the set of I_p's. And take the type for sentence complements in general to be all the sets obtained by taking arbitrary intersections, unions, and complements of the I_p's. That set is provably the set of all sets of properties of propositions, i.e. P(P(Pr)), a power set boolean algebra. Of course conjunctions, disjunctions, and negations of *that*+S's are just the meets, joins, and complements of the conjuncts, etc. as usual. So for example, in (6a) above *be strange* is predicated of $I_{(p \wedge q)}$ the set of functions which assign $(p \wedge q)$ value true, where p is the proposition which interprets *John passed* and q the one which interprets *Mary failed*.

In (6b) on the other hand, *be strange* is predicated once of I_p and once of I_q. Now given that neither p nor q are $\leq$ to the other, so p,q, and $(p \wedge q)$ are all distinct propositions, it follows that I_p, I_q, and $I_{(p \wedge q)}$ are different sets of properties. In particular $I_{(p \wedge q)}$ will contain that function which maps $(p \wedge q)$ onto t and everything else, e.g. both p and q, onto f. Hence, correctly, (6a) on this account will not entail (6b).

Notice now that all sentence complement taking predicates are homomorphisms! Thus (7a) and (7b) will be logically equivalent, and this judgment is correct:

(7) a. Both that John passed and that Mary failed are strange

b. That John passed is strange and that Mary failed is strange.

Similarly, this analysis predicts the logical equivalence of (8a) and (8b).

(8) a. That John passed and not that Mary failed is strange

b. That John passed is strange and it is not the case that it is strange that Mary failed.

Notice also that elements in the type for sentence complements ($\bar{S}$) will be denoted by expressions other than *that*+S's and their boolean combination. At a first guess for example *everything* in *John believes everything* should be the intersection of the I_p taken over all p in Pr. And *something that Harry said* might be represented as the union of the I_p which are such that Harry said I_p is true, etc. And the $\bar{S}$ taking predicates also behave homomorphically on such elements of $T_{\bar{S}}$, e.g. (9a) and (9b) below are judged logically equivalent:

(9) a. Something but not everything is strange

b. Something is strange but it it not the case that everything is strange.

(Many intriguing questions arise here concerning the binding of possible world indices, the validity of sentences like *John believes everything that he believes*, the entailment between *everything which is so is strange* therefore *the fact that everything which is so is strange is itself strange*.)

Consider now the specifically factive character of *be strange, ironic, surprising, pleasing,* etc. Clearly *that John passed is strange* entails *John passed*. So we want to restrict the possible interpretations of $\bar{S}$ taking VPs like *be strange* etc. to those functions from P(P(Pr)) into Pr which are both homomorphisms and *factive*, as defined in:

<u>DEFINITION 5</u>. A function f from P(P(Pr)) into Pr is *factive* iff for all p in Pr, $f(I_p) \leq p$.

Chosing *be strange* from the factive functions then we are guaranteed that the truth value of *that John passed is strange* in j $\leq$ the truth value of *John passed* in j, and thus the former sentence entails the latter.

Now consider boolean combinations of factive predicates. Clearly the a- and b-sentences below are judged logically equivalent:

(10) a. That John left early is both strange and unpleasant
b. That John left early is strange and that John left early is unpleasant.

(11) a. That John left early is not strange
b. John left early and it is not the case that it is strange that he did.

(10a) says that the value of a conjunction of factive predicates on a propositional individual is defined pointwise on the individuals (functions on the individuals extending uniquely to complete homomorphisms). Thus for f and g factive functions $(f \wedge g)(I_p) = f(I_p) \wedge g(I_p)$, meets on the right being taken in the algebra Pr of propositions.

More interestingly, (11) says that the complement of a factive predicate is still factive. That is, $(f')(I_p) = p \wedge (f(I_p))'$, (meets and complements on the right taken in Pr of course). It is straightforward to show that meets and complements of factive functions are factive, and that meet and complement as defined satisfy the axioms of boolean algebra. (Joins are defined pointwise on the I_p's like meets.)

Thus (treating $\bar{S}$'s as NP's) our definition of presupposition applies to factive predicates, since they are homomorphisms, and, moreover, yields the correct results: Both *that* P *is strange* and *that* P *(not is strange)* entail P. And *It is not the case that* [*that* P *is strange*] does not, as it only says that *that* P *is strange* is false, which it will be if P is false.

In summary, being able to take complements in the types for essentially all categories has enabled us to define a presupposition relation which captures the clear cases, leaves extensions to less clear ones open, and depends crucially on sentential and ordinary VPs being boolean categories with the consequent distinction between VP and S negation. Moreover, our judgments of presupposition here are accounted for by the independently motivated assignment of boolean structure to the categories. No additional apparatus (3 truth values, gaps, supervaluations) are needed then to capture the clear cases of presupposition: they follow from the boolean structure of the relevant categories.

4. ENRICHING THE TYPES

4.1. Since the types for most categories are boolean algebras it is not surprising that subcategories of a category may be defined according as the expressions in them must satisfy one or another boolean conditions on their denotations. For example, the APs *female, tall, skillful, fake,* and *alleged* all belong to logically distinct subcategories of AP, distinguished in terms of the boolean properties of the functions which can interpret them. (And each can be booleanly distinguished from the irrealis APs like *imaginary* if the semantics suggested in Section 1 is adoped.) And one of the ultimate aims of a semantic investigation of a language is to state the "meanings" of each expression in the language. While we are very far from that goal, being able to discriminate subclasses of expression which are grossly semantically similar is a positive step. So the fact that many expressions can be semantically distinguished in terms of the boolean properties of their denotations is a positive recommendation for a boolean approach to semantics.

Furthermore, most categories will be interpreted by functions from one boolean algebra to another, and hence distinct categories may be compared as to whether the conditions used to distinguish their subcategories are the same or not. And many striking similarities emerge. E.g. the logical subcategory features we need for APs overlap very significantly with those we need for adverbs, but almost not at all for those we need for VPs. VPs on the other hand are logically very similar in terms of subcategorization to TVPs, heads of possessives, etc. Grouping together categories which share many logical subcategorization features we find that they correspond reasonably well to natural syntactic classes (or super classes), which further supports the claim that the syntactic structure of a language reveals its logical structure.

Table I below is a first and very incomplete attempt to state these syntactic and logic correlations. On the left we give three syntactically defined classes of expressions and in the columns on the right which classes are subcategorized for which logical properties. The syntactic classes are: Modifiers, Predicatives, and Specifiers. *Modifiers* (Mods) are expressions which combine with elements of various categories to yield expressions in the same category. They will include APs, adverbs, PPs, ad-adjectives (e.g. *very* in *very tall,* etc.) and perhaps some ad-determiners (e.g. *very, too,* etc. in *very many, too few*).

Predicatives (Preds) are expressions which combine with full NPs and various "nominalized" structures, e.g. nominalized S's (including $\bar{S}$'s), VPs, etc. They include the VPs, TVPs, Ditransitive VPs, Prepositions, "transitive"

CNPs e.g. *relative (of)*, *employer (of)*, etc. and heads of possessives (somewhat debatably). E.g. we analyze *('s) father* as something which combines with an NP such as *every man* to form an NP, *every man's father*. Preds are further distinguished from Mods in that only Preds may impose case on their (NP) arguments.

Specifiers (Specs) hardly constitute a super class since the only clear cases are the Determiners (Dets): *every, a, no, some but not all,* etc., though possibly the *to* in *to smoke is unhealthy* and possibly sentence complementizers will ultimately be included here as well. Semantically Specs map a set into a set of a higher type (extensionally, its power set).

TABLE I

Syntactic Class	Logical Subcategorization Features				
	restrict. absolute	preserve structure (Hom.)	increase decrease conservative	reverse polarity	transparent
Mods	yes	never	-----	(?) no	yes
Preds	(never)	yes	-----	yes	yes
Specs	-----	never	yes	(?) yes	yes

I shall first discuss the interpretation of the entire Table and then present some of its entries in more detail, defining the logical subcategorization features.

Column 1 says that Mods are commonly subcategorized according as they are restricting or absolute. Preds never are, though the line below *never* emphasizes that most of the failure is definitional. *restricting,* etc. are features only defined on functions from an algebra into itself, and most Preds are not functions like that, though those that are, like heads of possessives, are never subcategorized as restricting, etc. Specs are never functions of the right sort.

In column 2 we see that Preds are commonly subcategorized according as they preserve the boolean structure of their domains, that is, are homomorphisms as defined earlier. Mods and Specs never are, though logically they could be.

Column 3 days that features like *increasing*, etc. only apply to functions of the sort that Specs are. Column 4, *polarity reversal* (and polarity preservation) logically applies to all three syntactic superclasses and, due to limited investigation on my part, all entries are tentative. Still it seems to me that Mods will never be so subcategorized, Preds very likely will be, and Specs I am unsure of. (Many of them do reverse polarity, but that normally follows from the independent constraint on the Spec, as most are logical constants, so it is not likely that we need subcategories here, that is, Specs which are not logical constants but which must be constrained to be interpreted by polarity reversal functions.)

And in column 5 we see that *transparency* may and does subcategorize all superclasses. Moreover, in distinction to the other features *transparency* is *not* defined with reference to the boolean structure of the types. I mention it because it interacts in interesting ways with the other features. E.g. while it is logically possible for Mods to be $\pm$ *restricting* and $\pm$ transparent in all combinations, note:

<u>UNIVERSAL GENERALIZATION 1</u>. All transparent modifiers are restricting.

Likely further study of the distribution of subcategory features within and across superclasses will reveal further constraints on the logical form of natural language which do not follow from the definitions of the logical features themselves. E.g. restricting slightly an observation due to Montague we may say that lexical VPs are always transparent (derived VPs like *be required to be a citizen* need not be).

Turning now to specifics, consider first the Mods. They are interpreted by functions from an algebra into itself (see LT 81 for a generalization of the notion of a modifier). We define:

<u>DEFINITION 6</u>. A function f in $F_{B/B}$ is *restricting* iff $f(b) \leq b$, all b in B.

The most common and most productively formed APs in a language will be restricting. Thus, mixing levels, *tall(man)* $\leq$ *man*, which is equivalent to saying that the set of individuals which have the *tall man* property is a subset, necessarily, of those with the *man* property. Similarly, both *female* and *skillful* are restricting, but *fake* and *alleged* are not. We may then distinguish a proper subcategory of AP, call it AP_{+r}, whose type will be the set of restricting functions from T_{CNP} into T_{CNP}. APs not in this category will not be required to be interpreted by restricting functions.

Note further that some Ad-adjectives will be restricting. E.g. *very(tall)* $\leq$ *tall* (i.e. for all properties p, (*very(tall)*)(p) $\leq$ *tall*(p).

Similarly, manner adverbs and many basic PPs (*in the garden*) will be restricting. E.g. the individuals who are working carefully are a subset of those who are working, those who are working in the garden are a subset of those who are working, etc. Some adverbs of course, e.g. *possibly, apparently, allegedly, ostensibly,* etc. are not restricting.

Moreover, the set of restricting functions from an algebra B into itself possesses a natural boolean structure, which I shall call a restricting algebra, as follows: for f and g restricting functions and b in B, $(f \wedge g)(b) =_{df} f(b) \wedge g(b)$; analogously for joins. And $(f')(b) = b \wedge (f(b))'$; $0(b) = 0$, and $1(b) = b$. The interesting property of a restricting algebra is the definition of complement (note the analogy to the factive predicates here). It says e.g. that *a (not diligent) student* is not simply an object (e.g. my pen) which fails to be a diligent student, it must be a student which fails to be a diligent student.

So the cross categorial generalizations of Table I can be extended regarding Mods as follows: languages may present modifiers of major categories, they may be subcategorized as restricting or not, and boolean combinations of restricting Mods will be interpreted as per the restricting algebra defined above.

And within the restricting Mods there are still quite general subcategory distinctions to make. Consider the logical differences between "absolute" APs like *male, female, mortal,* etc. and "merely restricting" ones like *tall* and *skillful*. The former do more than merely restrict the property they modify, they actually determine another property. Thus to say that Mary is a female lawyer is to say that Mary is both a lawyer and a female individual. But to say that Mary is a tall lawyer does not entail that Mary is a tall individual (e.g. suppose that lawyers are all short compared to individuals generally). So Mary is a tall lawyer only commits us to Mary's being tall relative to lawyers, not to individuals generally. More generally the following argument is valid, but ceases to be when *female* is replaced by *tall* or by *skillful*: Mary is a female lawyer and Mary is an artist; therefore Mary is a female artist.

To build these observations into our semantics in a general way, note that the property of being an individual (i.e. existing) is the unit property of the set of extensional properties P. E.g. it is U if P is taken as P(U). So we define:

<u>DEFINITION 7</u>. f in $F_{B/B}$ is *absolute* iff $f(b) = b \wedge f(1)$, all b in B.

This says directly that a female lawyer is a lawyer and a female existent (individual). Absolute functions are obviously restricting (since b meet anything ≤ b). Moreover, the reader may check that meets, joins and complements of absolute functions are absolute, so these functions actually form a subalgebra of the restricting algebra, one which is in fact isomorphic to B! (Note that two absolute functions differ iff their value at 1 is different; that can be anything we like, so there are as many absolute functions (but not more) than elements of B.)

Not surprisingly then we find that absolute APs behave syntactically much more like CNPs than do the merely restricting ones. E.g. they occur closer to the CNP than the others *a tall female lawyer* but ?? *a female tall lawyer*; they occur predicatively without CNP heads, e.g. *Mary is a female,* but **Mary is a tall*. Like CNPs but in distinction to merely restricting APs they do not naturally take degree modifiers or have comparative or superlative forms. Note that the existence of superlative forms for non-absolute restricting Mods is a natural expectation on this semantics. E.g. *tall tall lawyer* is generally a different property from *tall lawyer,* as the former requires being tall relative to tall lawyers whereas the latter only requires being tall relative to lawyers. However, it follows from the definition of absolute that *female female lawyer* is the same property as *female lawyer*. So iteration is logically pointless. Yet iteration seems to be the natural interpretation for the basic degree modifier *very*. E.g. as a first step in the semantics for *very* we might require that *very*(f)(b) maps b onto f(f(b)).

Notice now the existence of absolute VP mods, and a generalization apparent on boolean approaches but perhaps not on others. First note that, holding the point in time constant (12a) and (12b) are L-equivalent:

(12) a. John [is singing in the garden (at t_0)]
b. John [(is singing and is in the garden) (at t_0)].

So extensionally *to sing in the garden* and *to sing and be in the garden* are the same. Thus the *in the garden* function maps a VP like *sing* onto the meet of *sing* with the fixed VP *be in the garden*. And this intransitive use of *be* means essentially exist (at a point in time), and that is just the unit element of the (extensional) VP algebra, that is, that is the VP homomorphism which assigns all individuals value t. So semantically a stative locative PP (e.g. ones formed from *in, on, at, near, behind*, etc.)

but not in general others is a function f mapping a VP interpretation h onto $(h \wedge f(1_{VP}))$. So to characterize this subcategory it is enough to categorize them as VP Mods, +absolute. The general definition of absolute for Mods takes care of the rest.

Note that on a non-boolean approach it is not at all apparent that there are any logical similarities between PP semantics and (absolute) AP semantics. And this similarity is supported by similarity in syntactic behaviour. E.g. just as absolute but not merely restricting APs may function like CNPs in certain contexts (mentioned above) so PPs (together with the constant *be*) function like VPs, (*John is in the garden,* etc.) but those merely restricting adverbs do not so function, **John is carefully*.

4.2. Homomorphisms

We have already discussed (transparent) VPs like *be bald, be a linguist,* which behave homomorphically on their arguments (subjects). Similarly, many (transparent) TVPs behave homomorphically. E.g. *kiss(Mary and Sally)* is semantically the same VP as *kiss(Mary) and kiss(Sally)*, so *kiss* preserves meets, etc. Of course many TVPs are neither transparent nor homomorphisms (homs), e.g. *seek*, etc. (To look for the President or the Vice Pres. is not necessarily the same as looking for the Pres. or looking for the Vice Pres., though in fact there seems to be an ambiguity here similar to that mentioned for the factive predicates in (6)). There is however an interesting cooccurrence generalization here. Namely, while it is logically possible for a TVP to be $\pm$ tr(ansparent) and $\pm$ hom, and yield VPs which are any combination of these two features, in fact it seems that +tr, +hom TVPs always yield +tr, +hom VPs. E.g. the VP *kiss(Mary)* is clearly transparent and a homomorphism. I can find no apparent logical reason for this, but it seems similar to the generalization noted earlier that lexical VPs are usually transparent. So it seems that when an n-place Pred is transparent on its argument it forms n-1 place ones which are transparent on theirs $(n > 0)$. Similarly perhaps for the feature +hom.

Note also that for the +tr +hom n-place Preds (n is the number of NPs they ultimately combine with to form a Sentence or 0-place Pred) we have a quite general syntax and a correspondingly general semantics: The 0-place Preds are the sentences, and their type is (extensionally) 2, a complete and atomic boolean algebra. The n+1 place Preds are, categorially, P_n/NP, and semantically their type is the homomorphisms from T_{NP}

into T_{P_n}, regarded as a boolean algebra itself, the operations being defined pointwise on the individuals.

And it seems that this general notion of an n-place Pred is linguistically enlightening. For example, Passive can be formulated (see KEENAN 1980 for details) as an operation which universally maps n+1 place preds onto n-place preds, with of course a uniform semantics - that in (13) will do as a first approximation.

$$(13) \qquad (\ldots(\text{pass}(p_{n+1}))(x_1)\ldots)(x_n) = \bigvee_y (\ldots(p_{n+1}(x_n))(x_2)\ldots(x_{n-1}))(y),$$

where the x_i and y range over individuals. So this view leads us to expect the existence of passives of intransitives, as in Turkish, Latin (*curritur* = being run), etc. of the category sentence! Similarly, Causative can be given a general definition as a class of operations deriving n+1 place Preds from n-place ones, etc.

Note also that it is not only the "verbals" among the Preds which may behave homomorphically. Some "transitive" CNPs will. E.g. if John has the property expressed by *friend of both Bill and Mary* then he must have the property expressed by *friend of Bill and friend of Mary*, and conversely. Moreover, noting that in LT the extensional type for CNP is isomorphic to that for VP, it will follow that the extensional type for the homomorphic "transitive" CNPs will be isomorphic that for the +hom TVPs, actually justifying our informal use of "transitive" here and giving to expect that e.g. "nominalizations" of TVPs will yield "transitive" CNPs, e.g. *destroy (the city) / destruction (of) the city*, etc. Much more could be said here.

So Preds are commonly subcategorized as +hom or -hom. (Some -hom VPs will be: *love each other, be the two students I know best, ate the whole cow (between them)*, etc. Note that subjects of such predicates formed with *and* will probably require a different *and* from the one we have been treating booleanly. Thus not all the *and*'s in *Both John and Mary and (also) Bill and Emily love each other* can be treated as intersections, otherwise we obtain as a reading that expressed by *John, Mary, Bill, and Emily all love each other*.

It is interesting to query here why Mods and Specs (≈ Dets) are never subcategorized as +hom. For the Mods the features we have used are largely incompatible with being a homomorphism. E.g. suppose that f is both restricting and a homomorphism. Since it is restricting we have that

$f(p') \leq p'$, and since a hom. $f(p') = (f(p))'$. So $(f(p))' \leq p'$, so $p \leq f(p)$. But since f is restricting, $f(p) \leq p$. Thus $f(p) = p$, and since p was arbitrary we have that f is the identity function. So only one restricting function is a homomorphism.

And it is easy to show that no *negatively restricting* function ($f(p) \leq p'$, all p), the interpretation for APs in the *fake*, etc. class, can be homomorphisms (since $f(1) = 0$). Similarly, the irrealis ones will not preserve the unit. But why aren't there other subcategories of AP, say the poorly understood class of *apparent, alleged, ostensible,* which behave homomorphically? And why do not Dets as a class (or as logical constants) preserve the boolean structure?

4.3. Determiners

In the simplest cases Dets are functions from B into P(B). We define:

DEFINITION 8. f in $F_{P(B)/B}$ is *increasing* iff for all p,q,r in B, if $p \in f(r)$ and $p \leq q$ then q in f(r). f is *decreasing* iff if $p \leq q$ and q in f(r) then p in f(r).

For example, without argument, Dets like *every, a, the, three, most, more, than half, uncountably many,* etc. are increasing. Negations of increasing ones are decreasing, e.g. *not a, not every,* etc. as well as *no* (= *not a*), *at most three, fewer than three,* etc. Dets like *all but three, some but not all, exactly three,* etc. are neither increasing nor decreasing. While the features *increasing* and *decreasing* have received the greatest attention in the literature on Dets (logically speaking) it is the property of being *conservative* defined below which actually seems to characterize the (one-place, transparent) Dets in English.

DEFINITION 9. f in $F_{P(B)/B}$ is *conservative* iff for all p,q in B, $p \in f(q)$, iff $(p \wedge q) \in f(q)$.

This definition may seem unintuitive and "mathematical" at first sight but in fact it is based on a sound intuition, one closely related to the Fregean compositionality condition. First consider that some simple Dets clearly meet it. Suppose that every student has a property p, and let s be the *student* property. Well, then clearly every student has both s and p, that is $(s \wedge p)$. And if every student has $(s \wedge p)$ then in particular every student has p. So $p \in$ every(s) iff $(p \wedge s) \in$ every(s). Other simple Dets are equally easily reasoned, as are more complex ones like *some but not all* and the italic portions of (14) below:

(14) a. *Every* student *but John* left.

b. *More* students signed *than teachers who didn't (sign)*.

We refer the reader to KEENAN & STAVI (to appear) for a thorough exposition of the treatment of Dets in English presented here.

The intuition behind the definition is this: if d is an English Det we expect the interpretation of e.g. d(*student*) to depend in a substantive way on which individuals have the *student* property, and to not depend on ones that do not. d(*student*) could not e.g. refer to the properties shared by all cats. And Definition 9 captures (perhaps not as perspicuously as it might) this intuition. For it follows from Definition 9 that the value of a Det f at a property p is determined by those properties q in f(p) which are $\leq$ p and thus ones which *only* individuals having p have. More explicitly:

<u>THEOREM 2</u>. Let $A = \{q \in f(p) : q \leq p\}$. Then $f(p) = \{r : r \wedge p \in A\}$.

<u>PROOF</u>. Let $s \in f(p)$. By the definition of conservative, $(s \wedge p) \in f(p)$, and since $(s \wedge p) \leq p$ then $(s \wedge p) \in A$, so $s \in \{r : (r \wedge p) \in A\}$. Going the other way, let $s \in \{r : (r \wedge p) \in A\}$. So $(s \wedge p) \in A$. By definition of A, then $(s \wedge p)$ in f(p). Thus $f(p) = \{r : (r \wedge p) \in A\}$. □

This characterization of Dets is interesting for two additional reasons. First it clarifies the difference between the kinds of functions Dets are as opposed to homomorphisms. Since the value of a Det at p depends only on properties $\leq$ p, it follows that a Det has the smallest range of possible values at the 0 property, the next smallest range of possible values at the atoms, etc. and the greatest range at the unit property. Homomorphisms are not at all like that. They must for example map the unit onto the unit so they have no choice at the unit property (though as there are Dets which map the unit onto the unit this fact is not incompatible with being a Det).

And second, this characterization brings out a certain similarity between Dets and the restricting APs, clearly the most widespread and productively formed of the APs. Namely, a restricting AP f must map p onto some $q \leq p$. And a Det(p) is a function of a set of $q \leq p$. This seems a reasonably natural analogue of "higher type" restricting AP (though the closest analogue would be a function mapping p simply onto a set of properties $\leq$ p, and such functions are not Dets).

4.4. Polarity reversal

Drawing on a number of insightful observations of FAUCONNIER (1979) and LADUSAW (1979) we define, for B and D any boolean algebras:

DEFINITION 10. f in $F_{D/B}$ *reverses polarity* iff for all x,y in B, if $x \le y$, then $f(y) \le f(x)$; and f *preserves polarity* if $f(x) \le f(y)$.

Although not described in boolean terms of polarity reversal, Fauconnier and Ladusaw have pointed out interesting correlations between the presence of negative polarity items and polarity reversal (pr) operators. Roughly, a sufficient (but not necessary) condition for negative polarity items to occur is that they be under the scope of a pr operator.

Notice that negation (boolean complement) is a pr operator from an algebra into itself. In fact there we have $x \le y$ iff $y' \le x'$. In this sense all categories have pr operators. But again negation is a logical constant, so this fact follows from its independently constrained interpretation. It is not clear that we need to subcategorize a category for such operators. The best candidates for such subcategorization will be sentential Preds, e.g. *implausible, doubtful,* etc. and in a slightly more restricted sense that I do not have the space to define, the negative factives like *strange, surprising,* etc. Among other Preds, possibly the TVP *suspect* and a few "transitive" APs like *suspicious (of), afraid (of),* have readings on which they reverse polarity. Ladusaw (op cit) further cites the interesting case of the negative preposition *without*.

Among the other superclasses, I know of no convincing cases of pr Mods ("transitive" APs etc. are not Mods, they are Preds). Among the Dets there are many. *every* reverses polarity and the complements of most other (not of *every*) "basic" Dets (see KEENAN & STAVI (op cit) for the definition of basic Det) do, e.g. *not a,* etc. But again as these are logical constants it is not clear that we shall have to specify a set of non-logically determinate Dets which are constrained to be interpreted by pr functions. Analogous claims hold for certain "transitive" S Mods, e.g. *if* (on its ordinary truth functional definition). (It is easily shown that for propositions p and q, if $p \le q$ then *if* q $\le$ *if* p in the sense that for all propositions r, *if* q *then* r $\le$ *if* p *then* r.)

So the advantage of our boolean characterization is that we can describe what a variety of elements in different categories have in common logically, namely they reverse polarity, and thus give a uniform statement to many (but not all) of the observations in FAUCONNIER (op cit) and LADUSAW (op cit). I might only note in conclusion here that not all "negative" items will reverse polarity. For example the complements of restricting functions are still restricting and do not reverse polarity. So e.g. the *(not diligent) students* are not necessarily a subset of the *(not diligent) young students,* even though the young students are necessarily a subset of the students; and this follows on our analysis.

Polarity preservation seems somewhat less interesting than polarity reversal, as there seems to be no correlated syntactic property such as triggering negative polarity items. And again while many operators preserve polarity, e.g. homomorphisms always do, it seems unlikely that we will have to constrain subcategories as preserving polarity independently of constraints on subcategories or constants which we need anyway.

4.5. Transparency

Non-rigorously we may say that a function f from X into Y is *transparent* if for all p,q in X and all possible worlds j, if the extension of p in j, ext(p,j), = ext(q,j) then ext(f(p),j) = ext(f(q),j). The "definition" assumes that X and Y are extensional, that is, that the *ext* function is defined on them (cross J). How extensions are defined depends a bit on the category, and of course many categories, i.e. that for TVPs like *seek, want, need,* etc., are not extensions. In general, if the type for C is a set of functions with domain J, then ext(f,j) is just f(j). The properly functional categories are extensional iff their functions are transparent.

As is clear from the above informal sketch, *transparency* is not a specifically boolean property. And part of our interest in it has already been mentioned. Namely certain generalizations concerning the distribution of transparency and other logical subcategorization features appear to impose constraints on the logical form of natural language.

Moreover, the vague claim that transparency and the other features are independent and interact in interesting ways is itself significant. In a certain sense, intensional logic is made up to distinguish transparent from non-transparent operators. And it is natural to wonder whether non-transparent operators (the transparent ones are all representable in an extensional logic) exhibit any interesting boolean behaviour. And they do.

For example, among the "merely restricting" Mods, almost all are non-transparent (*tall*, and few other one dimensional physical object modifiers are the exception here). So e.g. if *doctor* and *lawyer* have the same extension in j, i.e. the doctors and the lawyers are the same individuals, it clearly does not follow that *skillful doctor* and *skillful lawyer* have the same extension, as a given individual might be a skillful doctor but an inept lawyer. But these APs are still restricting, e.g. the skillful doctors in j must be a subset of the doctors in j, all j, so *skillful doctor* ≤ *doctor*, where the type for CNPs in the intensional logic is $F_{P/J}$, taken as a boolean algebra defined pointwise on J.

Similarly we can expect that there will be non-transparent homomorphisms among the Preds, though no examples were given in LT (78). And an algebraic observation (due to Edit Doron, pc) tells us where to look. For suppose that h is a transparent homomorphism. Construct a non-transparent one as follows: Fix a particular k in J, and define f_h by $f_h(x)(j) = h(x)(k)$. In other words, the value of the new function at an argument has as its extension in any j the value of h at that argument in a fixed world k. So if we had a way of referring to possible worlds in our language we could construct such non-transparent homomorphisms. And many candidates suggest themselves.

Consider for example date names. Arguably they specify possible worlds (perhaps sets of them). So from a transparent homomorphism like *be a woman* we should be able to form a non-transparent one like *be a woman in 1972*. This latter Pred clearly seems to be a homomorphism. E.g. if The President and the Vice President were women in 1972 then The President was and so was the Vice President, etc., so the Pred preserves meets, etc. But the function also seems clearly to not be transparent. For example, in our world (1980) *the Prime Minister of Israel* and *Begin* have the same extension. But *be a woman in 1972* holds of the former but not the latter, so it is not transparent.

And this observation generalizes to large classes of subordinate clauses (pointed out to me by David Gil, pc) assumed here to be VP Mods, as is standard in generative grammar. Thus we obtain judgments similar to the ones above if *in 1972* is replaced by *when Nixon was President*, etc. *if* clauses behave similarly.

Yet another case of "possible world fixing" is illustrated by the "picture PP's" discussed in REINHART (1976). Thus from a transparent homomorphism like *cry* we may form a non-transparent one like *cry in Ben's picture*. Clearly, if the President and the Vice President are crying in

Ben's picture then the President etc. is, and conversely. So it preserves meets. And if no political figure is crying in Ben's picture then it is not the case that a political figure is crying in Ben's picture, so it appears to preserve complements. And equally it is not transparent. For if, say in our world, the President is the commander of the armed forces, we cannot infer from *the President is crying in Ben's picture* that a five starred general is, since Ben's picture may not have portrayed the one crying as a military figure at all but only as a civilian one.

Among other Preds, it seems likely that many transitive CNPs are non-transparent homs. Arguably if the property of being *onerous* is a member of *the duties of the President and the Vice President* then it is among *the duties of the President* and also among *the duties of the Vice President*, so arguably (but more work needed here) *duties (of)* preserves meets, etc. But it is clearly not transparent since if the President is the commander of the Army it will not follow that the duties of the President are the same as the duties of the commander of the Army, since *duties of* and many other such expressions pertain to roles or offices, not the individuals which hold them.

So we may infer here that non-transparent operators will also present an interesting boolean behaviour and that in particular Preds wil exhibit members in all combinations of subcategories $\pm$transparent, $\pm$homomorphism (the non-homomorphisms presented earlier are transparent).

5. CONCLUSION

Given that essentially all types are boolean it is not surprising that this boolean structure is used in natural language for reasons other than merely interpreting conjunctions, disjunctions, and negations. And Boole's suggestion that these operations represent "Laws of Thought" seems reasonable.

On the other hand we can expect that specific categories will present structure specific to what they describe, structure that is not specifically boolean. For example, an explicit semantics for Det will require features (or constants) defined in terms of cardinalities, properties which are not specifically boolean. Probably the merely restricting AP semantics will require discussions of "scalar" functions which are possibly not entirely definable in terms of the boolean $\leq$ relation (though we get far here).

Doubtless the semantics for place and time adverbials and verbal subclasses will require analysis of our "natural geometry"; the semantic differences between verbs of motion, desire, and perception will doubtless require serious analysis of motion, desire, and perception, and there is no reason to think they are specifically boolean.

So a boolean approach to semantics is clearly not the whole story, but it is an important chapter.

FOOTNOTES

* The research for this paper was supported by a National Science Foundation Grant no. 202/357-7696.

1. Syntactically the analysis proposed here is similar to that developed in Delacruz' 'Factives and Proposition Level Constructions in Montague Grammar'(DELACRUZ 1976). Semantically there are major differences however. In particular I take the type for sentence complements to be the boolean closure of the I_p's, not simply the set of I_p's as in Delacruz. It is this which allows denotations for *everything that John believes,* etc. More importantly, I distinguish the subcategory of *strange* from that of *true, obvious,* etc. the two classes of adjective lying in different subalgebras of the set of homomorphisms from $T_{\bar{S}}$ into T_S, and for me it is this subcategory difference which accounts for the difference in presupposition. For Delacruz it is the interpretation of *the fact that S* which accounts for its presuppositional nature. Thanks to Susan Ben Chorin, Jeroen Groenendijk and Martin Stokhof for pointing out the similarity to Delacruz (op cit).
2. Note that as regards elements of P (as opposed to $F_{P/J}$, the type for CNP in a standard intensional logic) we can not normally distinguish extensional properties from others, *if* P is complete and atomic as is usual. But if P is not required to be atomic it will now be the case that there will be many, in fact infinitely many, properties in exactly one individual, no matter what one we chose. Only one of those however will be an atom (the one whose meet with the meet of the complements of all the atoms is the zero property). It is of course in the larger class of properties that we will take denotations of *imaginary horse,* etc. The extension of a property may be defined as its meet with the join of all the atoms, and a property will be called *extensional* iff it is $\leq$ to the

join of all the atoms. If we think of the join of the atoms as being the denotation for *real* then p is extensional iff $p = p \wedge real$. Clearly, the English expressions I cited as being naturally interpreted by atoms will meet this condition. That is, when we speak of the *tallest man* we are not normally comparing against Paul Bunyan, etc. So the fact that we cannot characterize atoms as the properties which are in exactly one individual (in a non-atomic algebra) does not argue that the notion of an atom is unclear pretheoretically. They are still the intended denotations of expressions like *tallest man, student who is the only student who passed,* etc.

3. Susan Ben Chorin provides me with the following simpler and direct proof: For x arbitrary in P, assume $x \leq \bigwedge I$. We must show that $x = 0$ or $x = \bigwedge I$, whence $\bigwedge I$ is an atom. From the definition of individual, $x \in I$ or $x' \in I$. Suppose $x \in I$. Then $\bigwedge I \leq x$, whence, from the assumption plus assymetry of $\leq$, $x = \bigwedge I$. Suppose $x' \in I$. Then $\bigwedge I \leq x'$, whence by transitivity of $\leq$, $x \leq x'$, so $x = x \wedge x' = 0$.

REFERENCES

BOOLE, G., 1847, *The mathematical analysis of logic,* MacMillan, Barclay and MacMillan, reprinted by Basil Blackwell, Oxford 1847.

DELACRUZ, E.B., 1976, 'Factives and propositional level constructions in Montague grammar', in: B. Partee (ed.), *Montague grammar,* Academic Press, 1976, pp.177-199.

FAUCONNIER, G., 1979, 'Implication reversal in natural language', in: F. Guenthner and S. Schmidt (eds), *Formal semantics and pragmatics in natural language,* Reidel, Dordrecht.

KARTUNNEN, L., 1973, 'Presuppositions of compound sentences', in: *Linguistic Inquiry* 4, 169-193.

KEENAN, E.L., 1980, 'Passive is phrasal, not sentential or lexical', in: T. Hoekstra, et al. (eds), *Lexical Grammar,* pp. 181-215, Foris Publications, Dordrecht, 1980.

KEENAN, E.L., to appear a, *Eliminating the universe: A properly Leibnizian approach to semantics,* in: UCLA Working Papers in Semantics, 1, 1981.

KEENAN, E.L., to appear b, *A new approach to intensional properties,* in: UCLA Working Papers in Semantics 1, 1981.

KEENAN, E.L. & L. FALTZ, 1978, *Logical types for natural language,* UCLA Occasional Papers in Linguistics 3, to appear in Reidel, Synthese Language Libary Series, 1981.

KEENAN, E.L. & J. STAVI, to appear, *An algebraic representation of quantification in natural language,* in: UCLA Working Papers in Semantics 1, 1981.

LADUSAW, W., 1979, *Polarity sensitivity as inherent scope relations,* Ph.D. diss. Univ. of Texas at Austin, to appear in Garland Press.

MENDELSON, E., 1970, *Boolean algebras and switching circuits,* Schaum's Outline Series, McGraw-Hill Book Co.

MONTAGUE, R., 1972, 'The proper treatment of quantification in ordinary English', in: J. Hintikka, J. Moravcsik and P. Suppes (eds), *Approaches to Natural Language,* 1973.

REINHART, T., 1976, *The syntactic domain of anaphora,* Ph.D. diss. MIT.

THE INTERPRETATION OF ADJECTIVAL, NOMINAL AND ADVERBIAL COMPARATIVES

by

Ewan Klein

1. INTRODUCTION

In this paper, I shall advance a general approach to the interpretation of comparative constructions. The kinds of construction that I am concerned with are the following:

(1) (a) *predicate adjectives*
Chris is taller than Alex is.

(b) *prenominal adjectives*
Norbert is a larger flea than Nat is.

(c) *'quantifiers'*
Jude bought as many apples as Steve did.

(d) *adverbs*
Gill walks as quickly as Peter does.

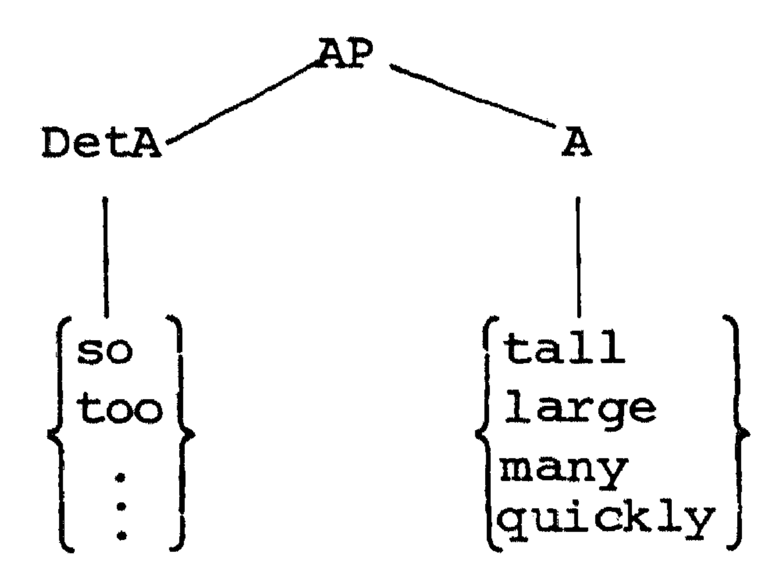
I shall start off by examining examples like (1a) in some detail. My assumption is that the fundamental properties of the comparative can be studied most directly in the simple adjectival construction. I shall then argue that an adequate analysis of (1a) provides the basis for a unified account of the remaining constructional types in (1).

The plausibility of generalizing from (1a) in this way rests on the claim that the head of a comparative clause always consists of some sort of predicate. Within the framework of Montague grammar, we might express this idea as follows:

(2) If α cooccurs with degree modifiers (e.g. *so, too, that, more/-er, as,...*), then α' (the translation of α) is of type $\langle \tau, \underline{t} \rangle$, for some type τ.

In other words, I am suggesting that the expressions *tall, large, many* and *quickly* in (1) are all to be analysed as predicates (though not all of the same type). The hypothesis in (2) will be worked out in more detail in the remainder of the paper. However, two refinements can be mentioned briefly at this point. First, I want to add that α must also be a vague predicate; that is, a predicate which admits truth-value gaps. Second, it is plausible to suppose that α is always of category A, i.e. an adjective. That is, if we take DetA to be the category of degree modifiers,[1] then *so, too,...* will only occur in configurations of the following sort:

(3)



My plan in this paper is to first formulate a semantics for constructions like (1a), and then show how the hypothesis (2) can be supported for each of the remaining sentence types. I will not carry the analysis into great detail here, for reasons of space. My main goal, as I mentioned at the outset, is to sketch a general approach, and to show that is both plausible and promising.

2. PREDICATE ADJECTIVES

In this section, I shall be concerned with measure adjectives - adjectives which cooccur with degree adverbs like *very* and the modifiers *so, too,...* mentioned earlier. I shall adopt the view that adjectives occurring in predicate position are to be analysed as predicates, rather than covert noun-modifiers. Consequently, $\underline{g}(A)$, the type associated with category A, is to be $\langle \underline{e},\underline{t} \rangle$.[2]

It is often observed that measure adjectives belong to the class of vague predicates, i.e. expressions which may not be definitely true nor definitely false of the things they are predicated of. From a formal point of view, the obvious way to represent this characteristic is to let such predicates denote partial functions on the relevant set. In addition, measure adjectives are also context-dependent, a fact I shall return to shortly.

Hence, if α is the translation of a measure adjective and c is a context of use, then $F_\alpha(c)$, the semantic value of α at c, is a partial function from the universe of discourse U to the set of truth values $\{0,1\}$. Let $X^{(Y)}$ be the set of all partial functions from Y to X, and let $Con_{g(A)}$ be the set of constants into which lexical items of category A are translated. Then

(4) If $\alpha \in Con_{g(A)}$ and $c \in C$, then $F_\alpha(c) \in \{0,1\}^{(U)}$

A function like $F_\alpha(c)$ allows us to demarcate two disjoint subsets of U:

(i) the positive extension of α at c, consisting of those elements in U of which α is definitely true; and

(ii) the negative extension of α c, consisting of those elements of which α is definitely false.

In addition, there may be further elements in U for which $F_\alpha(c)$ yields no value. In this case, α has an extension gap at c.

Let us consider an example. Suppose that U consists of the people listed, together with their heights, in (5):

(5)

Chris	6' 2"
Steve	6'
Jude	5' 8"
Gill	5' 6"
Alex	5' 5"

Taking c to be a fairly standard context, $F_{tall}(c)$ will partition U something like this:

(6) *tall:*

+		−
Chriss Steve	Jude	Gill Alex

In this diagram, '+' indicates the positive extension of *tall*, and '-' the negative extension.

I want to turn now to the semantic role of adjective modifiers. To begin with, I shall consider measure phrases. In broad terms, such expressions will map adjective meanings into adjective meanings. For example, *five foot six* will combine with *tall* to produce a complex predicate *five foot six tall*. This predicate will be true of an individual u just in case

$\underline{u}$ is as tall as the standard measure 5' 6".[3] Its interpretation can be represented by the same kind of diagram as I used above for *tall*. (In the interests of simplicity, I am treating *five foot six* as an unanalysable whole.)

(7) *five-foot-six(tall)*:

+				−
Chris	Steve	Jude	Gil	Alex

The effect of the measure phrase is to shift the boundary 'downwards'. Moreover, it eliminates the extension gap associated with its argument. A similar shifting of the boundary, though in the opposite direction, is triggered by the measure phrase *six foot two*:

(8) *six-foot-two(tall)*:

+	−			
Chris	Steve	Jude	Gill	Alex

It should be obvious that for any pair of individuals with distinct heights, there is some measure phrase δ such that δ(*tall*) is true of one of those individuals and false of the other. If we take measure phrases to be dominated by the category DetA, then the associated type $\underline{g}$(DetA) will be $<\underline{g}(A),\underline{g}(AP)> = <<\underline{e},\underline{t}> <\underline{e},\underline{t}>>$. If we assume, in addition, that degree modifiers always close the extension gaps associated with their arguments,[4] then whenever δ is an expression of type $\underline{g}$(DetA), $\underline{F}_{\delta}(\underline{c})$ will be a function which takes a partial function in the set $\{0,1\}^{(\underline{U})}$ and turns it into a total function in $\{0,1\}^{\underline{U}}$.

It might be objected that we have reached a very general characterization of the interpretation of degree modifiers on the basis of a quite atypical class of expressions. For degree modifiers can combine with only a small proportion of measure adjectives; and there do not appear to be modifiers for adjectives like *clever*, say, which can shift the demarcation between positive and negative extensions in a comparable way.

Consider, however, the anaphoric role played by *that* in the following sentences:

(9) (a) Steve is *six foot* tall, but nobody else I know is *that* tall.
(b) You have to be *very* clever to pass this exam, but most of the candidates are *that* clever.
(c) If Jude is late *enough to miss the train*, Alex will probably be *that* late too.

(d) The council is *too* mean *to contribute any funds*, but our friends certainly aren't *that* mean.

(e) Alex is *as* successful *as Howard Hughes*, and I would like to be *that* successful too.

In (9 a,b), the antecedents of *that* are modifiers of the sort I have already discussed. In the remaining sentences, however, the antecedents are complex, consisting of an initial degree word (*enough, too, as*) and a postadjectival complement. It seems fairly clear that these are discontinuous modifiers, which function as semantic units; indeed on some accounts (BOWERS 1975, BRESNAN 1973, CHOMSKY 1965), they also function as syntactic constituents in underlying structure. But of course it is possible to construct indefinitely many such complex modifiers, and consequently *that* will have indefinitely many potential antecedents. Suppose, then, that *that* is a DetA proform. Semantically, it will be treated as a variable, ranging over the same class of functions as those denoted by the complex modifiers in (9). But this class of functions will be exactly the same as we required for the interpretation of measure phrases. Hence, our earlier generalization from measure phrases to degree modifiers as a whole appears to be justified.

However, we cannot let $\underline{D}_{\underline{g}(\mathrm{DetA})}$, the range of possible denotations of DetA expressions, be the complete set $\underline{H} = \{\underline{h} | \underline{h}: \{0,1\}^{(\underline{U})} \rightarrow \{0,1\}^{\underline{U}}\}$. For the latter will contain some functions which violate the grading requirements of measure adjectives. Suppose, for instance, that d is a variable of type $\underline{g}(\mathrm{DetA})$, α is the translation of a measure adjective, and that for some value of d, the universe gets partitioned by $d(\alpha)$ in the following way:

(10)

+	−
Chris Steve	Jude Gill Alex

Then there should be no further value of d such that $d(\alpha)$ induces the partition indicated by the dotted line:

(11)

+	−
Chris Steve	Jude Gill Alex

That is, suppose there is some value of *that* such that *Steve is that tall* is true, while *Jude is that tall* is false. Then there should be no further value of *that* which reverses these truth values, i.e. makes *Jude is that tall* true and *Steve is that tall* false.

Clearly, $\underline{D}_{\underline{g}(\mathrm{DetA})}$ must not contain such mutually inconsistent functions. Moreover, we must also exclude values of d such that for a given vague predicate α, $d(\alpha)$ is inconsistent with the interpretation of α itself. These two requirements are expressed in the following statement:

(12) In any model $\underline{M}$ based on $\underline{U}$, $\underline{D}_{\underline{g}(\mathrm{DetA})}$ is a maximal subset of $\underline{H}$ such that (i) and (ii) are true in $\underline{M}$:[5]

(i) $\forall x \forall y \forall Q[\exists d[d(Q)(x) \wedge \neg d(Q)(y)] \rightarrow \forall d[d(Q)(y) \rightarrow d(Q)(x)]]$

(ii) $\forall Q \exists d[\forall x[Q(x) \rightarrow d(Q)(x)] \wedge \forall x[\neg Q(x) \rightarrow \neg d(Q)(x)]]$.

Once $\underline{g}$(DetA) variables have been introduced, it is straightforward to provide an analysis of adjectival comparatives. My proposal is that (1a), repeated here as (13a), should be assigned a logical structure very similar to that proposed by SEUREN (1973), namely (13b):

(13) (a) Chris is taller than Alex is.

(b) $\exists d[d(\mathit{tall})(\mathit{Chris}) \wedge \neg d(\mathit{tall})(\mathit{Alex})]$.

This says, in effect, that (13a) is true iff there is some value of *that* such that *Chris is that tall* is true while *Alex is that tall* is false.

Comparatives with *as* and *less* can be integrated neatly into this treatment. They are analysed as follows:

(14) (a) Alex is as tall as Chris is.

(b) $\forall d[d(\mathit{tall})(\mathit{Chris}) \rightarrow d(\mathit{tall})(\mathit{Alex})]$.

(15) (a) Alex is less tall than Chris is.

(b) $\exists d[\neg d(\mathit{tall})(\mathit{Alex}) \wedge d(\mathit{tall})(\mathit{Chris})]$.

Given these translations, familiar rules of quantifier logic predict that the following equivalences will hold:

(16) Chris is taller than Alex is $\Longleftrightarrow$
Alex is not as tall as Chris is $\Longleftrightarrow$
Alex is less tall than Chris is.

For further discussion of this analysis, see KLEIN (forthcoming b).

3. CONTEXT DEPENDENCE AND PRENOMINAL ADJECTIVES

I mentioned above that the interpretation of a measure adjective will be dependent on contextual factors. There seem to be basically two ways in which the context can play a semantic role in this connection.

Consider an adjective like *skilful*. When evaluating a sentence containing this expression, we usually require an answer to the question: skilful at performing what activity? Sometimes this information will be supplied in the sentence itself: *Alex is skilful at drawing*. But in interpreting the less explicit

(17) Alex is skilful,

we must look to the wider nonlinguistic context to find out what kind of activity Alex is skilful at.

Second, even when we have established a particular dimension of skill relevant to the interpretation of (17), we still require information about the appropriate comparison (or reference) class. Again, this may be given linguistically. Thus, in *Alex is skilful for a four year old,* we are judging Alex's level of skill relative to that of other four year olds. But again, if this information is not given explicitly, it must be sought in the context of use.

These two kinds of context dependence are usefully discussed in BARTSCH & VENNEMANN (1972), KAMP (1975), LAKOFF (1972), McCONNELL-GINET (1973), SIEGEL (1979) and WHEELER (1972). In KLEIN (forthcoming a), I have attempted to provide a precise formal modelling of them. In the present paper, I shall confine myself to some brief remarks on the topic of comparison classes.

I have already assumed that model-theoretic interpretation of a natural language will be relativized to contexts of use. Let $\mathcal{U}$ be a function from the set of contexts $\underline{C}$ to subsets of $\underline{U}$. Intuitively, if $\underline{c} \in \underline{C}$, then $\mathcal{U}(\underline{c})$ is the comparison class which is relevant to the discourse taking place in $\underline{c}$. Suppose $\mathcal{U}(\underline{c}) = \underline{X}$. When a vague predicate α is evaluated at $\underline{c}$, we want $\underline{F}_\alpha(\underline{c})$ to be 'focussed' on $\underline{X}$. That is, it should partition $\underline{X}$ into a positive and negative partition, disregarding anything outside $\underline{X}$. For the purposes of discourse that occurs in $\underline{c}$, $\underline{X}$ counts as the whole universe. If we suppose, for simplicity, that $\underline{F}_\alpha(\underline{c})$ is undefined for all arguments outside the comparison class[6], this idea can be expressed as follows:

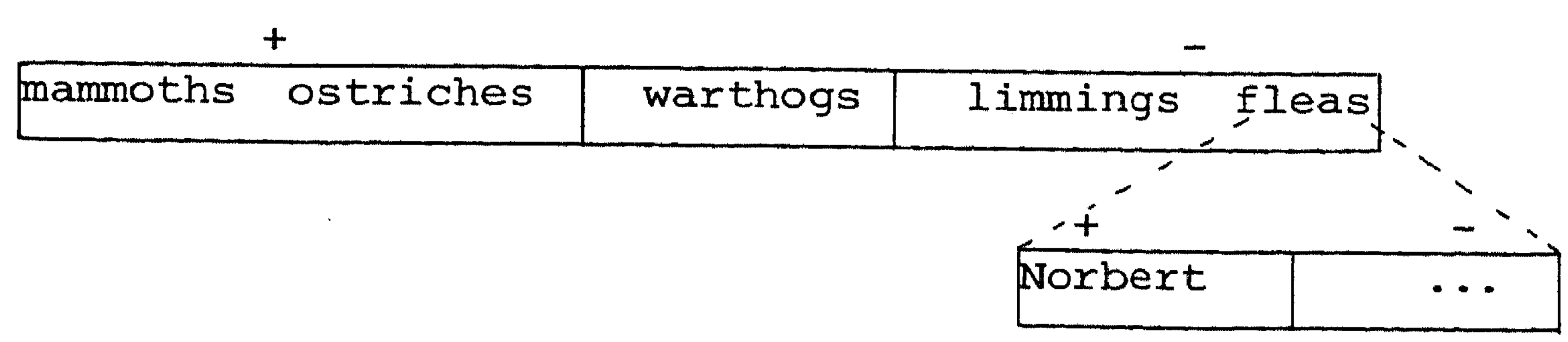
(18) Whenever $\underline{c} \in \underline{C}$, $\mathcal{U}(\underline{c}) \subseteq \underline{U}$, and for all $\alpha \in Con_{\underline{g}(A)}$, $\underline{F}_{\alpha}(\underline{c}) \in \{0,1\}^{(\mathcal{U}(\underline{c}))}$.

Let me turn now to the case where a measure adjective occurs prenominally. Following KAMP (1975), I should like to suggest that the semantic contribution of the head noun is mediated by the context. In other words, the interpretation of the adjective is still context dependent, but we find in addition that the head noun has an important modifying effect on the context. In combination with a suitable device for enabling the context to select a particular criterion of application, this provides a novel means for analysing the well-known example *skilful cobbler*. Suppose this phrase is interpreted in a context $\underline{c}$; then *cobbler* has the effect of modifying $\underline{c}$ to a new context $\underline{c}'$ where the relevant dimension of skill is that of mending shoes. Since *cobbler* will (indirectly) select a different sense of the word skilful from that (indirectly) selected by *darts player, skilful cobbler* need not be coextensive with *skilful darts player* even though *cobbler* is coextensive with *darts player*.

A similar phenomenon arises with the other sort of context dependence. On the most natural reading, *Alex is a skilful child* means that Alex is skilful when compared to other children (as opposed to 'skilful at being a child'). Here, I want to say that the head noun introduces a new context in which the relevant comparison class for evaluating *skilful* is the set of children. More generally, if $[_{\bar{N}}[_{A}\alpha][_{\bar{N}}\beta]]$ is evaluated at a context $\underline{c}$, then its value is the same as $[_{A}\alpha]$ evaluated at $\underline{c}'$, where $\underline{c}'$ is just like $\underline{c}$ except that $\mathcal{U}(\underline{c}')$ is the set denoted by $[_{\bar{N}}\beta]$.[7]

This approach gives us a means of dealing with another familiar problem involving prenominal adjectives. If *large* is taken to be a predicate in *Norbert is a large flea,* there is a danger that, given the additional premise *every flea is an animal,* we will end up with the unwanted conclusion *Norbert is a large animal*. But there is no diffuculty in categorizing prenominal *large* as a predicate so long as it is stipulated that the head noun determines the appropriate comparison class. Let $\underline{c}$[flea] and $\underline{c}$[animal] be two contexts which are the same except that the respective comparison classes consist of fleas and animals. Then clearly it can be the case $\underline{F}_{\mathit{large}}(\underline{c}[\text{flea}])(\text{Norbert}) = 1$ while $\underline{F}_{\mathit{large}}(\underline{c}[\text{animal}])(\text{Norbert}) = 0$. This situation is illustrated in (19), where animals (above) and fleas (below) are partitioned by *large*:

(19)



Although limitations of space have prevented me from developing a detailed proposal, I have attempted in this section to suggest that there is a viable alternative to the prevailing view that prenominal adjectives play the semantic role of common noun modifiers. We can instead analyse them as predicates whose context of use is modified by the head noun.

4. QUANTIFIERS

The expressions *many* and *few* are distinct, syntactically and semantically, from 'classical' quantifiers such as *every, a, some, all, no,* etc. On distributional grounds, it is quite plausible to group them with adjectives rather than classical quantifiers. Unlike the latter, they occur in predicate position, after definite determiners, and cooccur with degree modifiers:

(20) (a) The problems are many/*all.
(b) Sam's many/some* friends were noisy.
(c) The chairs were too few/*some to accomodate us.

Data of this sort is discussed, for example, in BARTSCH (1973), BOWERS (1975), HOGG (1977), JACKENDOFF (1968) and PARTEE (1970).

Let us suppose, then, that *many* is classified as a measure adjective in (20). It follows from my earlier remarks that it will be a vague predicate. It is only necessary to add that it is also a plural adjective, and hence to be interpreted as a predicate of sets, not individuals.

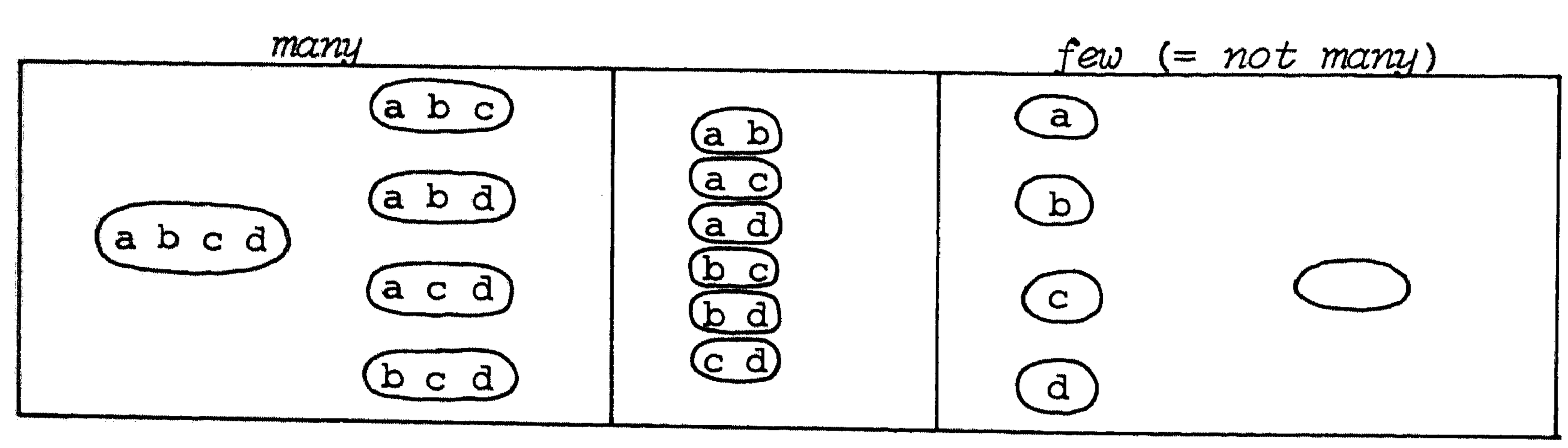
Consider once more the extension of *tall* at a context $\underline{c}$. Given a universe $\underline{U}$, $\underline{F}_{tall}(\underline{c})$ will yield the value 1 for some members of $\underline{U}$, 0 for other memebers of $\underline{U}$, and will possibly be undefined for yet other members. Suppose now that pow($\underline{U}$) is the power set of $\underline{U}$. $\underline{F}_{many}(\underline{c})$ is a function on pow($\underline{U}$). And $\underline{F}_{many}(\underline{c})$ will yield 1 for some sets in pow($\underline{U}$), 0 for other sets, and possibly be undefined for yet others. Moreover, if $\underline{F}_{tall}(\underline{c})(\underline{u})$ is true, and u' has the same height as u, then F_{tall}(c)(u') will also be true. Similarly, if F_{many}(c)(X) is true, for some X ∈ pow(U), and Y has the same cardinality as X, then F_{many}(c)(Y) will also be true.[8] In other words, while *tall* grades along the dimension of height, *many* grades along the dimension of cardinality.

Few can be defined in terms of *many*:[9]

(21) $\forall X[few(X) \leftrightarrow \neg many(X)]$.

Notice, however, that the matrix of (21) will be undefined for any value of X such that $\neg many(X)$ is undefined. On the one hand, this means that (21) should be stipulated to be true only under those valuations which eliminate the extension gap associated with *many*. And on the other hand, (21) is compatible with a situation in which there are sets X which belong to the positive extension of neither *many* or *few*. Setting U to be {a,b,c,d}, this kind of state of affairs is pictured below:[10]

(22)



In order to say something about the logical structure of sentences containing *many* and *few*, I shall introduce into the object language an operator corresponding to pow, namely P. If α is an expression of type $<\tau,t>$, then $P(\alpha)$ is of type $<<\tau,t>,t>$ and denotes the set of subsets of the extension of α. Using '$\subseteq$' in a sloppy but, I hope, intelligible way, this can be expressed as follows:

(23) $\forall X[P(\alpha)(X) \leftrightarrow X \subseteq \alpha]$.

I am going to let P stand in for a plural operator on one-place predicates. So, for example, if *problem* translates as *problem*, the plural noun *problems* will translate as P(*problem*). A set will belong to the extension of P(*problem*) just in case it is a set of problems.

At this point I want to say something briefly about plural definite descriptions. On an intuitive level, *the problems* denotes the set of all problems (in a given context). But it is hardly satisfactory to translate *the problems* as *problem*, even though this will have the desired denotation. On the one hand, *problem* lacks the quantificational structure which we require if we are to capture familiar scopal ambiguities. On the other hand, it makes singular and plural *the* seem totally unrelated in their semantics. Although I do not have space to justify the proposal here, I am going to assume that plural *the* $\bar{N}$ parallels singular *the* $\bar{N}$ in the following way: loosely speaking, it denotes the unique maximal set which satisfies the descriptive predicate $\bar{N}$. So, in particular, *the problems* will translate as (24), where Q is a variable of type $<<\underline{e},\underline{t}>,\underline{t}>$:

(24) $$\lambda Q \exists X[\forall Y[P(problem)(Y) \leftrightarrow Y \subseteq X] \wedge Q(X)].$$

Accordingly, *the problems are many* will translate as (25):

(25) $$\exists X[\forall Y[P(problem)(Y) \leftrightarrow Y \subseteq X] \wedge many(X)].$$

Next, consider (26):

(26) The many unlucky students failed.

As CARDEN (1970) has pointed out, *many* must be interpreted nonrestrictively in this position. I assume, therefore, that in the translation of (26), *many* will fall outside the scope of the uniqueness subformula of the definite determiner. This gives us

(27) $$\exists X[\forall Y[P(unlucky)(Y) \wedge P(student)(Y) \leftrightarrow Y \subseteq X] \wedge many(X) \wedge P(fail)(X)].$$

So far, I have ignored sentences like *many students failed* in which *many* and *few* occupy an NP-initial position. Here, there are good grounds for thinking that these plural adjectives do indeed play the semantic role of determiners. However, it is straightforward to express this second interpretation as a function of their basic adjectival interpretation. Suppose the grammar contains a phrase structure rule of the following sort:

(28) $$\mathrm{NP} \longrightarrow \underset{\begin{bmatrix}+\mathrm{quant}\\ \alpha\mathrm{neg}\end{bmatrix}}{\mathrm{AP}} \quad \bar{\mathrm{N}}\,.$$

The feature [+quant] serves to subcategorize APs which contain quantifying adjectives like *many* and *few* as their lexical heads. [-neg] indicates *many* as head, while [+neg] indicates *few*. There are two corresponding translation rules (where AP', $\bar{N}$' are the translations of AP, $\bar{N}$):

(29) $[-neg]: \lambda Q \exists X[AP'(X) \wedge \bar{N}'(X) \wedge Q(X)]$

$[+neg]: \lambda Q \forall X[\bar{N}'(X) \wedge Q(X) \rightarrow AP'(X)]$.

Accordingly, we get the following translations for sentences containing *many* and *few* in determiner position:

(30) (a) Many students failed.

(b) $\exists X[many(X) \wedge P(student)(X) \wedge P(fail)(X)]$.

(31) (a) Few students failed.

(b) $\forall X[P(student)(X) \wedge P(fail)(X) \rightarrow few(X)]$.

Notice that on this analysis, (31) differs from (30) in having no existential entailments. Hence, we predict that (31) is a logical consequence of (32):

(32) (a) No students failed.

(b) $\neg\exists X[P(student)(X) \wedge P(fail)(X)]$.

Moreover, given the definition of adjectival *few* in (21), we also predict that (31) is equivalent to (33):

(33) (a) Not many students failed.

(b) $\neg\exists X[many(X) \wedge P(student)(X) \wedge P(fail)(X)]$.

In this section, I have argued that *many* and *few* should be classified as plural adjectives which can occupy determiner position. Their quantifying properties can be adequately explained without simply categorizing them as quantifiers. Two factors are responsible: (i) they grade along the dimension of cardinality, and (ii) in determiner position, they introduce quantification over sets. Most of what I have said here also applies to *much* and *little*. The main points of difference are that *much* and *little* are, of course, mass adjectives, not plural, and that they can also play an adverbial role in negative polarity environments.

5. ADVERBS

There are basically three semantic theories of adverbs in the literature. They have been analysed as

(i) predicates of properties,

(ii) predicates of events,

(iii) predicate modifiers.

The first two ideas can be traced back to REICHENBACH (1947). (ii), of course, is also familiar from the work of DAVIDSON (1967), and has received an interesting formulation by CRESSWELL (1974). BARTSCH (1972) adopts a variant of (ii) in which adverbs are treated as predicates of processes. The third alternative is argued for by PARSONS (1972) and is adopted by Montague in various papers. It is developed at some length by THOMASON & STALNAKER (1975), and is further discussed by RICHARDS (1976) and CRESSWELL (1979).

In terms of the argument I have developed in this paper, it is perhaps sufficient to note that (i) and (ii) are viable approaches; on either account, we get the result that adverbs are analysed as predicates of some kind. A more convincing case would be made if I could show that a semantic theory of English which adopts either (i) or (ii) is at least as adequate as one which adopts (iii). Unfortunately, this lies beyond the scope of the present study. Instead, I shall briefly develop a version of (i), and indicate how it copes with certain problems noted by Parsons and Davidson.

To begin with, let me first present Reichenbach's analysis, transposed into the notational conventions of this paper. In his discussion of activities, REICHENBACH (1947: 302) distinguishes between two kinds of property. First, 'general', unspecified second-order properties; and second, 'specific' properties which are delimited in various ways and which hold of individuals who participate in particular activities. The various specific properties of walking at a certain speed, in a certain direction, and so on, have in common the general property of being a walking. Suppose, then, that we interpret *walk* as a function from indices to sets of specific properties, and let P be a first-order property variable. Then the Reichenbachian translation of *Sue walks* will look something like this:

(34) $\exists P[{}^{\vee}walk(P) \wedge {}^{\vee}P(Sue)]$.

We can perhaps loosely gloss (34) as 'there is a particular activity which is a walking and which Sue is involved in'.

An adverb like *slowly* will be a predicate of a first-order property. Let us use *slow* as the appropriate constant. Then *Sue walks slowly* will be translated as

(35) $\exists P[{}^{\vee}walk(P) \wedge slow(P) \wedge {}^{\vee}P(Sue)]$.

I want to suggest a couple of modifications to this scheme. First, in line with my treatment so far, I will revert to an extensional representation. Second, suppose *walk* is again taken to denote a set $\underline{X}$ of individuals, i.e. the individuals who walk. Form the power set of $\underline{X}$. One element of pow($\underline{X}$) will contain all the individuals who walk slowly, another will contain all the people who walk towards Tehran, and so on. I suggest that *slowly* should be interpreted as a predicate of sets, such that for any $\underline{Y} \in$ pow($\underline{X}$), $\underline{Y}$ is in the positive extension of *slowly* just in case every member of $\underline{Y}$ walks slowly. If *slowly* is now translated as *slowly*, (35) can be replaced by (36):

(36) $\exists X[P(walk)(X) \wedge slowly(X) \wedge X(Sue)]$.

EMONDS (1976) has suggested that *ly* adverbs should be assigned to the category of adjectives. For convenience, then, I shall suppose that $\underset{[+ly]}{\text{AP}}$ is the node which dominate adverb phrases. They can be introduced into VP by the following phrase structure rule:

(37) $\text{VP} \longrightarrow \text{VP} \;\; \underset{[+ly]}{\text{AP}}$.

The corresponding translation rule is (38):

(38) $\lambda x\exists X[P(\text{VP}')(X) \wedge \text{AP}'(X) \wedge X(x)]$.

PARSONS (1972: 131) criticizes Reichenbach's analysis on the grounds that it fails to cope with reiterated adverbs, as in

(39) John painstakingly wrote illegibly.

I do not have anything to say here about the placing of adverbs within VP. However, there seems to be no problem about representing the wider scope of *painstakingly*:

(40) $\exists X[P(\lambda x\exists Y[P(write)(Y) \wedge illegibly(Y) \wedge Y(x)])(X) \wedge$
$\wedge\, painstakingly(X) \wedge X(John)]$.

Since I have made the semantics of adverbs extensional, I have to say something about data which parallels the *skilful cobbler* case. That is, we do not want *mend shoes skilfully* to be coextensive with *play darts skilfully* even in those situations where *mend shoes* is coextensive with *play darts*. Clearly, adverbs are context dependent in just the same way as adjectives. Hence, if the solution I sketched for prenominal adjectives is satisfactory, it will also be applicable to adverbs.

DAVIDSON (1967) points out an aspect of the interpretation of adverbs which resembles the *large flea* problem. Given the premises *June swam the channel quickly* and *everyone who swims the channel crosses the channel*, there is a danger of deriving the unwanted conclusion *June crossed the channel quickly*. The solution to this problem can again be found in the notion of comparison class. In the general case, the comparison class of an adverb introduced by (37) will be the extension of $P(\mathrm{VP}')$. Let $\underline{c}[\mathrm{pow(swim)}]$ be that context just like $\underline{c}$ except that the comparison class is the power set of the extension of *swim the channel*; and analogously for $\underline{c}[\mathrm{pow(cross)}]$. Then we may well have $\underline{F}_{quickly}(\underline{c}[\mathrm{pow(swim)}])(\underline{X}) = 1$ but $\underline{F}_{quickly}(\underline{c}[\mathrm{pow(cross)}])(\underline{X}) = 0$. That is, compared with all other swimmings, $\underline{X}$ is a quick swimming, but compared with all other crossings, $\underline{X}$ is not a quick crossing.

6. CONCLUSION

In Section 2, I presented a semantics for adjectival comparative constructions which has the following virtues:

(i) The interpretation of the comparative adjective is given as a function of its positive counterpart. Since morphological evidence across a wide range of languages suggests that the positive is the basic form, this is preferable to any semantic theory which treats the positive as an implicit comparative.

(ii) No reference is made to abstractions such as degrees or extents. The only semantical extensions involved in the present theory of comparatives are independently required for (a) a treatment of vagueness, and (b) the interpretation of the DetA anaphor *that*.

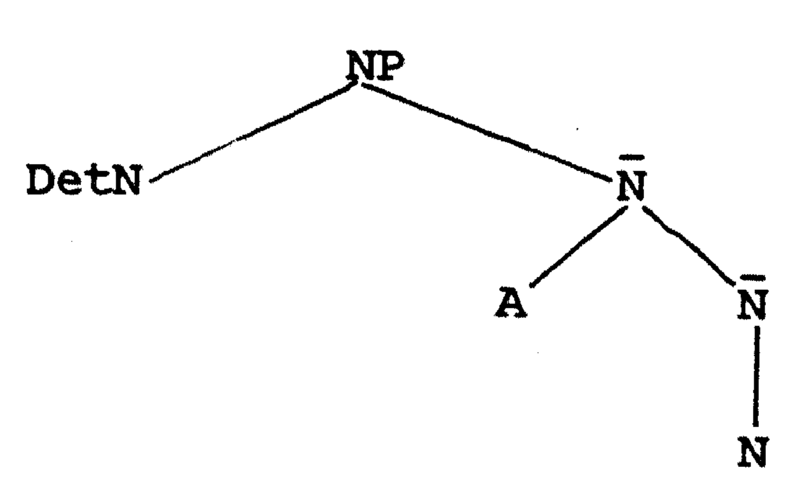
(iii) The semantic interconnections between comparatives containing *more/-er*, *less*, and *as* can be stated in a natural and revealing way.

(iv) The semantics is compatible with a concrete syntactic analysis, as demonstrated by the phrase structure treatments in GAZDAR (1980), KLEIN (forthcoming a).

In Sections 3 - 5, I argued that the heads of the other major comparative constructions -- those involving prenominal adjectives, *many* and *few* and adverbs -- could plausibly be interpreted as predicates and categorized as A. The analyses presented in each of these three sections are independently well-supported. They are substantially strengthened, I believe, by the fact that the corresponding comparative constructions can be subsumed under a single coherent theory.

FOOTNOTES

1. 'DetA' stands for Determiner of Adjective, by contrast with 'DetN', Determiner of Noun.
2. If VP is assigned the type $<\underline{g}(\mathrm{NP}),\underline{t}>$, as suggested by Montague in 'Universal Grammar' and subsequently advocated by KEENAN & FALTZ (1978), then one might want to assign this higher type to A and AP as well.
3. For a discussion of conventional metrics, see KLEIN (forthcoming a).
4. This assumption, which is probably too strong, can be dispensed with; however, it simplifies exposition considerably.
5. Q is a variable ranging over functions in $\{0,1\}^{(\underline{U})}$.
6. An alternative, possibly superior, is to assign the value 0 to all arguments outside $U(\underline{c})$.
7. I am assuming that the structure of NP is something like this:



8. This observation only holds if the context is held constant. One hundred people would count as many at a party, but not at a football match.
9. Here and in the sequel, X, Y are variables of type $<\underline{e},\underline{t}>$.
10. Each vertical column is an equivalence class under the relation 'exactly as many as'.

REFERENCES

BARTSCH, R., 1972, *Adverbialsemantik,* Athenaum, Frankfurt.

BARTSCH, R., 1973, 'The semantics and Syntax of Number and Numbers', in: J.P. Kimball (ed.), *Syntax and Semantics* 2, Academic Press, New York, 1973.

BARTSCH, R. & T. VENNEMANN, 1972, *Semantic Structures,* Athenaum, Frankfurt.

BOWERS, J.S., 1975, 'Adjectives and Adverbs in English', *Foundations of Language* 13, 529-562.

BRESNAN, J.W., 1973, 'Syntax of the Comparative Clause Construction in English', *Linguistic Inquiry* 4, 275-343.

CARDEN, G., 1970, 'On Post-Determiner Quantifiers', *Linguistic Inquiry* 1, 415-427.

CRESSWELL, M.J., 1974, 'Adverbs and Events', *Synthese* 28, 455-481.

CRESSWELL, M.J., 1979, 'Adverbs of Space and Time', in: F. Guenthner & S.J. Schmidt (eds), *Formal Semantics and Pragmatics for Natural Languages,* Reidel, Dordrecht, 1979.

DAVIDSON, D., 1967, 'The Logic of Action Sentences', in: N. Rescher (ed.), *The Logic of Decision and Action*, University of Pittsburgh Press, Pittsburgh, 1967.

EMONDS, J.E., 1976, *A Transformational Approach to English Syntax,* Academic Press, New York.

GAZDAR, G., 1980, 'A Phrase Structure Syntax for Comparative Clauses', in: T. Hoekstra, H. van der Hulst & M. Moortgat (eds), *Lexical Grammar,* Foris Publications, Dordrecht, 1980.

HOGG, R., 1977, *English Quantifier Systems,* North-Holland, Amsterdam.

JACKENDOFF, R., 1968, 'Quantifiers in English', *Foundations of Language* 4, 422-442.

KAMP, H., 1975, 'Two Theories of Adjectives', in: E.L. Keenan (ed.), *Formal Semantics of Natural Language,* Cambridge University Press, Cambridge.

KEENAN, E.L. & L. FALTZ, 1978, *Logical Types for Natural Langue,* UCLA Occasional Papers in Syntax 3.

KLEIN, E.H., forthcoming a, 'A Semantics for Positive and Comparative Adjectives', to appear in *Linguistics and Philosophy*.

KLEIN, E.H., forthcoming b, *The Interpretation of Adjectival Comparatives*.

LAKOFF, G., 1972, 'Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts', *CLS* 8, 183-228.

McCONNELL-GINET, S., 1973, *Comparative Constructions in English*, Ph.D. Thesis, University of Rochester.

PARTEE, B.H., 1970, 'Negation, Conjunction and Quantifiers: Syntax vs. Semantics', *Foundations of language* 6, 153-165.

PARSONS, T., 1972, 'Some Problems Concerning the Logic of Grammatical Modifiers', in: D. Davidson & G.H. Harman (eds), *Semantics of Natural Language*, Reidel, Dordrecht.

REICHENBACH, H., 1947, *Elements of Symbolic Logic*, MacMillan, New York.

RICHARDS, B., 1976, 'Adverbs: From a Logical Point of View', *Synthese* 32, 329-372.

SEUREN, P.A.M., 1973, 'The Comparative', in: F. Kiefer & N. Ruwet (eds), *Generative Grammar in Europe*, Reidel, Dordrecht, 1973.

SIEGEL, M., 1979, 'Measure Adjectives in Montague Grammar', in: S. Davis & M. Mithun (eds), *Linguistics, Philosophy and Montague Grammar*, University of Texas Press, Austin, 1979.

THOMASON, R.H. & R.C. STALNAKER, 1975, 'A Semantic Theory of Adverbs', *Linguistic Inquiry* 4, 195-220.

WHEELER, S., 1972, 'Attributives and Their Modifiers', *Nous* 6, 310-334.

ADAPTATION OF MONTAGUE GRAMMAR TO THE REQUIREMENTS OF PARSING

by

Jan Landsbergen

The paper describes a variant of Montague grammar, of which the composition rules have analytical counterparts on which a parsing algorithm can be based. Separate attention is given to the consequences of including rule schemes and syntactic variables in the grammar.

1. INTRODUCTION

In this paper a variant of Montague grammar, called M-grammar, is developed, for which an effective parsing procedure can be designed.

By "Montague grammar" I mean the formalism for defining the syntax and semantics of both natural and formal languages that is described in Montague's paper "Universal Grammar" (MONTAGUE, 1970), henceforth UG. The grammar of the English fragment in "The Proper Treatment of Quantification in Ordinary English" (MONTAGUE, 1973), henceforth PTQ, is a well-known application of this formalism. Montague grammar is especially attractive because of its elegant way of defining an interpretation for a natural language, which it does by means of a syntax-directed translation into a logical language for which an interpretation is defined directly. This makes Montague grammar suitable, in principle, for application in question-answering systems, where the translation from a natural language into a logical language is a useful intermediate step (cf., for instance, BRONNENBERG et al., 1980).

However, Montague grammar is a generative formalism. It generates natural language sentences and their logical forms "in parallel". In a question-answering system a parser is needed, i.e. an effective procedure which assigns to an input question the syntactic derivation tree(s) required for the translation into the logical language. For applications like this the framework of UG has to be adjusted in such a way that for the grammars within the modified framework a parser can be designed.

Obviously such a parser would also be a useful testing tool during the development of a large grammar.

Although the present paper mainly deals with syntactic problems, it gives due attention to the conditions that have to be fulfilled in order to maintain the systematic correspondence between syntax and semantics.

The paper is organized as follows.

In Section 2 M-grammars are defined. The rules of an M-grammar, called M-rules, are compositional rules which construct a labelled bracketing, called S-tree, from "smaller" S-trees and ultimately from basic expressions. In having rules operating on labelled bracketings instead of strings, as in PTQ, I follow the proposal in PARTEE (1973). As Montague's operations on strings can always be reformulated in terms of labelled bracketings, this is not a restriction. The domain in which the M-rules operate, i.e. the set of labelled bracketings that are allowed as possible inputs and outputs of these rules, is defined separately, by means of a context-free grammar. The language defined by the M-grammar is a subset of the language defined by this context-free grammar.

Each compositional M-rule has to obey the condition that it has a unique analytical counterpart. In Section 3 an analytical version of an M-grammar is defined on the basis of these analytical rules. It is proved that the compositional and the analytical version of the grammar define the same language.

Section 4 describes a parsing algorithm based on the analytical version of the grammar. This parser operates as follows. First a context-free parser generates the "surface trees" for the input string, according to the context-free grammar. Then the analytical M-rules are applied to each surface tree.

In Section 5 the framework of M-grammar is extended by the introduction of rule schemes, abbreviations for a possibly infinite set of rules. A modified version of the parser which is able to deal with these rule schemes is defined.

An important feature of Montague grammar is the use of syntactic variables (he_0, he_1, etc.). They are important tools for attaining a systematic relation between syntax and semantics, but they cause problems for a parsing system. These problems and their solution are discussed in Section 6.

In Section 7 a comparison is made with Friedman and Warren's parser for PTQ and Petrick's parser for Transformational Grammar.

2. M-GRAMMARS

In this section I will define M-grammars and compare them with Montague grammars as described in UG and PTQ.

An M-grammar is a triple $<G_s,B,R>$, where G_s is a loop-free[1] context-free grammar, B is a set of basic expressions, R is a set of M-rules.

The context-free grammar G_s is a quadruple $<V_N,V_T,S,P>$, where V_N is a set of syntactic categories, V_T is a set of terminal symbols, S is a distinguished syntactic category and P is a set of production rules. G_S defines a set L_s of labelled ordered trees, called S-trees, in the usual way:

- a node labelled by a terminal symbol is an S-tree.
- if $A \rightarrow B_1 \ldots B_n$ is a production rule and $t_1,\ldots,t_n$ are S-trees with top nodes labelled $B_1,\ldots,B_n$ respectively, then $A[t_1,\ldots,t_n]$ is an S-tree.

I use the labelled bracketing notation $A[t_1,\ldots,t_n]$ for a tree with top node labelled A and $t_1,\ldots,t_n$ as immediate sub-trees. The set of basic expressions B is a subset of L_s. An M-rule $R_i \in R$ has the form of a condition-action pair: $<C_i,A_i>$, where C_i is a predicate on n-tuples of S-trees $<u_1,\ldots,u_n>$ (if R_i is an n-ary rule). A_i is a function applicable to the n-tuples of S-trees for which C_i holds; $A_i(<u_1,\ldots,u_n>)$, i.e. the result of applying function A_i to $<u_1,\ldots,u_n>$, is an S-tree. C_i and A_i must be finitely characterizable.

In order to make parsing possible, the M-rules must obey two conditions:

<u>CONDITION C1</u>. For each composition M-rule $R_i = <C_i,A_i>$ there is an analytical counterpart $R_i' = <C_i',A_i'>$ where C_i' is a predicate applicable to S-trees and A_i' is a function from S-trees to n-tuples of S-trees, such that

(i) C_i' and A_i' can be expressed by means of effective procedures;

(ii) for all $u_1,\ldots,u_n$, $v \in L_s$: $C_i(<u_1,\ldots,u_n>)$ and $v = A_i(<u_1,\ldots,u_n>)$ if and only if $C_i'(v)$ and $<u_1,\ldots,u_n> = A_i'(v)$.

Condition C1 requires that each M-rule has a unique inverse rule. This is a severe condition. In several cases where, from a purely generative point of view, only one rule is needed, we need a rule *scheme* in M-grammars (cf. Section 5).

<u>CONDITION C2</u>. There is a measure for S-trees, i.e. a function μ from S-trees to non-negative integers, such that for each rule r_i holds:
if $C_i(<u_1,\ldots,u_n>)$ and $v = A_i(<u_1,\ldots,u_n>)$, then for each u_k $(1 \le k \le n)$: $\mu(v) > \mu(u_k)$.

An example of a measure is the number of nodes of an S-tree. Condition C2 requires in that case that the results of the application of a compositional rule is an S-tree that is bigger than any of the input S-trees.

In Montague grammar each rule operates on strings of specific syntactic categories and delivers a string of a specific category. In order to maintain this property for M-grammars, we define the syntactic category of an S-tree as the category of its top node and impose Condition C3.

<u>CONDITION C3</u>. For each n-ary rule R_i there are syntactic categories $P_1,\ldots,P_n$, $P_r \in V_N$ such that if $C_i(<u_1,\ldots,u_n>)$ holds, then the categories of the top nodes of $u_1,\ldots,u_n$ are $P_1,\ldots,P_n$, respectively, and the category of the top node of $A_i(<u_1,\ldots,u_n>)$ is P_r.

Because of Condition C3 each rule can be written in a way which is close to the notation in UG:

$$R_i = <<C_i,A_i>,<P_1,\ldots,P_n>,P_r>.$$

The desired systematic relation between syntactic and semantic rules can now be achieved in the usual way: there has to be a mapping ϕ from syntactic categories to semantic types and the semantic composition rule corresponding with R_i must be applicable to expressions of types $\phi(P_1),\ldots,\phi(P_n)$ and build a logical expression of type $\phi(P_r)$.

M-rules that satisfy Condition C3, still differ from the rules in UG in being *partial* rules.[2] The actions A_i only have to be applicable to S-trees that satisfy Condition C_i.

The M-rules are the actual rules of the grammar. They are composition rules, which build an S-tree, starting from the basic expressions. The context-free grammar G_s defines the domain in which the M-rules operate.

The way in which the M-rules construct an S-tree can be indicated by means of a derivation tree, or D-tree. A D-tree is a labelled ordered tree with basic expressions at the terminal nodes and indices i of M-rules R_i at the non-terminal nodes. I will use the notation $i<d_i,\ldots,d_n>$ for a derivation tree with index i at the top node and immediate sub-trees $d_1,\ldots,d_n$.

With each S-tree v that can be generated by the M-rules a D-tree corresponds which shows the derivational history of v. There may be more than one D-tree for v; in that case v is ambiguous with respect to the M-rules.

It should be noted that the applicability of an M-rule depends only on the input S-trees, not on their D-trees.

The set of D-trees of an arbitrary S-tree v is defined as follows:

$$\begin{aligned} \text{D-trees}(v) =_{\text{def.}}\ & \text{if } v \in B \text{ then } \{v\} \\ & \text{else } \{i\langle d_1,\ldots,d_n\rangle \mid \exists u_1,\ldots,u_n \in L_s,\ R_i \in R: \\ & C_i(\langle u_1,\ldots,u_n\rangle) \wedge v = A_i(\langle u_1,\ldots,u_n\rangle) \\ & \wedge \forall j: 1 \le j \le n \Rightarrow d_j \in \text{D-trees}(u_j)\} \end{aligned}$$

Less formally: if v is a basic expression, v is its own derivation tree. The derivation trees of an S-tree v that is not a basic expression have a top node labelled with an index i of an M-rule and $d_1,\ldots,d_n$ as immediate sub-trees. In that case there must be a tuple of S-trees $\langle u_1,\ldots,u_n\rangle$ such that $d_1,\ldots,d_n$ are their D-trees and such that M-rule R_i is applicable, delivering the S-tree v.

Let $\sigma(v)$ be the terminal string of v: the sequence of terminal symbols at the terminal nodes of the S-tree v. The language L defined by an M-grammar is the set of terminal strings of S-trees with top node S that have a derivation tree.

$$\begin{aligned} L =_{\text{def.}} \{\phi \mid \exists v \in L_s: \phi = \sigma(v) \wedge \text{D-trees}(v) \neq \emptyset \wedge \\ \wedge \text{top-node-cat}(v) = S\}. \end{aligned}$$

<u>Example of a simple M-grammar</u>

$$G_M = \langle G_s, B, R\rangle.$$

The contex-free grammar G_s is defined as follows:

$V_N = \{t, T, IV\}$
$V_T = \{\text{John, Mary, and, walk, talk}\}$
The distinguished symbol is t.
The production rules are:

$$\begin{aligned} t &\to T + IV \\ T &\to T + \text{and} + T \\ T &\to \begin{Bmatrix}\text{John}\\ \text{Mary}\end{Bmatrix} \\ IV &\to \begin{Bmatrix}\text{walk}\\ \text{talk}\end{Bmatrix} \end{aligned}$$

The basic expressions of the M-grammar are:

$$B = \{T[\text{John}], T[\text{Mary}], IV[\text{walk}], IV[\text{talk}]\}.$$

The set of M-rules R consists of two rules: R1 and R2. In this paper no specific notation for the M-rules is prescribed. I will explain the notation used in the examples in an ad hoc manner.

RULE R1. $C_1(<u_1,u_2>) =_{def.} u_1 = T[T[\ ], and, T[\ ]] \wedge u_2 = IV[\ ]$.

Here IV[ ] is the notation for an arbitrary S-tree with category IV at the top. So Rule R1 is applicable to a pair $<u_1,u_2>$ if u_1 is of category IV.

$A_i(<u_1,u_2>) =_{def.} t[u_1,u_2]$.

So the result of A_1 is an S-tree with category t at the top and u_1, u_2 as immediate sub-trees.

RULE R2. $C_2(<u_1,u_2>) =_{def.} u_1 = T[\ ] \wedge u_2 = T[\ ]$

$A_2(<u_1,u_2>) =_{def.} T[u_1, and, u_2]$.

The context-free grammar generates S-trees like t[T[John], IV[walk]] (fig. 1), t[T[T[John], and, T[Mary]], IV[walk]] (fig. 2). The S-tree of fig. 1 cannot be generated by the M-rules. The S-tree of fig. 2 can be generated by the M-rules and the corresponding derivation tree is 1<2<T[John], T[Mary]>, IV[walk]> (fig. 3).

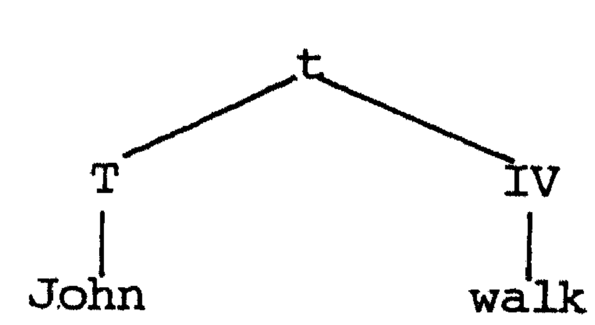


fig. 1

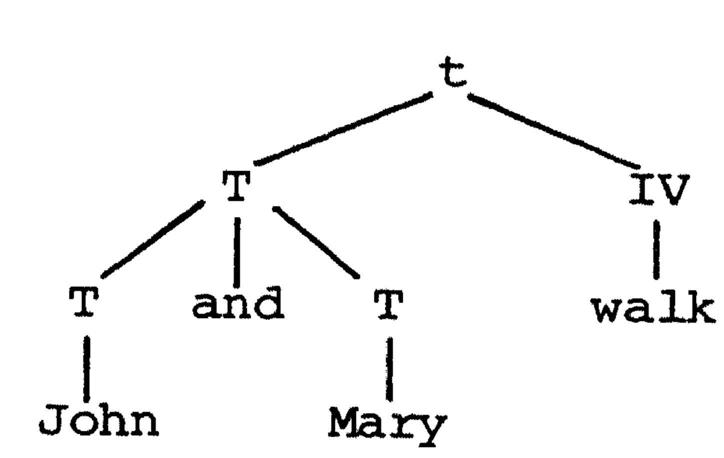


fig. 2

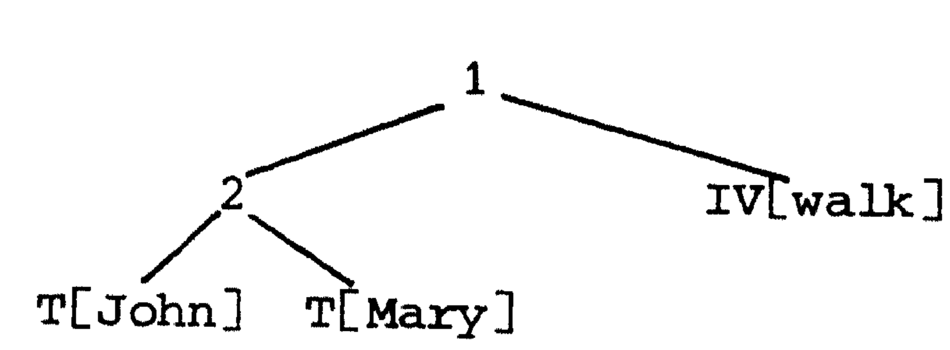


fig. 3

It can easily be checked that Rules R1 and R2 obey Conditions C1, C2 and C3, if we choose the number of nodes as the measure function.

The analytical versions of Rules R1 and R2 are:

RULE R1'. $C_1'(v) =_{def.} v = t[T[T[\ ], and, T[\ ]], IV[\ ]].$

$A_1'(v) =_{def.} \langle v.1, v.2\rangle.$

Here $v_j.i$ is the notation for the i-th immediate sub-tree of v_j.

RULE R2'. $C_2'(v) =_{def.} v = T[T[\ ], and, T[\ ]]$

$A_2'(V) =_{def.} \langle v.1, v.3\rangle.$

3. THE ANALYTICAL VERSION OF AN M-GRAMMAR AND ITS EQUIVALENCE TO THE COMPOSITIONAL VERSION

Because the rules of an M-grammar have to satisfy Condition C1, there is an analytical Rule $R_i' = \langle C_i', A_i'\rangle$ corresponding with each composition rule $R_i = \langle C_i, A_i\rangle$. If R' is a set of such analytical rules we will call $\langle G_s, B, R'\rangle$ an analytical M-grammar. An M-grammar as defined in Section 2 will be called a compositional grammar. For each compositional grammar there is an analytical version.

The set of derivation trees that an analytical grammar assigns to an S-tree is defined as follows:

$$\begin{aligned}\text{D-trees}_{an}(v) =_{def.}\ & \text{if } v \in B \text{ then } \{v\}\\ & \text{else } \{i\langle d_1,\dots,d_n\rangle \mid \exists u_1,\dots,u_n \in L_s,\ R_i' \in R':\\ & \quad C_i'(v) \wedge \langle u_1,\dots,u_n\rangle = A_i'(v) \wedge\\ & \quad \forall j: 1 \le j \le n \Rightarrow d_j \in \text{D-trees}_{an}(u_j)\}.\end{aligned}$$

The definition suggests how the derivation trees of v can be found by top-down application of the analytical rules. In Section 4 a parser will be described based upon this definition. In this section I prove the equivalence of the analytical and the compositional version of an M-grammar.

I will call the two versions equivalent if they assign to each string the same set of derivation trees. As both versions use the same context-free grammar G_s, it is sufficient to prove that they assign to any S-tree v the same set of derivation trees.

LEMMA 1. $\forall v \in L_s$: $d \in \text{D-trees}(v) \Rightarrow d \in \text{D-trees}_{an}(v)$.

PROOF. We use induction on the number of nodes in d: #nodes(d).

1. #nodes(d) = 1.
If a derivation tree d in D-trees(v) consists of one node, v must be a basic expression. In that case

$$\text{D-trees}(v) = \text{D-trees}_{an}(v) = v.$$

So the theorem holds for all d with #nodes(d) = 1.
2. Assume that the theorem holds for all d with

$$\#\text{nodes}(d) \leq k.$$

3. Proof for #nodes(d) = k + 1.
d has more than one node and therefore must have the form $i\langle d_1,\dots,d_n\rangle$.
According to the definition of D-trees(v):

$$(1)\qquad \exists u_1,\dots,u_n \in L_s,\ R_i \in R \text{ such that } C_i(\langle u_1,\dots,u_n\rangle),$$
$$v = A_i(\langle u_1,\dots,u_n\rangle) \text{ and } d_j \in \text{D-trees}(u_j) \quad (1 \leq j \leq n).$$

From $C_i(\langle u_1,\dots,u_n\rangle)$ and $v = A_i(\langle u_1,\dots,u_n\rangle)$ it follows (Condition C1) that $\exists R_i' \in R'$ such that $C_i'(v)$ holds and $\langle u_1,\dots,u_n\rangle = A_i'(v)$. Because #nodes(d) = k + 1, #nodes(d_j) ≤ k. According to the induction hypothesis: $d_j \in \text{D-trees}(u_j) \Rightarrow d_j \in \text{D-trees}_{an}(u_j)$. So from (1) we can deduce (2).

$$(2)\qquad \exists u_1,\dots,u_n \in L_s,\ R_i' \in R' \text{ such that } C_i'(v),$$
$$\langle u_1,\dots,u_n\rangle = A_i'(v) \wedge d_j \in \text{D-trees}_{an}(u_j), \quad (1 \leq j \leq n). \qquad \square$$

From (2) it follows immediately that $d \in \text{D-trees}_{an}(v)$.

LEMMA 2. $\forall v \in L_s$: $d \in \text{D-trees}_{an}(v) \Rightarrow d \in \text{D-trees}(v)$.

PROOF. Completely analogous to the proof of Lemma 1. □

THEOREM. $\forall v \in L_s$: $\text{D-trees}(v) = \text{D-trees}_{an}(v)$.

PROOF. The theorem follows immediately from Lemma 1 and Lemma 2. □

4. A PARSER FOR M-GRAMMARS

On the basis of the definition of $\text{D-trees}_{an}(v)$, a procedure M-PARSER can be designed which assigns to an S-tree v its set of derivation trees. I present here the main structure of this procedure. It is so close to the definition of D-trees_{an} that one can trust that it indeed delivers the set of D-trees defined by the analytical M-grammar. The proof of Section 3 guarantees that this is the set of D-trees defined by the compositional M-grammar.

```
M-PARSER(v):
begin
S_D := Ø;
if v ∈ B
then S_D := {v}
else for each analytical rule R'_i ∈ R' do
     if C'_i(v)
     then begin
         <u_1,...,u_n> := A'_i(v);
         for each tuple <d_1,...,d_n> ∈
         M-PARSER(u_1) × ... × M-PARSER(u_n) do
         S_D := S_D ∪ {i<d_1,...,d_n>}
         end;
M-PARSER := S_D
end
```

M-PARSER applies the analytical M-rules to the S-tree v in a top-to-bottom fashion. S_D is the set of D-trees, originally empty. Successful application of a rule R'_i to v results in a tuple $\langle u_1,\ldots,u_n\rangle$. M-PARSER is then applied to $u_1,\ldots,u_n$. Each application of M-PARSER to a u_j gives a (possibly empty) set of D-trees for u_j. For each tuple of D-trees $\langle d_1,\ldots,d_n\rangle$ in the cartesian product of these sets, a D-tree $i\langle d_1,\ldots,d_n\rangle$ is constructed and added to S_D. M-PARSER comes to a successful end if it is ultimately, at the deepest level of recursion, applied to basic expressions.

The procedure M-PARSER is effective, i.e. it ends after a finite number of steps. This follows immediately from the fact that C'_i and A'_i are effective procedures and that each application of $A'_i(v)$ results in a tuple of S-trees with a measure smaller than $\mu(v)$. Because of this the maximal recursion depth of M-PARSER cannot be more than the measure of the S-tree to which it is applied.

I assume here that the number of rules R'_i is finite. In Section 5 I will discuss rule schemes, which may define an infinite number of rules.

M-PARSER is not yet a complete parser. It has to be preceded by an ordinary context-free parser, called CF-PARSER, which assigns to a string s all S-trees that have s as their terminal string and a top node labelled S. Thanks to the requirement that the context-free grammar is loop-free, there is always such a parser. The complete algorithm is:

PARSER(s):

<u>begin</u> $S_D := \emptyset$;

<u>for</u> <u>each</u> $v \in$ CF-PARSER(s) <u>do</u>

$S_D := S_D \cup$ M-PARSER(v);

PARSER := S_D

<u>end</u>

5. RULE SCHEMES

The parsing procedure of Section 4 is only effective if the number of M-rules is finite. In this section I will define rule schemes, abbreviations of a (possibly) infinite set of rules, and describe a parser for an M-grammar with such rule schemes. In Montague grammars rule schemes occur: "rule" S3 of PTQ, for instance, is in fact a rule scheme, with an instance for each variable index. In M-grammars rule schemes are needed more often than in purely generative grammars, because of the condition that for each single rule there is a unique inverse rule.

A rule scheme is a triple $S = \langle P,I,A\rangle$.

- P is a possibly infinite set of parameter values.
- I is a function from n-tuples of S-trees $\langle u_1,\ldots,u_n\rangle$ to subsets of P, for an n-ary rule scheme.
- A is a function with two arguments: a parameter value p and an n-tuple of S-trees $\langle u_1,\ldots,u_n\rangle$; the result of the application of A is an S-tree.

P, I and A must be finitely characterizable.

In an ordinary M-rule $\langle C,A\rangle$, condition C decides whether or not the rule is applicable to a particular tuple $\langle u_1,\ldots,u_n\rangle$ of S-trees. A rule scheme $\langle P,I,A\rangle$ may be applicable in several ways and each parameter value in $I(\langle u_1,\ldots,u_n\rangle)$ determines one way in which the rule scheme is applicable.

Each rule scheme $\langle P,I,A\rangle$ defines a set of M-rules $\{\langle C_p,A_p\rangle \mid p \in P\}$, where

$$C_p(\langle u_1,\ldots,u_n\rangle) =_{\text{def.}} p \in I(\langle u_1,\ldots,u_n\rangle)$$
$$A_p(\langle u_1,\ldots,u_n\rangle) =_{\text{def.}} A(p,\langle u_1,\ldots,u_n\rangle).$$

A rule $\langle C_p,A_p\rangle$ is called an instance of the rule scheme.

Rule schemes have to obey three conditions: CS1, CS2 and CS3.

CONDITION CS1. For each rule scheme $S = \langle P,I,A \rangle$ there is an analytical rule scheme $S' = \langle P,I',A' \rangle$, such that

- I' is a function from S-trees to *finite* subsets of P.
- A' is a function with two arguments: a parameter $p \in P$ and an S-tree; the value of A' is an n-tuple of S-trees.
- I' and A' can be expressed by means of effective procedures.
- $\forall u_1,\ldots,u_n,\ v \in L_S$:
  $p \in I(\langle u_1,\ldots,u_n \rangle) \wedge v = A(p,\langle u_1,\ldots,u_n \rangle)$ if and only if
  $p \in I'(v) \wedge \langle u_1,\ldots,u_n \rangle = A'(p,v)$.

CONDITION CS2. Condition C2 must hold for all instances of the rule scheme.

CONDITION CS3. Condition C3 must hold for all instances of the rule scheme.

An analytical rule scheme $\langle P,I',A' \rangle$ defines a set of analytical M-rules $\{\langle C'_p,A'_p \rangle \mid p \in P\}$, where

$$C'_p(v) =_{\text{def.}} p \in I'(v)$$

$$A'_p(v) =_{\text{def.}} A'(p,v).$$

From these definitions and Condition CS1 it follows immediately that each instance $\langle C_p,A_p \rangle$ obeys Condition C1 and that $\langle C'_p,A'_p \rangle$ is its analytical counterpart.

The conclusion can be that from an M-grammar with rule schemes a compositional and an analytical M-grammar, as defined in Sections 2 and 3, can be derived. The functions D-trees and D-trees$_{an}$ can be expressed in terms of these derived M-grammars. The equivalence proof of Section 3 is valid for M-grammars with an infinite set of rules and therefore is also valid for grammars with rule schemes.

The procedure M-PARSER of Section 4 has to be adjusted. For a grammar with a finite number of rule schemes S'_i it becomes:

```
M-PARSER'(v):
begin
S_D := Ø;
if v ∈ B
then S_D := {v}
else for each rule scheme S'_i do
```

```
        for each parameter value p ∈ I'_i(v) do
        begin
        <u_1,...,u_n> := A'_i(p,v);
        for each tuple <d_1,...,d_n> ∈ M-PARSER'(u_1) × ... × M-PARSER'(u_n) do
        S_D := S_D ∪ {i_p <d_1,...,d_n>}
        end;
    M-PARSER' := S_D
    end
```

The effectiveness of M-PARSER' can be established in a similar way as was done for M-PARSER, taking into account that there is a finite number of rule schemes S_i and that $I'_i(v)$ is a finite set.

The correctness of M-PARSER' can be established in the same way as the correctness of M-PARSER, if we realize that all parameter values $p \in I'_i(v)$ specify exactly the applicable instances of the rule schemes S'_i.

The parser has been described in terms of rule schemes only. This is no restriction, because an M-rule can always be considered as a special case of a rule scheme, where the set P of parameter values has only one element.

Example

I will now describe an M-rule scheme for PTQ rule S4. The original rule is:

RULE S4. If u_1 is an expression of category T and u_2 is an expression of category IV, then $u_1u'_2$ is an expression of category t, where u'_2 is the result of replacing the first basic verb in u_2 by its third person singular present.

This rule cannot be represented by a single M-rule, because it does not have a unique inverse rule. The point is that the first basic verb in the IV might, in principle, be preceded by a verb that is already in third person singular present form.

The M-rule scheme for S4 is the triple $<P_4,I_4,A_4>$, where P_4 is the set of all possible paths in S-trees. A path is a sequence of branches from the top to a sub-tree. As each branch can be represented by a positive integer (n represents the branch to the n-th daughter), a path can be represented by a sequence of positive integers $i_1...i_n$. So P_4 is the set of all such sequences. We define:

$$i_1 \ldots i_m < j_1 \ldots j_n \text{ if } \exists k\ (1 \le k \le n): i_1 = j_1, \ldots, i_k = j_k, \; i_{k+1} < j_{k+1}.$$

$$I_4(<u_1,u_2>) =_{def.} \{p \mid u_1 = T[\ ] \wedge u_2 = IV[\ ] \wedge u_2.p \text{ is the first basic verb in } u_2\}.$$

For a given pair $<u_1,u_2>$ there is at most one such path p, which illustrates that from a purely generative point of view a rule scheme is not needed.

$A_4(<u_1,u_2>) =_{def.}$ $t[u_1,u_2']$ where u_2' is the result of replacing in u_2 the basic verb $u_2.p$ by its third person singular present.

The inverse rule scheme is $<P_4, I_4', A_4'>$.

$I_4'(v) =_{def.}$ $\{p \mid v = t[T[\ ], IV[\ ]] \wedge v.2.p$ is a verb in third person singular present form $\wedge \neg\exists p_i < p$ such that $v.2.p_i$ is a basic verb$\}$.

The definition of I_4' shows that if, for instance, v.2 contains exactly two verbs, both in third person singular present form, two instances of the inverse rule scheme are applicable.

$A_4'(v) =_{def.}$ $<v.1, v_2'>$, where v_2' is the result of replacing in v.2 the verb v.2.p by its basic form.

6. SYNTACTIC VARIABLES

An important feature of Montague grammar is the use of syntactic variables. The PTQ grammar for instance, contains an infinite number of variables he_0, he_1, he_2, etc. as basic expressions.[3] It also contains rules that eliminate a variable by substituting a term for it. PTQ rule S3 is such a 'variable-removing' rule; more precisely, it is a rule scheme, with instances for each variable. The problem with rule schemes like S3 is that they assign an infinite number of derivation trees to each sentence, most of them only differing in the choice of their variables.

If we assume that there is a systematic correspondence between syntactic variables and variables of the logic, as expressed by Janssen's Variable Principle (JANSSEN, 1980), we are able to partition the infinite set of derivation trees of a sentence into a finite number of equivalence classes. The members of an equivalence class of derivation trees differ only in the choice of their variables and, therefore, correspond with logical expressions that are logically equivalent. From this point of view only one 'canonical' derivation tree of each equivalence class is of interest. I will define this canonical derivation tree as follows.

First the *level* of a node in a D-tree is defined as the length of the path (the number of branches) from the top node to that node.

A *variable-removing node* in a D-tree is a node corresponding with a rule that eliminates some variable he_k. According to the Variable Principle the corresponding translation rule into the logic binds the corresponding logical variable.

A *canonical D-tree* is a D-tree where a variable-removing node at level N removes the variable he_N.[4]

Figure 4 shows one of the D-trees assigned by the rules of PTQ to the sentence 'John loves every woman'. Figure 5 shows the corresponding canonical D-tree. At the nodes of these D-trees the index of the corresponding PTQ rules and the removed variables are indicated (index 2^a stands for the first sub-rule of S2, where syntactic operation F_0 places 'every' before a CN).

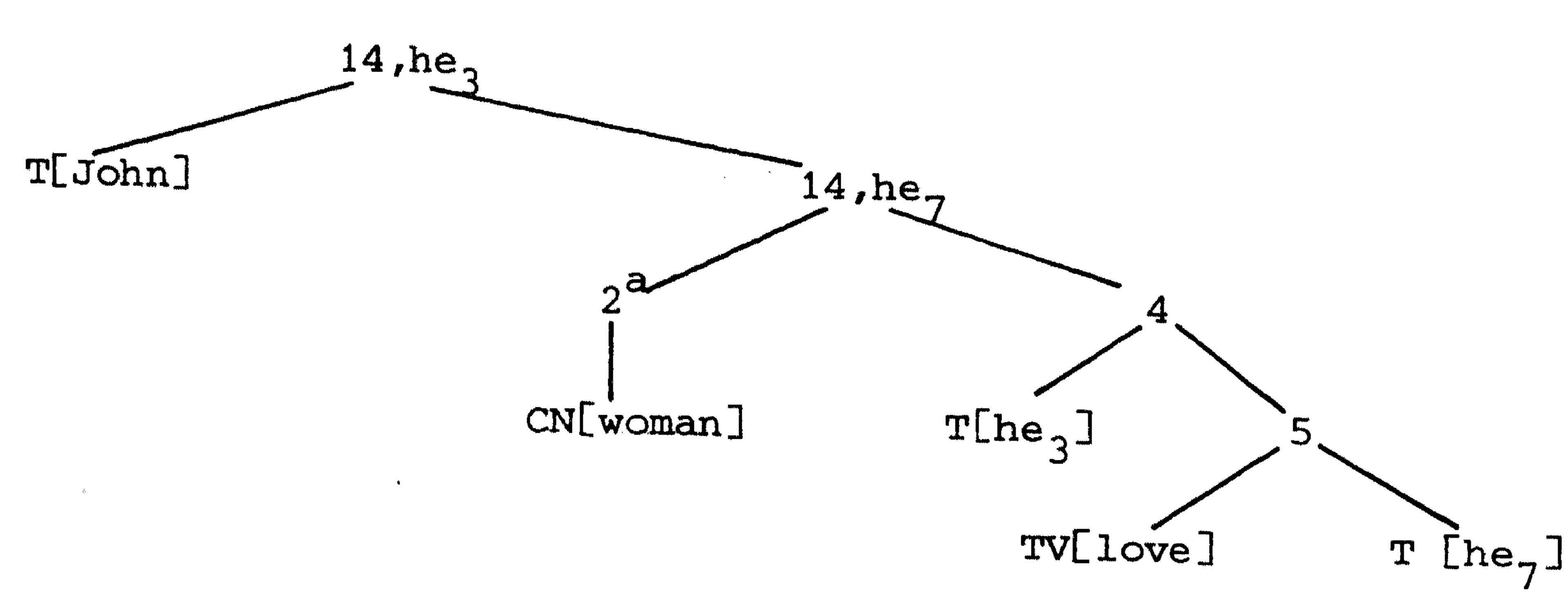


Figure 4

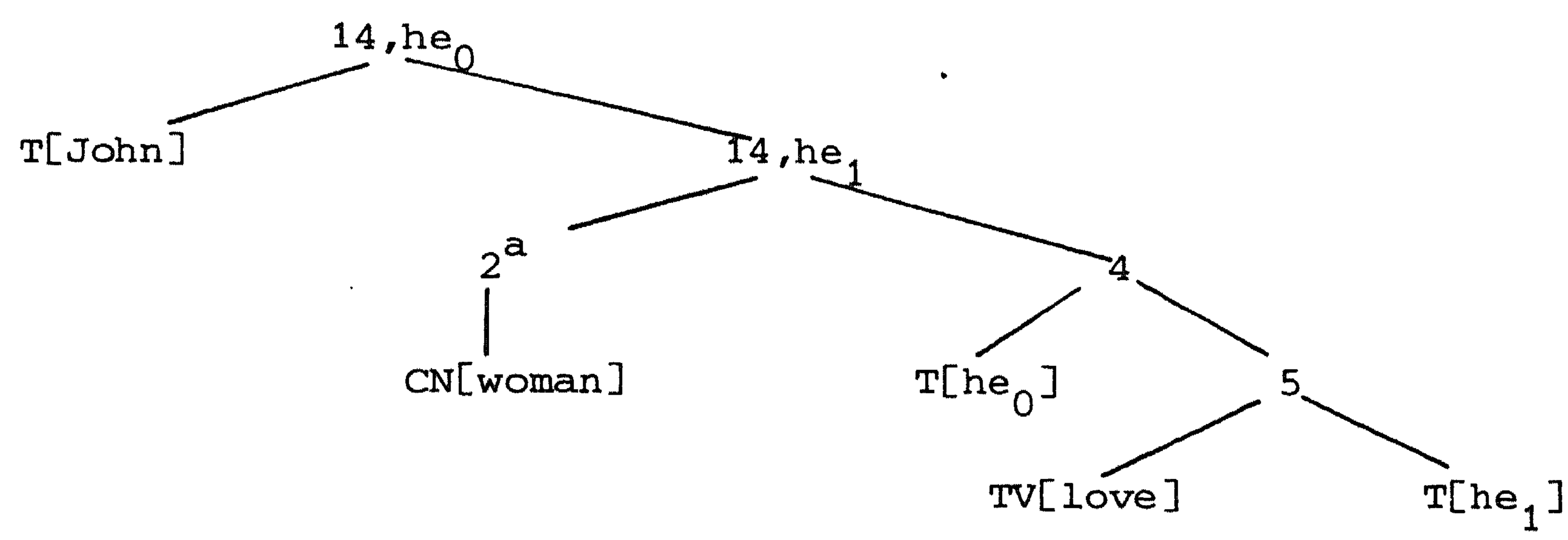


Figure 5

I will assume here, without proof, that each of the above-mentioned equivalence classes of D-trees contains one canonical D-tree. In that case a parser only needs to generate canonical D-trees. The assumption is only correct if each variable-removing rule is an instance of a rule scheme with instances for each variable. Let $\langle P, I', A' \rangle$ be the analytical version of such a rule scheme (each instance is a 'variable-introducing' rule).

We will replace I' by a function I", which has a second argument, N. Now, I"(v,N) does not give all parameter values for v, but only those that are relevant for the rule instances that introduce variable he_N. For analytical rule schemes that do not introduce a variable, I"(v,N) is equal to the original I'(v).

If the parser tries to construct a new node of the derivation tree at level N_1 by applying rule scheme $\langle P, I', A' \rangle$ it has to call $I''(v, N_1)$. This can be achieved by giving the new parsing procedure, M-PARSER", the level N as a second argument. If the parsing process is started by calling M-PARSER"(v,0), only canonical derivation trees are generated.

M-PARSER"(v,N):
begin
$S_D := \emptyset$;
if $v \in B$
then $S_D := \{v\}$
else *for* *each* rule scheme S_i' *do*
    for *each* parameter value $p \in I_i''(v,N)$ *do*
    begin
    $\langle u_1, \ldots, u_n \rangle := A_i'(p,v)$;
    for *each* tuple $\langle d_1, \ldots, d_n \rangle \in$ M-PARSER"$(u_1, N+1) \times \ldots \times$ M-PARSER"$(u_n, N+1)$

```
        do S_D := S_D ∪ {i_p <d_1,...,d_n>}
        end;
   M-PARSER" := S_D
   end
```

The complete parser becomes:

```
   PARSER"(s):
   begin S_D := ∅;
         for each v ∈ CF-PARSER(s) do
             S_D := S_D ∪ M-PARSER"(v,0);
         PARSER" := S_D
   end
```

The effectivity of M-PARSER" follows immediately from the effectivity of M-PARSER', and the fact that $I_i'(v,N)$ delivers a finite set of parameter values.

However, M-PARSER' and M-PARSER" are only effective if a measure can be defined such that each rule satisfies Condition C2. The number of nodes is not an adequate measure here, because rules that substitute a single term ('John' for example) for a variable, like PTQ rule S14(i), would not increase the number of nodes. Therefore, we choose the number of 'non-variable' nodes as the measure. This still precludes rules like PTQ rule (scheme) S14(ii), which only substitutes a variable for another variable. These rules are superfluous from a semantic point of view. They seem to have been included by Montague in order to make S14 a total rule. This measure also precludes the 'vacuous' applications of S14 that arise when a term is substituted for a variable that is not present in the expression. FRIEDMAN & WARREN (1978) already noticed that these applications of the rule lead to incorrect translations into the logic.

Example

As an example of a rule scheme involving syntactic variables I will give the M-grammar formulation of PTQ rule scheme S3. The original PTQ rule is:

RULE S3. If u_1 is an expression of category CN and u_2 is an expression of category t, then "u_1 such that u_2'" is an expression of category CN, where u_2' comes from u_2 by replacing each occurrence of he_n or

him_n by he, she, it or him, her, it, respectively, according as the first basic expression of category CN in u_1 is of masc., fem. or neuter gender.

The M-rule scheme for S3 is $<P_3, I_3, A_3>$, where

$$P_3 = \{<g,Q,n> \mid g \in \{\text{masc., fem., neuter}\},\ Q \text{ is a set of paths},\ n \text{ is a variable index}\}.$$

Each parameter value is a triple consisting of a gender, a set of paths and a variable index. (It is not absolutely necessary to include the gender in the parameter, but it facilitates the formulation of the rule scheme.)

$$I_3(<u_1,u_2>) =_{def.} \{<g,Q,n> \mid n \geq 0,\ g \text{ is the gender of the first terminal CN in } u_1,\ Q = \{p \mid u_2.p = he_n \vee u_2.p = him_n\}\}.$$

So for a given pair $<u_1,u_2>$ there is a parameter value $<g,Q,n>$ for each variable index n, where Q is the set of all paths from the top of u_2 to variables with this index. Obviously, $I_3(<u_1,u_2>)$ is an infinite set, which contains an infinite number of triples of the form $<g,\emptyset,n>$, where n is the index of a variable not occurring in u_2, and $\emptyset$ is the empty set.

$A_3(<g,Q,n>, <u_1,u_2>) =_{def.}$
CN[u_1, such that, u_2'], where u_2' is the result of replacing in u_2 for each $p_i \in Q$ the variable at $u_2.p_i$ by the pronoun of the appropriate case and gender.

The inverse rule scheme is $<P_3, I_3', A_3'>$, where

$I_3'(v) =_{def.} \{<g,Q,n> \mid$ v = CN[CN[ ], such that, t[ ]] $\wedge$ g is the gender of first terminal CN in v.1 $\wedge$ v.3 does not contain variables with index n $\wedge$ Q is a subset of the set of paths to pronouns with gender g in v.3$\}$.

$A_3'(<g,Q,n>,v) =_{def.}$ $<u_1,u_2>$, where u_1 = v.1 and u_2 is the result of replacing in v.2 for all $p_i \in Q$ the pronoun v.2.p_i by the variable he_n or him_n according to the case of the pronoun.

$I'_3(v)$ is an infinite set. The subset of parameter values relevant for the rule instances that introduce variable he_N or him_N is defined as:

$$I'_3(v,N) =_{def.} \{\langle g,Q,n\rangle \mid \langle g,Q,n\rangle \in I'_3(v) \wedge n = N\}.$$

7. CONCLUDING REMARKS

The notion 'M-grammar' as described in this paper meets the standard I formulated in the introduction: it is a variant of Montague grammar for which effective parsers can be designed, and which maintains the systematic relation between the syntactic rules and the translation rules into the logic.

As an exercise in the new formalism, the complete grammar of PTQ has been rewritten in the form of an M-grammar and the corresponding parser has been programmed.

It may be interesting to compare the parser described here with Friedman and Warren's parsing method for Montague grammars (FRIEDMAN & WARREN, 1978), though this comparison is made difficult because of an important difference between their approach and mine. F. & W.'s parser is more or less tuned to PTQ, while my goal was primarily to define a class of grammars for which the same kind of parsing algorithm can be used. F. & W. describe the implementation of the individual PTQ rules in detail, whereas I confine myself to the global algorithmic structure of the program. A major problem with F. & W.'s parser - based upon an Augmented Transition Network representation - is that it is difficult to establish its correctness. Especially after reading the passage in F. & W.'s paper (pp.366-368) about the way in which false parsers are avoided in the case of terms nested in other terms, one is left with the uneasy feeling that other complicated cases might have been overlooked.

In the case of M-grammars the parser can be derived systematically from the analytical version of the grammar, which can be proved to be equivalent to the original compositional version. The equivalence of the compositional and the analytical version of each individual rule must be checked, but one does not have to bother about the interaction with other rules. Therefore, it is possible to have confidence in the correctness of the parser. The price paid for this is that there is some redundancy in the grammar and accordingly in the parser. If there are M-rules that are in fact context-free (as is the case for several PTQ rules), they are duplicated in the context-free grammar.

Another comparison that can be made is with Petrick's recognition procedure for Transformational Grammar (PETRICK, 1965). In order to make effective recognition and parsing possible, Petrick imposes two conditions on TG:

1. a recoverability condition on transformations, comparable with but weaker than my Condition C1;
2. a condition on the depth of S-embedding in the deep structure.

The second condition has the same objective as my condition C2: to guarantee that the parsing procedure comes to an end after a finite number of steps. However, Petrick's second condition is not a restriction on the individual rules, but on the whole collection of rules.

Obviously, requiring that transformations make the syntactic trees bigger, according to some measure, would be unacceptable in TG. Transformations usually have a kind of paraphrasing character and may even involve deletions. In Montague grammar, the rules are basically compositional, they build a new expression from parts that are intuitively smaller. Condition C2 requires expressions that are intuitively smaller to be smaller in a technical sense as well. Though I follow Partee's suggestion to apply Montague rules to labelled trees instead of strings, I do not support her proposal to incorporate transformation rules in Montague grammar. It seems worth while to investigate first what is possible within the more restrictive framework of compositional rules.

ACKNOWLEDGEMENTS. I wish to thank W.J.H.J. Bronnenberg, H.C. Bunt, J.H.G. Rous, R.J.H. Scha and editor T.M.V. Janssen for their comments on an earlier version of this paper. J.H.G. Rous programmed the PTQ parser.

FOOTNOTES

1. A context-free grammar is called loop-free if derivations of the form $A \overset{*}{\Rightarrow} A$ are not possible.
2. If the rules are total, i.e. applicable to all expressions of the required syntactic categories, the translation into the logic can be defined in an elegant way, i.e. as a homomorphism between algebras. This makes it possible to prove that an interpretation of the logic (again a homomorphism into an algebra) induces an interpretation of the natural

language. Therefore, it may be interesting to note that total rules can always be derived from the partial rules of M-grammar in a trivial way, by redefining the set of syntactic categories:

- Introduce a category $P(u_k)$ for each S-tree u_k. S-tree u_k is the one and only S-tree of category $P(u_k)$.
- Replace each rule R_i by an infinite number of new rules, as follows: if $C_i(<u_1,\dots,u_n>)$ holds, then
  $<A_{i,P(u_1),\dots,P(u_n)}, <P(u_1),\dots,P(u_n)>, P(u_r)>$
  is a new, total, rule, where
  $A_{i,P(u_1),\dots,P(u_n)}(<u_1,\dots,u_n>) =_{def.} A_i(<u_1,\dots,u_n>)$.

On the basis of these rules a many-sorted algebra of derivation trees can be defined, with a sort for each category $P(u_i)$ and operations $A_{i,P(u_1),\dots,P(u_n)}$. The translation into the logical language can then be defined as a homomorphism from this algebra into an algebra of logical expressions.

If total rules are preferred for other reasons (cf. JANSSEN, 1978), M-grammar allows them, of course, as special cases of partial rules.

3. The introduction of an infinite set of variables is not in conflict with the condition that the set of possible S-trees must be defined by means of a context-free grammar, with a finite set of terminals. The variables are basic expressions, but not necessarily terminals. They may be compound S-trees, defined with the help of context-free rules.
4. I assume here, for the sake of simplicity, that all variables have the same syntactic category (as in PTQ).

REFERENCES

BRONNENBERG, W.J.H.C., BUNT, H.C., LANDSBERGEN, S.P.J., SCHA, R.J.H., SCHOENMAKERS, W.J. & E.P.C. VAN UTTEREN, 1980, 'The question-answering system PHLIQA 1', <u>in</u>: L. Bolc (ed.), *Natural Language Question-Answering Systems*, Carl Hanser Verlag, München & Wien, 1980, pp.217-305.

FRIEDMAN, J. & D.S. WARREN, 1978, 'A parsing method for Montague grammars', *Linguistics and Philosophy*, <u>2</u>.

JANSSEN, T.M.V., 1980, 'On problems concerning the quantification rules in Montague grammar', <u>in</u>: G. Rohrer (ed.), *Time, Tense and Quantifiers*, Max Niemeyer Verlag, Tübingen, 1980, pp. 113-134.

JANSSEN, T.M.V., 1978, 'Compositionality and the form of the rules in Montague grammar, in: J. Groenendijk & M. Stokhof (eds), *Amsterdam papers in formal grammar II, Proceedings of the second Amsterdam colloquium on Montague grammar and related topics,* Amsterdam papers in formal grammar 2, Centrale Interfaculteit Univ. of Amsterdam, 1978, pp.211-234.

MONTAGUE, R., 1970, 'Universal grammar', reprinted in: R.H. Thomason (ed.), *Formal Philosophy,* Yale University Press, New Haven, 1974, pp.222-246.

MONTAGUE, R., 1973, 'The proper treatment of quantification in ordinary English', reprinted in: R.H. Thomason (ed.), *Formal Philosophy,* Yale University Press, New Haven, 1974, pp.247-270.

PARTEE, B.H., 1973, 'Some transformational extensions of Montague grammar', reprinted in: B.H. Partee (ed.), *Montague Grammar,* Academic Press, New York, 1976, pp.51-76.

PETRICK, S.R., 1965, *A recognition procedure for transformational grammars,* Ph.D. Thesis, MIT, Cambridge, Mass.

AN INTENSIONAL LOGIC FOR MASS TERMS

by

Alice ter Meulen

1. INTRODUCTION

In this paper I present a formal semantics for mass terms in natural language, which is based directly on their syntax. I claim not only that a typed intensional logic with the usual set-theoretic interpretation provides a sufficiently rich model-structure to account for the interpretation of mass terms, but also that semantical theories of mass terms based on mereological set-theory fail to account for some characteristic intensional aspects of the interpretation of mass terms, and for some phenomena connected with their syntactic and semantic properties.

A fundamental distinction is made between *nominal* mass terms and *predicative* mass terms, supported by syntactic and semantic evidence. Since the semantic interpretation presented in this paper is based directly on the syntactic structure of the sentences with mass terms, no primitive relations or special notions are used that do not correspond to the syntactic constituents of these sentences. In this respect the present analysis is strictly compositional, which distinguishes it from many other recent proposals for the semantics of mass terms.

Generalized quantifiers are introduced, providing a general and linguistically motivated account of quantification, including non-logical quantifiers. They also serve to maintain at the logical level the distinction between quantities and individuals, both basic entities, without resorting to a logic with multi-sorted variables.

The intensional logic used in the interpretation of mass terms contains an S4-relation on reference points, which is interpreted as *change of situations*. Without modifying much in the syntax of the formal language of PTQ, a gain of expressive power is thus obtained, which is claimed to meet the requirements of the interpretation of natural language more adequately than the model-theory of Montague's PTQ. Meaning-postulates are formulated to restrict the models, as required by the logical properties of the tation of mass terms (e.g. homogeneous reference).

2. NOMINAL AND PREDICATIVE MASS NOUNS

Mass nouns occur in natural language in two syntactically and se tically distinguished ways. Consider the following sentences.

(1) Gold is an element.

(2) My tooth is filled with gold.

In (1) 'gold' is a mass noun that behaves in syntactic and semantic re like a proper name of an abstract entity. A mass noun used in this way a name is called a *nominal* mass noun, and the abstract entity which it denotes the *substance*. In (2) 'gold' is a mass noun that functions mor like a predicate. It takes quantifiers and various modifiers and denot a *set of quantities* of the substance. What quantities are will be disc in more detail below. A mass noun so used is called a *predicative* mass Often predicative mass terms do not contain an explicit determiner wit predicative mass noun. They are called *bare predicative mass terms*, an they must be distinguished clearly from nominal mass terms.

The distinction between nominal and predicative mass nouns is fun tal to the interpretation of mass nouns in this paper. Nominal mass no can neither syntactically nor semantically, be reduced to predicative c nor vice versa. The relation between the two uses of mass nouns, i.e. between substances and their quantities, is to be explained in the mode theory as a relation between their respective denotations.

Two syntactic arguments supply evidence supporting this distinctio between nominal and predicative mass nouns. Terms constructed from nomi mass nouns exhibit anaphoric behaviour similar to that of proper names allowing for *backwards pronominalization*.[1] Predicative mass terms do no typically allow for a coreferential reading with a preceding anaphor wi the same clause. Compare the sentences in the following sets.

(3) Sarah betrays the man she loves.
Water is defined by its chemical formula.

(4) The man she loves betrays Sarah.
Its chemical formula defines water.

The sentences in (3) and (4) allow for a coreferential reading of the pronoun and the proper name or nominal mass term. Quantified noun-phrases, including predicative mass terms, do not generally allow for such backwards pronominalization within the same clause. For example, the sentences in (5) allow for a coreferential reading of the indefinite quantified noun-phrase and the pronoun occurring after it, whereas the sentences with backwards pronominalization in (6) lack such a reading.

(5) Some women betray the men they love.
Some gold is sold by the person who finds it.

(6) The men they love are betrayed by some women.
The person who finds it sells some gold.

The sentences in (6) are relatively difficult to give a coreferential interpretation of the pronoun and the indefinite quantified noun-phrase occurring after it.[2] But the sentences in (5) can easily be given a coreferential interpretation.

This brief discussion of the evidence shows that only nominal mass terms allow like proper names for backwards pronominalization, whereas predicative mass terms, unless they are definite terms, do not allow a coreferential reading of a preceding anaphor.

The second syntactic argument supporting the distinction between nominal and predicative mass terms is based on the *type-restricted binding* of these expressions. Nominal mass terms bind only pronouns that are interpreted as denoting the substance, whereas predicative mass terms bind only pronouns, if they are interpreted as denoting a set of quantities. Consider the following sentences:

(7) Water is H_2O and it is muddy.
Ice is a solid and it is melting.

(8) Little gold is mined in Alaska and it is an element.
Some water is muddy and it is H_2O.

The anaphoric binding in these sentences is deviant, since substances, being abstract entities, cannot be muddy or melting, and sets of quantities cannot have properties of substances, namely being an element or having a certain atomic number. An account for the incorrect binding in (7) and (8) follows from the distinction between nominal and predicative mass terms.

The grammar does not generate sentences like these in the syntax, and hence they are not even interpretable.[3]

Two semantic arguments support the distinction between nominal and predicative mass terms. Theories of mass terms that fail to make the distinction necessarily fail to account for these facts.

Nominal mass terms denote the same substance at all reference-points in the same interpretation. I.e. nominal mass terms are, just like proper names, *rigid designators*.[4] Once the reference of a nominal mass terms is fixed in an interpretation, it remains the same at all reference-points. What water is in the sense of what the substance water is, is once and for all established in an interpretation of the language. But what water is, in the sense of what quantities of water there are, is a contingent matter. The denotation of the predicate mass term changes from one reference-point to another, as usual for non-logical predicates. In other words, our concept of 'water', for example, is fixed, but what water there is changes all the time. Substances determine what their quantities are, but the quantities do not make up the substance. The nature of this relationship is explained in the present interpretation as a relation between an intensional object and its extension at the reference-point.

Only by distinguishing between nominal and predicative mass terms it can be explained how the substance is rigidly denoted, whereas the quantities of the substance are referred to contingently.

Consistent with the rigidity of nominal mass terms is the semantic fact that nominal mass terms usually take *wide scope* over other terms and intensional operators. But predicative mass terms preferably have *narrow scope* over quantifiers and intensional operators.
Consider the following examples:

(9) Every child drinks some milk.
Mary found little gold but Jane found it in abundance.

(10) Everyone knows that gold is an elementary substance.
Mary thinks that water is H_2O but Jane thinks it is something else.

The sentences in (9) do not have a wide scope reading of the predicative mass term over the other quantified noun-phrases. But the sentences in (10) show that nominal mass terms have wide scope over quantifiers and intensional operators. These scope differences could not be adequately explained if there would be only one category of mass terms.

The syntactic arguments about backwards pronominalization and the type-restricted binding together with the semantic arguments based on rigid designation of nominal mass terms and the scope differences of predicative and nominal mass terms support the distinction between the two uses of mass terms at the linguistic level. This linguistic distinction is reflected in the philosophical distinction between substances and sets of quantities of substances. These facts could not be accounted for if there would be only one category of mass terms in the grammar or if substances would be equated with sets of quantities. Further philosophical arguments against such an equation are given in the next section.

3. ARGUMENTS AGAINST REDUCTIONISTIC SEMANTICS FOR MASS TERMS

Most semantic theories of mass terms recognize only one category of mass terms. Hence only one interpretation is assigned to all mass terms.[5] Although it is often realized that mass terms can be used either quantified or without quantifier, most theories reduce the one kind to the other. To account for the difference philosophical tricks are devised that have no reflection in the syntax of natural language, and lead to an ontological bias that need not be assumed in the interpretation of mass terms. Either the substance is reduced to an individual with scattered parts (Quine), or some notion of fusion is employed that explains the substance as a mereological whole formed from the set of quantities.

Mereology, initially developed, like type-theory, to avoid the set-theoretic antinomies, is a set-theory that allows for the formation of an individual whole from a set of its parts, without ascension in type. The whole is the same kind of entity as the individual parts of it are. Substances are explained as such mereological wholes, or fusions, or the totality of the quantities of the substance. Quantities are part of the substance, the part-of relation being a relation between entities of the same ontological level. In this way substances are made to depend essentially on their quantities.

Interpretations of mass terms based on mereological set-theory assume just one syntactic category of mass terms, if they take syntax into consideration at all, and interpret all occurrences of mass terms on a par.

By not distinguishing between the two uses of mass terms, on a syntactic and a semantic level, these theories fail to account for the syntactic and semantic differences between nominal and predicative mass terms that I discussed above.

They often require primitive relations in their formal semantics, relating substances to quantities with the part-of relation, or relating individuals to quantities they are constituted by. A ring of gold is analyzed as a ring that is constituted by a quantity of gold. The substance gold is the whole of which all quantities of gold are part. Although the various theories that give a reductionistic semantics for mass terms differ in detail on what primitives are assumed, most of them introduce primitive constants in the semantics that have no counterpart in the syntax of mass terms. It is not clear how many distinct primitives are needed, since individuals and quantities can have various relations to each other. The quantity can be constitutive of the individual, as in 'the ring is gold', but in 'hamburgers are food' the relation between the individuals and the quantities of food is clearly not constitutive, whatever it is. Another problem that a mereological semantics faces is to give a mathematically coherent account of the connection between the regular set-theory, used for the interpretation of count terms and the mereology.

This leads to an interpretation of mass terms that has only an indirect relation to the syntactic structure in which mass terms occur. An account of the interpretation of mass terms without such primitive unnatural notions, constructing the interpretation as an exact map on the syntactic trees is definitely preferable. My account of the semantics of mass terms attempts to adhere to such central methodological principles of formal semantics and Montague grammar.

As I said above arguments can be given against treating substances as sets of quantities or as special constructions based on sets of quantities. First, this would make reference to the substance as contingent as reference to the sets of its quantities, conflicting with the rigidity of nominal mass terms, that I briefly argued for above. But also in the case of coinciding sets of quantities, i.e. in the case in which all gold has been used to make jewelry, no distinction could be made between the substance gold and the substance jewelry. Similarly, if there were no water, there would be no substance water on such a mereological account of substances. This is particularly counterintuitive, since properties of substances can continue to be true of the substance, even when there happen to be no quantities of the substance. Water is H_2O even when there is no water available.

Semantic theories of mass nouns that are based on mereological set-theory take it to be a central question what the parts of the mereological whole are, whether there are any smallest parts and how a largest part (i.e. the whole) is constructed. The largest quantity is the fusion of all quantities to the mereological whole. Parts have a homogeneous structure, i.e. they consist of smaller parts of the same whole. This characteristic property of the denotation of mass nouns I will call, in accordance with established practice, *the property of homogeneous reference*. The present interpretation of mass terms takes this property also as distinctive of predicative mass nouns, but it is imposed on the interpretation by a meaning-postulate. To take this property of predicative mass nouns as the foundation of a set-theory, i.e. mereology, is, I argue, not only needed for the interpretation of mass terms, but it also leads to the wrong kinds of questions. Although quantities have a certain homogeneous structure, it is not this structure that is studied in the interpretation of mass terms in natural language. Exactly how complex a smallest quantity of a substance is, is a question of physics, just as the question by what processes larger wholes are formed. Mereological theories often assume an atomistic base, which is not required for the interpretation of mass terms.

Instead of constructing quantities and substances of such an atomistic base[6], I take quantities to be basic entities, that differ from individuals (i.e. the denotations of count nouns) only by having the property of homogeneous reference. In this way the strong parallel between count nouns and mass nouns is preserved, and there is no need to introduce special relations between individuals and quantities, as they both have as intension a property and as extension a set of basic entities. The property which is the intension of the predicative mass noun, is the extension of the corresponding nominal mass noun. The intensional object of a property is thus referred to extensionally by the nominal mass noun. The intension of the nominal mass noun is consequently a second-order property. The central point in this construction is that *ex*tensional reference to *in*tensional objects becomes possible. There is no need to account for more of the structure of quantities; they can be simple or complex, whatever you like.

It is not impossible to enrich mereological theories to allow for such extensional reference to intensional objects. Any first-order set-theory can be generalized by the full theory of types. However, such a solution strikes me as most unnatural for the following reasons.

First of all, mereology was invented as a nominalist answer to the set-theoretic paradoxes. Leonard and Goodman, as well as other mereologists, had good, philosophical reasons to adhere to a strictly first-order set-theory. The addition of a type-theory would not only be completely contrary to such spirit, but it would also be superfluous from a mathematical point of view. Montague grammar, in its extreme Platonism strongly antagonistic to such nominalists bias, would, on the other hand, not be served by a mereological interpretation of its first-order theory for the various reasons I have already mentioned above. There is no need in the modeltheoretic semantics for mass terms to construct larger objects from smaller ones. The only way to demonstrate such a need is to give valid arguments that hinge on such mereological assumptions, and that would be invalid in a regular set-theoretic interpretation restricted by a meaning-postulate for the property of homogeneous reference. I fail to see that there is such an argument.

4. GENERALIZED QUANTIFIERS IN NATURAL LANGUAGE

Predicative mass nouns denote sets of quantities, whereas regular count nouns denote sets of individuals. So the basic entities in the model-theoretic semantics are either quantities or individuals. Keeping to the methodological principles for the semantics mentioned earlier, this distinction should not be introduced in the formal language as a logic with two-sorted variables. Such a solution would be misguided and unnecessary, since it is a syntactic solution to what is essentially a semantic distinction. The logical language needs to have only one kind of individual variables, if it imposes the homogeneous reference of the denotation of predicative mass nouns by a meaning-postulate. If the distinction between individuals and quantities is imposed on the interpretation by such meaning-postulates, no multisorted logic is needed.

However, at the formal level a distinction must be made between properties of quantities and properties of individuals. E.g. inferences like 'this table is of wood', 'the wood is bought in a store' to 'the table is bought in a store' must be blocked. Furthermore, the fact that the interpretation of non-logical determiners like 'much' and 'little' (or their count equivalents 'many' and 'few') depend upon the noun with which they occur in the term must be accounted for. What is much or little can only be determined if we know what the relevant subset of the domain is that is being referred to. Given a certain contextually and conventionally determined norm and an intended domain, we can say truly or falsely that much water is polluted, for instance. Such essential dependency of a determiner upon the reference of the noun can be found in many other non-logical determiners, e.g. 'more than half', 'almost as much', 'infinitely many' that have no first-order representation. For example, for the sentence 'more than half the people in this room read Synthese' to be true, it does not matter what is the case in the entire domain of the interpretation, but it matters only what the people in this room read. Only the familiar logical quantifiers 'every' and 'some' do not depend upon such a subset of the domain, but upon the entire domain.

The insight that quantification in natural language depends only on subsets of the domain, in contrast to quantification in standard first-order predicate logic, forms the foundation of the theory of generalized quantifiers.[7] This theory provides a nice and general framework for the interpretation of quantifiers in natural language. Because it is essentially higher-order, it accounts in a general way for the compositional dependency between quantifiers and nouns in terms. In natural language the entire terms function as quantifying elements. Terms consisting of a noun and a determiner are the constituents in the syntax that bear scope and bind anaphora. Nouns are translated in the formal language to predicates, which are interpreted as functions having sets of entities as their values. Determiners or quantifiers, which are syntactically basic lexical items, are translated to functions that take the noun-phrase sets as arguments and yield families of such sets. The compositionality of terms is thus accounted for in treating all terms as generalized quantifiers.

In Montague's PTQ terms denote functions from reference-points to sets of sets of individuals. Their denotation is based on the denotation of the predicate in the term, i.e. a set of individuals, that is the translation of the noun. Determiners are introduced syncategorematically. They are not basic expressions of any syntactic category, and hence they are not given any interpretation in the model-theory by themselves. Instead, they are introduced in the application of syntactic rules, which receive special translations in the corresponding translation rules. In EFL (MONTAGUE, 1974 chapter 6) determiners are of the same category as adjectives.[8] In UG (MONTAGUE, 1974 chapter 7) determiners are again syncategorematic expressions. Several fragments[9] have been developed in which determiners are basic expressions of the lexicon that have a separate interpretation. The question whether to treat determiners syncategorematically or as lexical items is usually considered to be a merely theoretical decision on grounds of simplicity and convenience. The interpretation of terms with non-logical quantifiers as generalized quantifiers, however, makes empirical factors bear on this issue. Only in a formal semantics restricted to the interpretation of just logical quantifiers is the choice of syncategorematic or basic lexical treatment of quantifiers arbitrary.

In treating determiners syncategorematically in PTQ Montague gave up the means to analyze the compositional nature of the interpretation of terms. The semantics of generalized quantifiers assumes, on the contrary, that the interpretation of terms is fundamentally compositional. The interpretation of a term is based on the interpretation of the noun and on the interpretation of the determiner. The composition of the interpretation of a determiner and a noun to obtain the interpretation of a term is simply functional application. So determiners are treated as basic lexical expressions of a category that takes nouns and forms terms. Semantically determiners will then be translated into quantifiers which are interpreted by complex functions corresponding to this category.

Consider the following very simple sentences with quantified predicative mass terms.

Some gold shines

and

All gold shines.

These mass terms 'some gold' and 'all gold' are syntactically composed of the mass noun 'gold' and the determiners 'some' and 'all', respectively. Semantically these mass terms denote sets of sets of quantities of gold at a reference point. 'Some gold' denotes the set consisting of all sets that contain some quantities of gold at a reference point. 'All gold' is interpreted by the set consisting of all sets that have the set of all quantities of gold as subset. Note that the interpretation of the terms does not depend upon the entire domain, but only on subsets of the domain or on a set of such subsets.

In the theory of generalized quantifiers this dependency is captured by requiring that any non-logical determiner in a term introduces an interpretation that *lives on* the interpretation of the noun in that term. The notion of the interpretation of a term living on a set is defined as follows.

DEFINITION. An interpretation T' of a term T *lives on a set* A iff for any set X: $X \in T' \Longleftrightarrow X \cap A \in T'$.

This definition requires that the interpretation of the entire term depends only upon the interpretation of the noun in that term. What sets there are in the interpretation of the term depends only on what sets intersect with the set that is the interpretation of the noun in that term.

In set-theoretic notation we can characterize the interpretation of the mass terms in the above sentences as follows:

$$\text{some gold} - \{X \subseteq A \mid X \cap \{x \mid \text{gold}(x)\} \neq \emptyset\}$$

$$\text{all gold} - \{X \subseteq A \mid \{x \mid \text{gold}(x)\} \subseteq X\},$$

where A is the domain of the interpretation.

Informally the sentence 'some gold shines' is interpreted as true on a reference-point, if the set that contains all sets that have non-empty intersections with the set of all quantities of gold has also as member the set of things that shine. At the level of intensional denotations we would say that the property of shining is amongst the properties of the sets that contain some quantities of gold.

Similarly, the sentence 'all gold shines' is true if the set of sets containing all quantities of gold has the set of things that shine as a member. Or intensionally, the set of properties of the sets containing all quantities of gold has the property of shining as an element.[10]

This interpretation of terms and quantifiers is analogous to the analysis of proper names in Montague's PTQ. Proper names are interpreted as the set of all properties of the individual to which the proper name refers. A property is true of that individual if and only if that property is amongst the properties of that individual. A proper name is thus analyzed as a generalized quantifier. In fact, Montague showed in PTQ that generalized quantifiers provided a universal analysis of quantification in natural language. But by analyzing only the logical quantifiers, he failed to realize that the interpretation of most quantifiers of natural language depends essentially on the interpretation of the nouns in the terms. This is just what generalized quantifiers account for, although Montague never explicitly endorsed this. The main idea of the semantic interpretation of generalized quantifiers is that all terms, proper names or quantified expressions, denote sets of subsets of the domain and that properties of entities are in that set if and only if the entity has that property.

Another important point in favor of the analysis of terms as generalized quantifiers is that the so-called bare terms, like 'tables' in 'tables have four legs' or 'water' in 'I am drinking water' or 'gold' in 'this ring is gold', can all be treated as scope-bearing elements. Terms do not have to contain a determiner in order to bear scope and bind anaphora. All first-order theories of quantification had to introduce null-determiners or other theoretical postulates to account for this phenomenon. In the syntax a syntactic operation will be introduced that constructs terms of proper names, nominal mass nouns and predicative mass nouns. In this way a general account is given of the interpretation of such expressions as generalized quantifiers. To implement this theory of generalized quantifiers in a PTQ style fragment the following modifications must be made.

Since there is no need for individual concepts (the temperature puzzle will receive a different treatment), both count (CN) and predicative mass nouns (PMN) are of a basic syntactic category, with a corresponding semantic type $<e,t>$. Two distinct categories of determiners are needed for count nouns and for predicative mass nouns. These categories are defined respectively as CT/CN and as PMT/PMN, corresponding to the types $<<s,<e,t>>,<<s,<e,t>> t>>$ (functions taking properties to sets of properties). Using these categories the following syntactic rules are introduced:

<u>SYNTACTIC RULE 1</u>. If $\alpha \in P_{CT/CN}$ and $\beta \in P_{CN}$ then $F_1(\alpha,\beta) \in P_{CT}$, where $F_1(\alpha,\beta) = (\alpha\ \beta')$ and β' equals β adjusted for number according to the number of α.

<u>SYNTACTIC RULE 2</u>. If $\alpha \in P_{PMT/PMN}$ and $\beta \in P_{PMN}$ then $F_2(\alpha,\beta) \in P_{PMT}$, where $F_2(\alpha,\beta) = \alpha\ \beta$.

Correspondingly the following two translation rules are needed:

<u>TRANSLATION RULE 1</u>. If $\alpha \in P_{CT/CN}$ and $\beta \in P_{CN}$ then $F_1(\alpha,\beta)$ translates to

$$\lambda P[{}^\wedge(\alpha'({}^\wedge\beta'))\{P\}].$$

<u>TRANSLATION RULE 2</u>. If $\alpha \in P_{PMT/PMN}$ and $\beta \in P_{PMN}$ then $F_2(\alpha,\beta)$ translates to

$$\lambda P[{}^\wedge(\alpha'({}^\wedge\beta'))\{P\}].$$

The term which denotes a set of properties is applied to a property-variable, over which a lambda-abstraction is made to construct the appropriate expression. Separate rules are given for bare predicative mass terms, but limitations of space keep me from discussing these expressions. More illustrations of these rules are given in the last section.

5. THE INTENSIONAL LOGIC FOR MASS TERMS

The formal language and its interpretation that I use for the formal semantics of mass terms is in many respect similar to the language and interpretation of PTQ.

An important difference is that instead of the complex reference-points $\langle i,j \rangle$ of PTQ, consisting of a possible world and a moment of time, a simple notion of a *situation* is used. There is no need for complex reference-points, since as is obvious from the interpretation of the intensional logic in PTQ, no use is made of the possibility to keep the moment of time fixed, while varying the possible world. The interpretation of the intensional operators require only that the possible world is kept fixed, while allowing the moment of time to vary. In the interpretation of natural language there is no need for the full-fledged structure of the cardinal product of the set of moments of time and a set of possible worlds. The notion of a situation conflates the complex reference-point of PTQ, to a simple reference-point that does not have more structure than is required for the interpretation.

Another difference with PTQ is that a reflexive and transitive relation on the reference-points is introduced, which is interpreted as *change* of the situations. It simultaneously captures the temporal ordering on the reference-points, which was already given in PTQ, and a 'causal' connection between the situations. Change is described by verbs in natural language, and this intensional logic is intended to be useful for the interpretation of the various classes of verbs as well, but it is beyond the scope of this paper to discuss such an application.

The unnatural requirement on the interpretations in PTQ that they all have the same domain of entities is relaxed in the intensional logic for mass terms by introducing formal objects for each type, that are referred to by non-denoting singular terms and certain complex expressions in which they occur. This amounts to having a third truthvalue 'undefined', denoted by atomic sentences in which non-denoting expressions occur. It can thus be expressed that individuals or quantities come into existence and cease to exist, without using a philosophically suspect notion of a logically possible individual or quantity.

There are many arguments why an interpretation with no ordering on the reference-points (like in PTQ), or only an ordering on the moments of time (like in PTQ), or an ordering which is an equivalence relation on the reference-points (an S5-structure), is too general for the interpretation of natural language. For sake of brevity, I mention here only a few.

First of all, in natural language it is easy to find intuitively invalid arguments, that are valid in an S5-structure. For example:

(i) if one can get gas now, it is necessary that one can get gas now;

(ii) if the water from the well is potable, it will always be potable.

The invalidity is based on the fact that the present situation can change in the relevant respects, or that it could have been different. It is contingent to the present situation that one can get gas, or that the water from the well is drinkable. But in the present situation it is certainly possible that the well gets distorted and the water from it will be polluted. Or the present situation could have been such that there was no gas available.

Other examples of intuitively invalid inferences can be given that are all natural language renditions of the S5-axiom or the Barcan formula $\forall x[\Box\phi(x) \rightarrow \Box\forall x\ \phi(x)]$, which is valid in an S5-structure with constant domains. If these are valid, the kind of change described in natural language is excluded in the models.

More formally, an interpretation of the intensional logic is a quintuple $\langle A,S,R,C,F \rangle$, where A is the domain of entities, S the set of situations and R a relation assigning every situation s a subset of A, i.e. the entities that exist at s, C the reflexive and transitive change-relation on situations, and F the usual interpretation-function for constants. The language contains an operator $\boxed{C}$ for causal necessity, which is interpreted as truth at all situations that a given situation can change into.

Together with the formal objects as denotation for non-denoting terms and some complex expressions in which they occur, this intensional logic is a first step towards adapting and constraining the intensional logic of PTQ to fit the requirements of natural language interpretation more narrowly. In trying to understand how natural language interpretation differs from the more universal interpretation of formal languages, we might gain insight in what is characteristic of human language, without thereby giving up on a precise and formal semantics.

6. MEANING-POSTULATES AND EXAMPLES

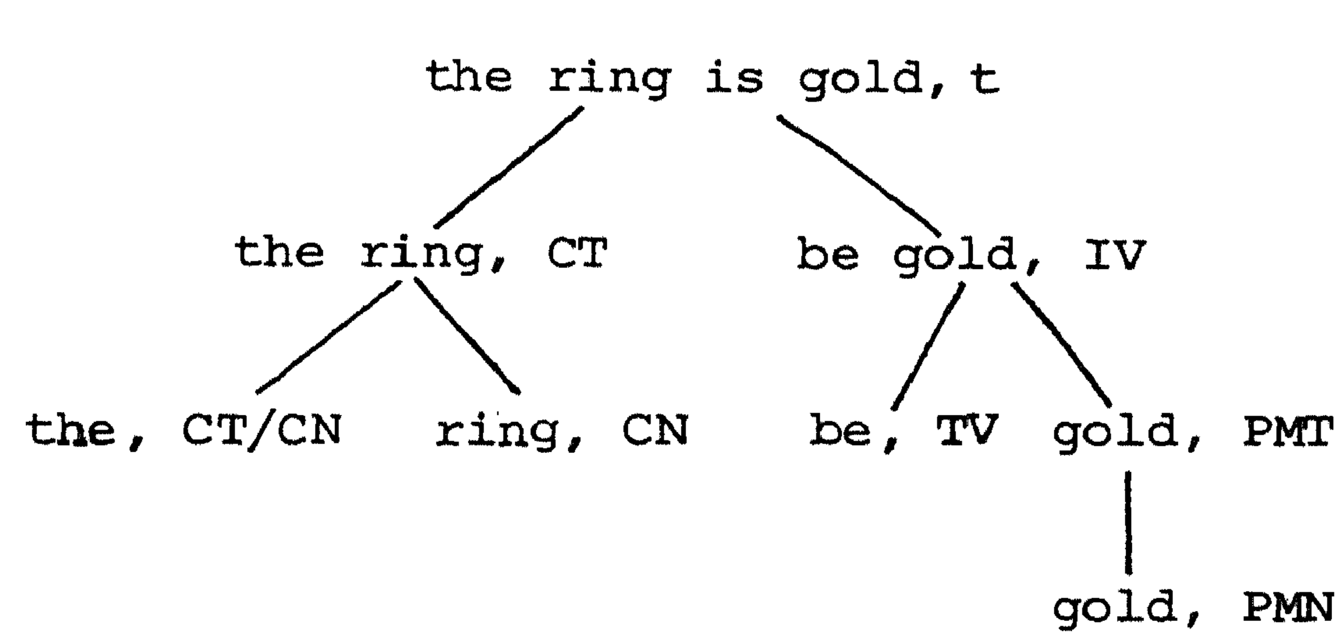
Although it is not possible within the limitations of this paper to present a full-fledged Montague grammar fragment with mass terms, I illustrate some of the main points of this paper by discussing some examples of translations and the meaning-postulates that characterize what are reasonable models for the interpretation of these expressions.[11] The central distinction between nominal and predicative mass terms comes out in the following two sentences.

The ring is gold

and

Gold is an element.

The first sentence has the following syntactic analysis tree:



The predicative mass noun 'gold' is made into a bare predicative mass term by a syntactic operation that does not modify the original expression, and the rest of the syntactic analysis is straightforward. The translation is obtained as follows. Since the rules are not explicitly stated, the types of the resulting expressions are given in the first steps in the translation for purposes of clarification.

$\text{ring} \rightsquigarrow \text{ring}' \; \langle e,t\rangle$

$\text{the} \rightsquigarrow \text{the}' \; \langle\langle s,\langle e,t\rangle\rangle,\langle\langle s,\langle e,t\rangle\rangle,t\rangle\rangle$

$\text{the ring} \rightsquigarrow \lambda P[{}^{\wedge}(\mathit{the}'({}^{\wedge}\mathit{ring}'))\{P\}] \; \langle\langle s,\langle e,t\rangle\rangle,t\rangle$

gold $\leadsto$ $gold_1'$ $\langle e,t\rangle$

gold $\leadsto$ $\lambda Q[{}^\wedge gold_2'\{Q\}]$ $\langle\langle s,\langle e,t\rangle\rangle,t\rangle$

Note that $gold_1' \neq gold_2'$, $gold_2'$ is the translation of the bare predicative mass term, and is of type $\langle\langle s,\langle e,t\rangle\rangle,t\rangle$. The process of obtaining $gold_2'$ from $gold_1'$ and its corresponding interpretation is not further specified here.

be $\leadsto$ $\lambda P\lambda x[P\{{}^\wedge\lambda y[x = y]\}]$ (as in PTQ)

be gold $\leadsto$ $\lambda P\lambda x[P\{{}^\wedge\lambda y[x = y]\}]({}^\wedge\lambda Q[{}^\wedge gold_2'\{Q\}])$,

which reduces to

$\lambda x[gold_2'({}^\wedge\lambda y[x = y])]$.

The ring is gold $\leadsto$

$\lambda P[{}^\wedge(the'({}^\wedge ring'))\{P\}]({}^\wedge\lambda x[gold_2'({}^\wedge\lambda y[x = y])])$

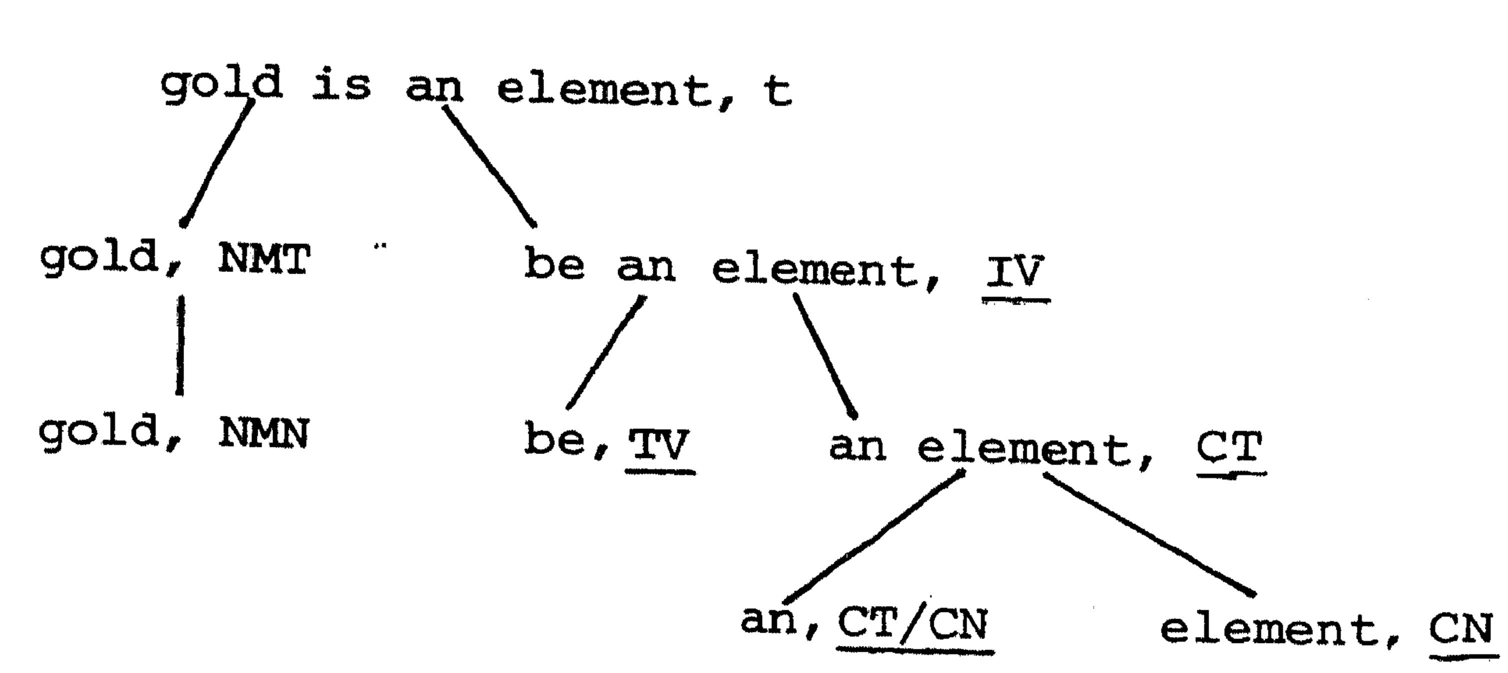
which reduces by λ-conversion and various extensionalization conventions to

$(the'(ring'))(\lambda x[gold_2'(\lambda y[x = y])])$

which says that the set of quantities of gold is a member of the set of sets that contain one ring.

Note that 'gold' as a predicative mass term is interpreted as a set of sets of quantities. 'The ring' is interpreted similarly as a set of properties, and the final translation states that 'being gold' is one of those properties of the ring.

The second sentence 'gold is an element' is translated into a second-order formula, since being an element is a property of substances, which are themselves first-order properties. The syntactic analysis is as follows:



By underlined categories is indicated that they are to be interpreted by second-order objects. I.e. besides the previously given syntactic rules 1 and 2 another rule for second-order term formation is needed.

<u>SYNTACTIC RULE 3</u>. If $\alpha \in P_{\underline{CT}/\underline{CN}}$ and $\beta \in P_{\underline{CN}}$ then $F_1(\alpha,\beta) \in P_{\underline{CT}}$, where $F_1(\alpha,\beta) = (\alpha\ \beta')$ and β' equals β adjusted for number according to the number of α.

<u>TRANSLATION RULE 3</u>. If $\alpha \in P_{\underline{CT}/\underline{CN}}$ and $\beta \in P_{\underline{CN}}$ then $F_1(\alpha,\beta)$ translates to

$$\lambda\mathcal{P}[{}^\wedge(\alpha'({}^\wedge\beta'))\{\mathcal{P}\}].$$

Similarly another rule is needed to construct nominal mass terms out of nominal mass nouns. Syntactically this construction does not alter anything, but semantically the property denoted by the noun is transferred to a set of properties (i.e. generalized quantifier). The formulation of such a rule is straightforward, as is the rule for forming sentences of second-order verbs and terms. The translation is as follows:

gold $\rightsquigarrow$ $gold'_3$ <s,<e,t>>

gold $\rightsquigarrow$ $\lambda\mathcal{P}\mathcal{P}\{gold'_3\}$ <<s,<e,t>>,t>

an element $\rightsquigarrow$ $\lambda Q[{}^\wedge(an'(element'))\{Q\}]$ <<s,<<s,<e,t>>,t>>,t>

be $\rightsquigarrow$ $\lambda\underline{P}\lambda P\underline{P}\{{}^\wedge\lambda Q[P=Q]\}$, where $\underline{P}$ is a third-order property-variable of type <s,<<s,<<s,<e,t>>,t>>,t>>,

be an element $\rightsquigarrow$ $\lambda\underline{P}\lambda P[\underline{P}\{{}^\wedge\lambda Q[P=Q]\}]({}^\wedge\lambda Q[{}^\wedge(an'({}^\wedge element'))\{Q\}])$,

which reduces by two conversions to

$\lambda P[(an'({}^\wedge element'))(\lambda Q[P=Q])]$.

gold is an element $\rightsquigarrow$ $\lambda PP\{gold'\}({}^\wedge\lambda P[(an'({}^\wedge element'))(\lambda Q[P=Q])])$,

which reduces by two conversions and extensionalization conventions to

$((an'(element'))(\lambda Q[gold'_3 = Q]))$.

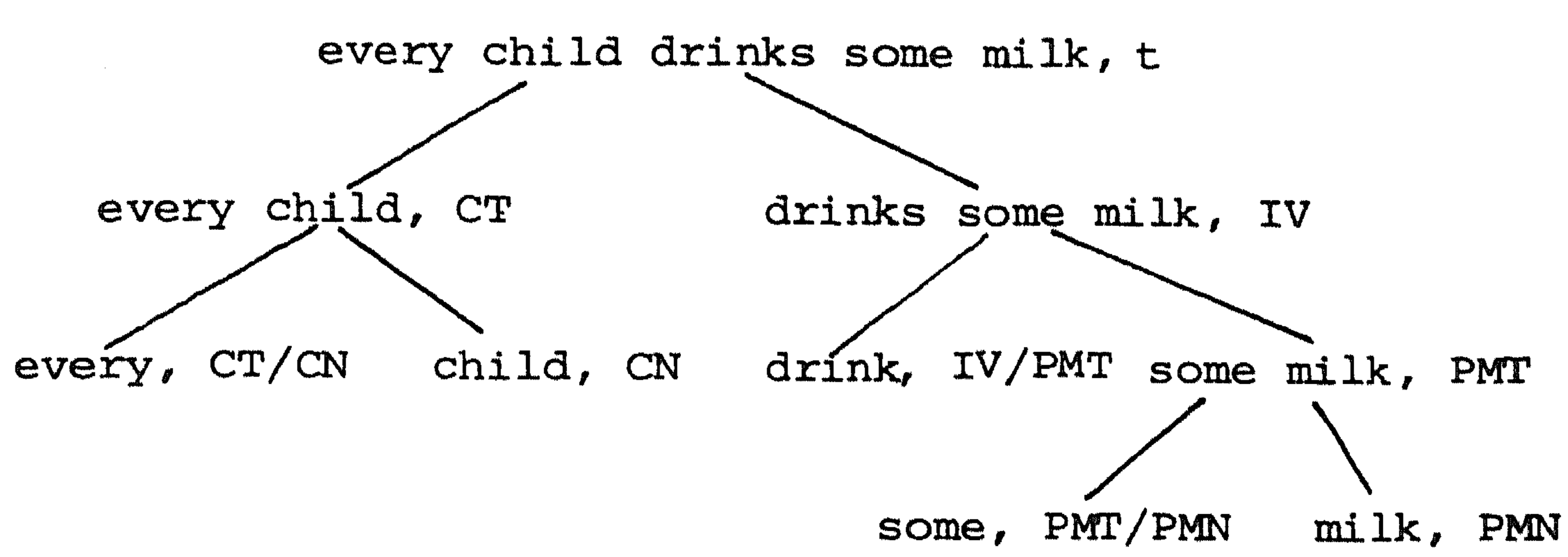
This expresses that the second-order property of being an element is a property of the first-order property gold.

Although these translations are not very transparent, it should be clear that the intension of the predicative mass noun gold is extensionally denoted by the nominal mass noun gold. Instead of constructing the nominal mass noun denotation out of the predicative mass noun denotation, I introduce a new name for the intension of the predicative mass noun, which is the translation of the nominal mass noun. There are various advantages to this explicit introduction of the denotation of nominal mass nouns. First of all, it captures the fundamental point that a concept (the substance) is not the same as the set of all possible entities that fall under the concept (the set of all logically possible quantities). Since it is impossible in set-theoretic possible world semantics to distinguish the substance water for example, from all the quantities of water in all possible worlds (i.e. the intension of the predicative mass noun 'water'), a constant of type $<s,<e,t>>$ is needed that will function as a name for the intension. This will also allow for two different names for two substances that have no quantities in any possible world, i.e. two fictional substances that have various properties, like being a metal or being an elementary substance, although the corresponding predicative mass nouns are interpreted by the same function (from possible worlds to the empty set). If a property or an intension is interpreted as a set of extensions in all possible worlds, as is commonly done, and the extension of the nominal mass noun is this property, the substance is again made to depend essentially on what happen to be the quantities in the possible worlds. This is what I have argued against at length. Furthermore, predicative mass terms in intensional constructions are to be interpreted as being about quantities, possible quantities, but not about the substance. E.g. in seeking gold we do not try to find the substance, or in thinking that the water in the well is polluted, we have no thoughts about any substance at all. A direct construction cannot account for this, and furthermore it would necessitate *ad hoc* restrictions on such constructions of nominal mass noun denotation out of the predicative mass noun denotation, since there are many complex predicative mass nouns that do not have an associated nominal mass noun that rigidly denotes the property which is the intension of the predicative mass noun. For example, gold that weights two ounces does not correspond to a rigidly denoted substance.

A sentence that illustrates the scope-properties of predicative mass terms nicely is

(9) Every child drinks some milk

which has the following syntactic analysis:



and accordingly the translation:

every child $\rightsquigarrow$ $\lambda P[^{\wedge}(\mathit{every'}\ \mathit{child'})\{P\}]$

drink $\rightsquigarrow$ drink'

some milk $\rightsquigarrow$ $\lambda Q[^{\wedge}(\mathit{some'}\ \mathit{milk'})\{Q\}]$

drink some milk $\rightsquigarrow$ $\mathit{drink'}(^{\wedge}\lambda Q[^{\wedge}(\mathit{some'}\ \mathit{milk'})\{Q\}])$

every child drinks some milk $\rightsquigarrow$

$\lambda P[^{\wedge}(\mathit{every'}\ \mathit{child'})\{P\}](^{\wedge}\mathit{drink'}(^{\wedge}\lambda Q(^{\wedge}\mathit{some'}\ \mathit{milk'})\{Q\}))]$

which converts to

$(\mathit{every'}\ \mathit{child'})(^{\wedge}\mathit{drinks'}(^{\wedge}\lambda Q[(\mathit{some'}\ \mathit{milk'})Q]))$.

It is not possible to make an indirect construction in which first a sentence 'every child drinks him' is made and the term 'some milk' is quantified into the unbound variable. The syntactic rule for quantifying in is restricted to count nouns only, which is a simple syntactic constraint which, nevertheless, does not disturb the universal interpretation of all terms as generalized quantifiers.

The meaning-postulates for mass terms that restrict the set of logically possible interpretations to the reasonable interpretations of the fragment with mass terms are formulated as follows:

<1> $\Box\forall x\forall y[x\ \mathit{part\ of'}\ y \rightarrow [P(x) \leftrightarrow P(y)]]$,
where P is a predicative mass noun

which imposes the requirement that the denotation of predicative mass nouns has the property of homogeneous reference.

<2> $\forall x\Box[\alpha'(x) \rightarrow \mathit{entity'}(x)]$,
where α' translates any count noun

<3> $\forall x^{\square}[\alpha'(x) \rightarrow \mathit{quantity}'(x)]$,
where α' translates any predicative mass noun

which regulate the logical properties of count-noun denotations and predicative mass noun denotations.

<4> $\exists x^{\square}[x = a]$, where a is any propername (type e)

<5> $\exists P^{\square}[P = a]$, where a is any nominal mass noun (type <s,<e,t>>)

which regulate rigid designation of proper names and nominal mass terms, respectively.

<6> $\square\forall x[P(x) \rightarrow \boxed{c}\, P(x)]$,
where P is any predicative mass noun

which requires that a quantity of P is, whereever it exists, a quantity of P.

<7> $\forall P^{\square}[\mathit{substance}'(P) \rightarrow \forall x[{}^{\vee}P(x) \rightarrow \mathit{quantity\ of}'(x, {}^{\vee}P)]]$

which regulates the relation between substances and their quantities.[12]

FOOTNOTES

1. This point is based on an argument of Chomsky against treating proper names and quantified noun-phrases semantically on a par, as is prominently done in Montague's PTQ. Rules for backwards pronominalization can easily be restricted in the syntax to apply to proper names only. This enables me to use Chomsky's evidence, in supporting the nominal/predicative distinction in mass terms, maintaining still that quantified noun-phrases can be treated semantically on a par with proper names. See CHOMSKY 1977 and CHOMSKY 1980.
2. If the determiner in the quantified noun-phrase is definite, or if the term refers to a specific, perhaps contextually given object, the backwards pronominalization seems to be more acceptable. E.g.

   The man she loves is betrayed by the woman
   Whoever finds it may keep the money

   It remains true, however, that indefinite predicative mass terms in general do not allow for backwards pronominalization. If the preceding anaphor occurs in a subordinate clause, it seems that backwards pronominalization - even with indefinite noun-phrases - is more acceptable. E.g.

> Although he knew it was polluted, John drank some water of the well.

3. BENNETT (1974) accounts for the incorrectness of these sentences by imposing a meaning-postulate that makes such sentences necessarily false. Any interpretation that treats substances and sets of quantities on a par will have to contain such an *ad hoc* explanation of these deviant sentences, or deny that they are deviant (see CARLSON, 1977).
4. See for more arguments PUTNAM (1975).
5. It is not the intention, due to limitations of space, to discuss here the various theories of mass terms that have been proposed in the literature. However, I have in mind in particular the work of Bennett, Bunt, Carlson, Moravcsik, Parsons and Quine (see references).
6. It should be obvious that I do not necessarily intend just physical atoms here.
7. BARWISE & COOPER (1980) contains an extensive exposition of the application of generalized quantification to natural language. My discussion draws heavily upon their results.
8. In EFL determiners are basic expressions, interpreted by functions from properties to properties. Nevertheless, rules of quantification construct terms with syncategorematic determiners. From a footnote on page 219 in MONTAGUE (1974) it appears that Montague had a direct construction of terms with a syntactic category of quantifiers in mind, although in EFL such construction is not fully realized.
9. E.g. THOMASON (1979) and HAUSSER (1974).
10. Although these examples contain only logical quantifiers, which have their familiar first-order representation, they illustrate here the interpretation of generalized quantifiers. For a discussion of non-logical determiners in generalized quantifiers, that demonstrates the need for such higher-order analysis of terms, see BARWISE & COOPER (1980).
11. In TER MEULEN (1980) a complete fragment is presented.
12. A copy of the complete fragment with mass terms developed in TER MEULEN (1980) is on request available from the author.

REFERENCES

BARWISE, J. & R. COOPER, 1980, *Generalized quantifiers and natural language,* Draft, Stanford University.

BENNETT, M., 1977, 'Mass nouns and mass terms in Montague grammar', Mithuen & Williams (eds), *Proceedings of the Montague grammar conference,* SUNY Albany, New York.

BENNETT, M., 1974, *Some extensions of a Montague fragment of English,* unpublished Ph.D. dissertation, UCLA.

BUNT, H., 1978, 'A formal semantic analysis of mass terms and amount terms', Groenendijk & Stokhof (eds), *Proceedings of the second Amsterdam colloquium on Montague grammar and related topics,* Amsterdam papers in formal grammar II, Centrale Interfaculteit, University of Amsterdam.

CARLSON, G., 1977, *Reference to kinds in English,* unpublished Ph.D. dissertation, University of Massachussetts, Amherst.

CHOMSKY, N., 1977, *Essays on form and interpretation,* North-Holland Publ. Co.

CHOMSKY, N., 1980, *Rules and representations,* Columbia University Press.

HAUSSER, R., 1974, *Quantification in an extended Montague grammar,* unpublished Ph.D. dissertation, University of Texas.

MONTAGUE, R., 1973, 'Comments on Moravcsik's paper', in: Hintikka et al. (eds), *Approaches to natural language,* Reidel Publ. Co.

MONTAGUE, R., 1974, *Formal Philosophy,* Yale University Press.

MORAVCSIK, J., 1973, 'Mass terms in English', in: Hintikka et al. (eds), *Approaches to natural language,* Reidel Publ. Co.

PARSONS, T., 1970, 'Analysis of mass and amount terms', *Foundations of Language* 6.

PELLETIER, F., 1979, *Mass terms: Some philosophical problems,* Synthese Language Library, Reidel Publ. Co.

PUTNAM, H., 1975, 'The meaning of "meaning"', *Language, Mind and Reality, Philosophical Papers,* Vol. 2, Cambridge University Press.

QUINE, W.V.O., 1960, *Word and object,* MIT Press, Cambridge.

SCHWARTZ, S., 1977, *Naming, necessity and natural kinds,* Cornell University Press.

TER MEULEN, A., 1980, *Substances, quantities and individuals,* A study in the formal semantics of mass terms, unpublished Ph.D. dissertation, Stanford University.
Available from the Indiana University Linguistics Club.

THOMASON, R., 1979, 'Home is where the heart is', French (ed.), *Contemporary Perspectives in the Philosophy of Language,* University of Minnesota Press.

QUANTIFICATION, PRONOUNS, AND VP ANAPHORA

by

Barbara Partee & Emmon Bach

1. STARTING POINTS

The central concern of this paper is the nature of semantic interpretation, in particular the issue of whether an intermediate level of "logical form" (e.g. a translation into intensional logic as in MONTAGUE 1973), mediating between natural language syntax and its model-theoretical interpretation, is dispensable. Such questions cannot, of course, be examined in isolation; in this section we will lay out a number of hypotheses that we will initially assume, the dispensability of translation into IL being just one of them. In Section 2 we state the central problems to be dealt with. In Section 3 we present a fragment of English which embodies our initial hypotheses and offers solutions to certain problems about pronouns and quantifiers. In Section 4 we attempt to extend the fragment to account for "VP deletion" phenomena (though not by deletion), and then show that our original hypotheses appear not to be cotenable. We end up with an apparent need for accepting a level of "logical form" as indispensable, even if we were to give up the uniform treatment of bound-variable pronouns and "free variable" pronouns described in Section 1.1 below. There is an additional technical problem which arises in our fragment even without the extension to VP deletion, which we discuss in an appendix and which may also signal the need for a level of logical form. Our particular conclusions are necessarily relative to our initial assumptions; our only general conclusion is that the nature of free and bound variables in semantic interpretation is very much in need of further elucidation before we can hope to settle the issue of the need for or desirability of a level of logical form in explaining the relation between syntax and semantics in natural languages.

1.1. Pronouns as variables

As one initial hypothesis we adopt the proposal of COOPER (1979) that each of the third-person singular non-reflexive pronouns *he, she, it* has the set of possible meanings represented by the following translation schema:

(1) $\{\lambda P {}^{\vee}P(x_i)$: i a natural number$\} \cup \{\lambda P \bigvee x[\bigwedge y[[{}^{\vee}\pi](y) \equiv y=x] \wedge {}^{\vee}P(x)]$: π a property-denoting expressing containing only free variables and parentheses$\}$.

Under this proposal the difference between "bound variable pronouns" and other pronouns (e.g. "discourse" pronouns) is not in the meanings of the pronouns themselves, but only in whether their initially free variables (x_i in the simple case, any variables free in the π-expression in the complex case) eventually get bound within the sentence or not. For example, *him* in both (2a) and (2b) below

(2) (a) Mary loves *him*.
(b) *Every man* loves a woman who loves *him*.

has as one of its interpretations[1] $\lambda P {}^{\vee}P(x_0)$; the difference is that in (2a) an appropriate value for the free x_0 must be determinable from the linguistic or non-linguistic context for an occurrence of the sentence to receive a determinate interpretation, whereas in (2b) x_0 may be bound by interpreting *every man* as "quantified in" with respect to x_0. (The complex part of Cooper's translation schema covers cases like the *it* of (3a) and (3b) below:

(3) (a) The man who gave his paycheck to his wife was wiser than the man who gave *it* to his mistress.
(b) Every man who owns a donkey beats *it*.

But we will not be much concerned with these cases here, and will ignore them for the most part in what follows.)

Cooper's proposal contrasts with the alternative urged by JANSSEN (1980) in which the bound variable *him* in (2b) is translated as $\lambda P {}^{\vee}P(x_0)$, as above, but the "free" *him* in (2a) is interpreted as something like $\lambda P {}^{\vee}P(c_0)$, where c_0 is a context-dependent constant or what is sometimes called a context variable: semantically a constant, but dependent for its interpretation on the context. The recent literature in formal semantics contains examples of both types of treatments of pronouns[2], but relatively few arguments for choosing between them besides those in COOPER (1979) and JANSSEN (1980). We return to this issue in Section 4.6 below.

1.2. General framework

Among our other initial assumptions are the following constraints on the overall theory of syntax and semantics:

(i) We assume Montague's general theory (MONTAGUE 1970b: UG), especially with respect to compositionality, except as modified by COOPER (1975) to permit the direct assignment of sets of interpretations to ambiguous sentences without a level of a disambiguated language, including Cooper's "storage" device (which amounts to a limited relaxation of the compositionality requirement).

(ii) We assume the well-formedness constraint of PARTEE (1979) strengthened to exclude indexed forms like he_0, he_1 ... from the syntax.

(iii) The syntax is limited to a (rich) context-free grammar[3] along the lines of recent proposals by GAZDAR (1979a,1979b), GAZDAR & SAG (this volume), SAENZ and ROSS (see SAENZ (forthcoming) and ROSS (forthcoming)), and PETERS and KARTTUNEN (see KARTTUNEN (this volume)).

(iv) The semantics is a direct model-theoretic interpretation of the syntax; an intermediate level of translation into intensional logic is dispensable (COOPER (1975)).[4] This last hypothesis accords with Montague's assertions but runs counter to most earlier and much current work in semantics by linguists, where the usual assumption is that the output of semantic interpretation is a "semantic representation", an assumption which is compatible with the rest of Montague's program provided that the "semantic representation" is capable of being interpreted model-theoretically.

1.3. VP-deletion

Our final starting point is the recent work on VP-deletion by WILLIAMS (1977) and SAG (1976). Sag and Williams both provide treatments which involve the basic principle that VP-deletion depends on "semantic identity" (McCAWLEY 1967, MORGAN 1970, KEENAN 1971, VERKUYL 1972, DAHL 1972, KOSTER 1979), Sag with a deletion rule, and Williams with an interpretive principle. Both interpret semantic identity as identity of "logical form" up to within alphabetic change of bound variables. LADUSAW (1979) suggests that the appropriate sense of "semantic identity" should be model-theoretically definable; if it is not, that would be evidence for the need for an intermediate level of logical form. This issue is taken up in Section 4 below.

2. THE PROBLEMS

2.1. Reflexive and non-reflexive pronouns

Montague's PTQ did not include reflexive pronouns; an adequate treatment must give the correct distribution of reflexive pronouns and correspondingly restrict the possible interpretations of non-reflexive pronouns. All of the approximately adequate treatments so far proposed seem to violate one or more of the constraints given in Section 1.2 above (e.g. BENNETT 1976, THOMASON 1976, CHOMSKY 1973). One of our aims here is to try to provide a descriptively adequate account of pronoun distribution and interpretation within the bounds of those constraints; this is the primary goal of the fragment presented in Section 3.

2.2. Semantic identity, pronouns, and variables

The chief difficulty we find in providing a model-theoretic version of "semantic identity" appropriate for the VP-deletion phenomena arises in trying to give a suitable interpretation to variables free within the "antecedent" VP. Details and examples will be found in Section 4; here we just sketch the problem very briefly. Stated in terms of "logical form", VP-deletion may involve a VP translation containing a free variable iff either (a) that variable remains free (i.e. gets its interpretation from the context) in the sentence containing the antecedent VP and the (possibly same) sentence containing the missing VP, or (b) the variable is bound in both VP's by the same (token) variable-binder. From this condition it appears that the relevant sameness of interpretation of two VP's cannot be determined independently of the larger context in which they appear, insofar as the interpretation of any variables free within them rests on whether and where they are bound in the containing sentence(s). Janssen's proposal for non-uniform pronoun interpretations would appear to have some advantages over Cooper's uniform treatment with respect to this problem; we discuss this issue in Section 4.6.

Our discussion of these problems below will not be conclusive; we hope at least to show the need for a better understanding of the interpretation of variables. The chief difficulties encountered in both Sections 3 and 4 seem to result from the "globally syncategorematic" nature of the interpretation of variables, i.e. from the fact that their interpretation in some sense depends on what binds them, which may be arbitrarily "far away" from where they occur in the semantic structure. The difficulties are resolvable in a system which includes a level of logical form and allows principles to be stated in terms of global properties of the logical form (as is apparently allowed in the work of Chomsky, Williams, Sag, Higginbotham, and others working in the Revised Extended Standard Theory); whether they are resolvable in a more constrained theory is an open question.

3. THE FIRST FRAGMENT: QUANTIFICATION, BOUND AND FREE PRONOUNS, REFLEXIVES, AND CONTROL

The fragment presented in this section is roughly comparable to that of PTQ in its coverage, but incorporates the distinction between reflexive and non-reflexive pronouns (including a version of the "non-coreference" restrictions of LASNIK 1976 and REINHART 1977) and a set of constraints on when a quantifier phrase can be interpreted as binding a pronoun which we consider more adequate than the simple "leftmost constraint" of PTQ. In Section 3.1 we describe the format of the fragment with illustration of its novel features; in Section 3.2 we present the fragment; and in Section 3.3 we provide some examples and brief discussion. Fuller discussion and motivation of many features of the analysis are found in BACH & PARTEE (1980).

3.1. Format

Our framework is a Saenz-Ross phrase structure grammar[5] with a Cooper store (see COOPER 1975) and two new stores. The grammar is a simultaneous recursive definition of 9-tuples:

(i) an English expression
(ii) a syntactic category
(iii) syntactic features (mostly omitted here)[6]
(iv) translation into intensional logic
(v) QST ("quantifier store" or "Cooper store"): a set of pairs (possibly empty) $\langle\alpha,i\rangle$ where α is an NP meaning or WH, SELF1, or

SELF2 and i is a natural number

(vi) semantic type (omitted here)

(vii) semantic features (omitted here)

(viii) LPST ("local pronoun store"): a set of natural numbers (described below)

(ix) SPST ("super pronoun store"): a set of natural numbers (described below).

Here is an example of a rule in this format:

(R6)	S = NP VP	$LPST(1) \cap LPST(2) = \emptyset$
	$0' = 2'(^{\wedge}1')$	$QST_R(2) = \emptyset$
	$QST(0) = QST(1) \cup QST(2)$	$LPST(0) = QST_1(0)$
		$SPST(0) = SPST(1) \cup SPST(2)$

Interpretation: The numbers 0,1,2 refer to the S, the NP, and the VP respectively (numbering mentioned categories from left to right); 0' means the translation of element 0 (the S in this case). The first two lines on the left are thus tantamount to the rules S4 and T4 of PTQ. The QST statement says that the stored part of the meaning for the whole sentence is the union of the stores of its parts (the usual case). The statements on the right put conditions on the contents of the stores of the NP and the VP and define the contents of the stores of the S; the functions of LPST and SPST are sketched briefly below and discussed with examples in Section 3.3.

The function of LPST is to recursively define the domains in which a non-reflexive pronoun cannot function as a variable bound to a potential antecedent; cf. the contrast between (4a) and (4b).

(4) (a) #*Every man* loves *him*. [# indicates anomaly on intended reading.]

(b) *Every man* believes that Mary loves *him*.

Our treatment of the contrast rests on the idea that two variables within the same "local context" (roughly, co-arguments of a verb or other predicate) cannot have the same index.[7]

Our QST functions basically like Cooper's storage device (COOPER 1975); all quantifier phrases are syntactically generated in their surface positions, but their meanings may be optionally stored in QST, to be "quantified in" at a higher S or VP level. We also use QST to obligatorily store special meanings for reflexive pronouns which enable them to be treated as "relation-reducers" (GEACH 1962, POTTS 1979) applying to VP's and TVP's

(Montague's IV's and TV's), and to store a special symbol WH corresponding to a gap (an empty NP) in a relative clause, which triggers the appropriate binding by the relative pronoun, which is generated in its surface position.[8] Following Cooper, we make use of QST to capture island constraints on both quantifier scope and WH-movement (not a movement rule here). For example, the restriction QST(2) = {<WH,i>} on the relative clause rule R18 simultaneously prevents WH-extraction out of relative clauses and prevents a quantifier inside a relative clause from having scope outside that relative clause (cf. RODMAN 1976).

$QST_I(x)$ and $QST_R(x)$ are defined in terms of QST as follows:

$$QST_I(x) = \{i \mid \langle\alpha,i\rangle \in QST(x)\}$$

$$QST_R(x) = \{i \mid \langle SELF1,i\rangle \text{ or } \langle SELF2,i\rangle \in QST(x)\}.$$

$QST_I(x)$ for any expression x contains the set of indices i of pronoun meanings $\lambda P[P\{x_i\}]$ in the translation of x which are eventually going to be bound by quantifying in a stored NP meaning (whose surface syntactic position is within x). By always including $QST_I(x)$ in LPST(x), we prevent quantifier phrases from binding pronouns "higher" in the tree (in a sense made precise by the rules.) QST_R is used in stating our analog of the "clausemate" condition on reflexivization. Since reflexivization in our treatment involves storing special reflexive meanings and then bringing them out of store at TVP or VP level by rules which are necessarily optional (see R20, R21), we need to insure that there are no reflexives still in QST when a VP is combined with another constituent. This is accomplished by putting the condition $QST_R(\alpha) = \emptyset$ on the NP-VP rule, R6, and on the infinitival rule, R19.

The third store, SPST, simply keeps track of the indices of all variables still free in the interpretation of each constituent; it is similar to the "BAG" of JANSSEN (1980). Whether such a device is really necessary is an open question; it is inessential to the fragment below except if we want to implement the (controversial) "leftmost" constraint on the binding of variables by quantifier phrases (JACOBSON 1977). It becomes potentially crucial in the discussion of VP-deletion in Section 4.

3.2. The fragment

Syntactic category	*Semantic type*
S, $\bar{S}$, REL, INFS	t
CN, CNP	$\langle\langle s,e\rangle,t\rangle$
N, NP, $PP_{[to]}$	$\langle\langle s,f(CNP)\rangle,t\rangle$
V, VP, $\overline{VP}$	$\langle\langle s,f(NP)\rangle,t\rangle$
TV, TVP	$\langle\langle s,f(NP)\rangle,f(VP)\rangle$
Det	$\langle\langle s,f(CNP)\rangle,f(NP)\rangle$
VP/INFS (try)	$\langle\langle s,t\rangle,f(VP)\rangle$
VP/$\bar{S}$ (believe)	etc.
TVP/$PP_{[to]}$ (give, explain)	
TVP/NP (give, envy)	
TVP/INFS (persuade, tell)	
TVP/$\bar{S}$ (persuade, tell)	
(VP/INFS)/NP (promise)	

Lexicon sample

(The first entry contains square brackets for parts of the 9-tuple omitted here.)

$\langle \textit{John}, N, [\ \underset{\text{features}}{\text{syntactic}}\ ], \lambda P[P\{{}^\wedge j\}], \underset{QST}{\emptyset}, [\ \underset{\text{type}}{\text{semantic}}\ ] [\ \underset{\text{features}}{\text{semantic}}\ ], \underset{LPST}{\emptyset}, \underset{SPST}{\emptyset} \rangle$

$\langle \textit{he}, N, \lambda P \bigvee x[\bigwedge y[R\{x_i,y\} \equiv y = x] \wedge P\{x\}],\emptyset,\emptyset,\{i\}\rangle$

(This is one of infinitely many instances of Cooper's π-schema.)

$\langle \textit{his}, Det, \textit{his}'_i, \emptyset, \{i\},\{i\}\rangle$,

where $\textit{his}'_i = \lambda P\lambda Q \bigvee x[\bigwedge y[[P\{y\} \wedge R\{y,x_i\}] \equiv x = y] \wedge Q\{y\}]$

$\langle \textit{himself}, N, \lambda P[P\{x_i\}],\{\langle SELF1,i\rangle\},\{i\},\{i\}\rangle$

$\langle \textit{himself}, N, \lambda P[P\{x_i\}],\{\langle SELF2,i\rangle\},\{i\},\{i\}\rangle$

where $SELF1 = \lambda R\lambda x[R(x)({}^\wedge\lambda P[P\{x\}])]$, R of type $\langle\langle s,e\rangle,f(VP)\rangle$

$SELF2 = \lambda R'\lambda x[R'(x)({}^\wedge\lambda P[P\{x\}])]$, R' of type $\langle\langle s,e\rangle,f(TVP)\rangle$

$\langle \textit{e}, N, \lambda P\ P\{x_i\},\{\langle WH,i\rangle\},\{i\},\{i\}\rangle$

$\langle \textit{man}, CN, man', \emptyset, \emptyset, \emptyset\rangle$

Rules

Simple unary rules

R1. XP = X (NP=N, etc.) — LPST(0) = LPST(1)

$0' = 1'$ — SPST(0) = SPST(1)

QST(0) = QST(1)

R2. $\bar{S}$ = that S — LPST(0) = SPST(2)

$0' = 2'$ — SPST(0) = SPST(2)

QST(0) = QST(2)

R3. $\overline{VP}$ = to $VP_{[-tns]}$ — same

$0' = 2'$

QST(0) = QST(2)

R4. $PP_{[to]}$ = to NP — same

$0' = 2'$

QST(0) = QST(2)

Function-argument rules, NP *and* S

R5. NP = DET CNP — $LPST(1) \cap LPST(2) = \emptyset$

$0' = 1'(^\wedge 2')$ — $LPST(0) = QST_I(0)$

$QST(0) = QST(1) \cup QST(2)$ — $SPST(0) = SPST(1) \cup SPST(2)$

R6. S = NP VP — Same as for R5, plus

$0' = 2'(^\wedge 1')$ — $QST_R(2) = \emptyset$

$QST(0) = QST(1) \cup QST(2)$

Function-argument rules, other than NP *and* S

R7. VP = TVP NP — $LPST(1) \cap LPST(2) = \emptyset$

$0' = 1'(^\wedge 2')$ — $LPST(0) = LPST(1) \cup LPST(2)$

$QST(0) = QST(1) \cup QST(2)$ — $SPST(0) = SPST(1) \cup SPST(2)$

R8. TVP = $TVP/PP_{[to]}$ $PP_{[to]}$ — same conditions

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R9. TVP = TVP/NP NP same

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R10. TVP = TVP/$\bar{S}$ $\bar{S}$ same

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R11. VP = VP/$\bar{S}$ $\bar{S}$ same

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R12. VP/INFS = (VP/INFS)/NP NP same

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R13. VP = VP/INFS [SUBJ CONTROL] INFS [SUBJ CONTROL] same

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R14. TVP = TVP/INFS [OBJ CONTROL] INFS [OBJ CONTROL] same

$0' = 1'(^\wedge 2')$

$QST(0) = QST(1) \cup QST(2)$

R15. CNP = CNP REL same

$0' = 2'(^\wedge 1')$

$QST(0) = QST(1) \cup QST(2)$

"Transformation" RIGHTWRAP: [TVP/x x NP] ⇒ [TVP/x NP x][9].

Rules involving variables: quantification, reflexives, and control

R16. ("Store NP")

NP = NP

$0' = \lambda P[P\{x_i\}],\ i = 0,1,\ldots$

$QST(0) = QST(1) \cup \{\langle 1', i\rangle\}$

$LPST(0) = QST_I(0)$ [$=LPST(1) \cup \{i\}$]

$SPST(0) = SPST(1) \cup \{i\}$

R17. (Quantifying In)

$S = S$ — $\langle\alpha, i\rangle \in QST(1)$, $\alpha \neq$ WH, SELF1, SELF 2

$0' = \alpha({}^\wedge\lambda x_i 1')$ — $LPST(0) = QST_I(0)$ $[=LPST(1) - \{i\}]$

$QST(0) = QST(1) - \{\langle\alpha, i\rangle\}$

R18. $REL = that\ S$ — $QST(2) = \{\langle WH, i\rangle\}$

$0' = \lambda P\, \lambda x_i[P\{x_i\} \wedge 2']$ — $LPST(0) = QST_I(0)$

$QST(0) = QST(2) - \{\langle WH, i\rangle\}$ $[=\emptyset]$ — $SPST(0) = SPST(2) - \{i\}$

R19. $INFS = \overline{VP}$ — $LPST(1) \cap \{i\} = \emptyset$

$0' = 1'({}^\wedge\lambda P[P\{x_i\}])$ — $QST_R(1) = \emptyset$

a. add [FREE CONTROL] to syn features of INFS — $LPST(0) = QST_I(0)$

b. add [SUBJ CONTROL] — $SPST(0) = SPST(1) \cap \{i\}$

c. add [OBJ CONTROL]

$QST(0) =$ a) $QST(1)$

b) $QST(1) \cup \{\langle SELF1, i\rangle\}$

c) $QST(1) \cup \{\langle SELF2, i\rangle\}$

R20. (Reflexive 1: Subject control)

$VP = VP$ — $\langle SELF1, i\rangle \in QST(1)$

$0' = \lambda P[P\{{}^\wedge SELF1\ (\lambda x_i 1')\}]$ — $LPST(0) = LPST(1) - \{i\}$

$QST(0) = QST(1) - \{\langle SELF1, i\rangle\}$ — $SPST(0) = SPST(1) - \{i\}$

R21. (Reflexive 2: Object control)

$TVP = TVP$ — $\langle SELF2, i\rangle \in QST(1)$

$\lambda P_2 \lambda P_1[P_2\{{}^\wedge\lambda y SELF2(\lambda x_i 1')(y)(P_1)\}]$ — $LPST(0) = LPST(1) - \{i\}$

$QST(0) = QST(1) - \{\langle SELF2, i\rangle\}$ — $SPST(0) = SPST(1) - \{i\}$.

Following Cooper's convention, QST must be empty in any final output for that output to count as well-formed. (The constraint of JANSSEN (1980) against allowing any free variables in the final output would amount here to requiring SPST also to be empty, but we do not incorporate that constraint in this fragment.)

3.3. Discussion and examples[10]

The grammar presented above meets a number of strong constraints on syntax, semantics, and the relation between them. The syntax is essentially context-free and meets the well-formedness constraint. Montague's compositionality requirements are almost met, except that QST provides a way for NP meanings to "escape" from their surface positions. There is no way, as far as we can see, of maintaining Montague's strong form of compositionality in conjunction with direct generation of well-formed surface structures by context-free rules, given the multiple scope ambiguities of NP's. The QST device appears to be the minimal weakening of the compositionality requirement sufficient to permit the full range of NP scope phenomena.

The main innovations in this fragment are the use of LPST and its interaction with QST, and the treatment of reflexives and control. We will limit our discussion here mainly to points relevant to VP-deletion phenomena, although much of the original motivation for this fragment was to show how a "rule-by-rule" semantics would potentially improve upon the treatments of "non-coreference" phenomena and restrictions on variable-binding originally discussed in the framework of "configurational" semantics (e.g. with Reinhart's use of c-command, REINHART 1976); this comparison is the focus of BACH & PARTEE (1980).

The shortest example of the effect of LPST is the restriction it places on the interpretation of sentence (5).

(5) # *He* sees *him*.

By virtue of the LPST conditions in the lexicon and in R6 and R7, the italicized pronouns cannot be interpreted as coindexed variables.[11] As a result we achieve the blocking of bound variable readings of (6 a,b), since R16, the NP Storage rule, leaves a pronoun meaning in the place corresponding to the surface position of the stored NP.

(6) (a) # *Every man* sees *him*

(b) # *He* sees *every man*.

What the LPST conditions in effect say is that two pronoun-meanings both "locally free" in the same NP or S domain cannot have the same variable index. In terms of function-argument structure, this amounts to saying that two immediate arguments of a verb cannot be coindexed, nor can two immediate arguments of a noun.[12] While ruling out (5), these conditions permit (7):

(7) *He* loves *his* mother.

The difference between (5) and (7) is that the LPST of *his mother* in the derivation of (7) is empty (R5); *his* is not an immediate ("local") argument of *loves*. As (5) led us to predict (6 a,b) to be bad, (7) would lead us to predict (8 a,b) to be good. As expected, (8a) is fine; but (8b) is not, which brings us to the additional factor of interaction between LPST and QST.

(8) (a) *Every man* loves *his* mother.
(b) # *He* loves *every man's* mother.

In the rules above, LPST(x) always includes $QST_I(x)$, for any expression x. This means that if an NP is interpreted as having higher scope than its surface position, via QST, it may bind pronouns that are higher in the surface tree, but not ones which occur as local arguments to any function expression which contains the surface position of that NP. In other words, pronoun meanings linked to QST elements stay "local" until they are bound, acting as if they were local co-arguments of any other NP up to the point in the tree where they are "quantified in" (retrieved from store). In the derivation of (8b), if <*every man'*,i> is stored in QST, then *i* is in LPST of the object NP and hence of the VP, so the subject *he* cannot be interpreted as $\lambda P[P\{x_i\}]$. (For differences between our constraint and either the "leftmost constraint" of JACOBSON (1977) or the C-command constraint of REINHART (1977), see BACH & PARTEE (1980); our constraint is closely related to the "functional principle" of KEENAN (1974).) It may well be that all three kinds of principles--relating to function-argument structure, left-right order, and surface syntactic domination relations--have some validity, since examples ruled out by all three are most clearly bad, and those where the principles make differing predictions often seem to provoke unclear judgments of acceptability.[13])

Our treatment of reflexives is designed to capture the fact that in English, reflexives can be controlled either by direct objects (arguments of TVP's) or subjects (arguments of VP's), and the fact that they always act as bound variables, never as referential NP's.[14] A reflexive pronoun is always given a two-part interpretation: an ordinary pronoun meaning as its direct translation, and a special stored meaning--either "SELF1" or "SELF2", corresponding to subject or object control. The rules that bring these meanings out of store serve to treat the reflexive as a "relation-reducer" (see POTTS 1979). For example, in deriving sentence (9),

(9) *Every woman* told *herself* that *she* was lucky,

we might first translate *she* as $\lambda P[P\{x_3\}]$, and then translate *herself* also as $\lambda P[P\{x_3\}]$, while putting <SELF1,3> into QST; then when we bring the stored meaning out of QST at the VP level, we obtain the translation (9a), which reduces to (9b).

(9) (a) *Every woman* told *herself* that *she* was lucky.

$$\lambda P[P\{x_3\}] \qquad (\lambda P[P\{x_3\}]$$

QST: {<SELF1,3>}

told herself that she was lucky:

$$\lambda P[P\{{}^\wedge SELF1(\lambda x_3[told'({}^\wedge lucky'({}^\wedge \lambda P[P\{x_3\}]))\ ({}^\wedge \lambda P[P\{x_3\}])])\}]$$

(b) $\equiv \lambda P[P\{{}^\wedge \lambda x\ told'({}^\wedge \lambda P[P\{x\}],\ {}^\wedge \lambda P[P\{x\},\ {}^\wedge lucky'({}^\wedge \lambda P[P\{x\}])])\}].$

On this kind of derivation, the subject *every woman* does not have to be quantified in; the meaning of the VP itself guanrantees that "every woman" must be (derivatively) bound to the reflexive, and hence to the non-reflexive *she* in this case.

Our LPST restrictions account for the difference between (9) above and (10), since *him* in (10) is a local argument of *sold*.

(10) # *Every dealer* sold *himself* to *him*.

The treatment of control of infinitives in this fragment departs from previous Montague framework treatments in which infinitives are treated semantically simply as VP's. In order to capture the distribution of reflexive and non-reflexive pronouns in controlled infinitives, we are forced to posit a semantic reflex of the "missing subject".[15] Consider the standard paradigm in (11),

(11) (a) *Every unicorn* persuaded every fish to kiss *it*.
(b) # *Every unicorn* persuaded every fish to kiss *itself*.
(c) # Every unicorn persuaded *every fish* to kiss *it*.
(d) Every unicorn persuaded *every fish* to kiss *itself*.

The rule R19 translates the "missing subject" as a pronoun, $\lambda P[P\{x_i\}]$, which acts as a local argument of the embedded verb with respect to LPST, while at the same time a reflexive meaning, either <SELF1,i> or <SELF2,i> is put into QST. Obligatory control is effected via R20 and R21; our fragment in effect provides an explicit semantics for a treatment like that of HELKE (1971) which posited a "SELF" morpheme as the subject of controlled infinitives, reflecting the fact that both reflexives and controlled "deletions" are obligatorily interpreted as bound variables which act as relation-reducers.

(In the case of "free control" the missing subject is also translated as a pronoun meaning, but nothing is added to QST, so these missing subjects are predicted to behave just like embedded subject pronouns.)

In sum, our first fragment gives a treatment of quantification scope and the binding of reflexive and non-reflexive pronouns and "missing subjects" within a highly constrained version of Montague grammar. With respect to the distribution of reflexive and non-reflexive pronouns, our account gives results similar to those obtained by various indexing devices proposed within the Revised Extended Standard Theory by CHOMSKY (1980), HIGGINBOTHAM (1980 and forthcoming), and others, and our LPST can perhaps be viewed as an "indexing device" insofar as it is not otherwise needed for the syntax or the semantics. The main difference between our approach and the REST approach, besides the general differences in the two theories, is that our "locality" principles (cf. KOSTER 1978) are in terms of semantic function-argument structure rather than in terms of syntactic C-command.

In the next section, we examine the result of attempting to extend the fragment to account for VP-deletion with particular attention to cases involving quantifiers and reflexive and non-reflexive pronouns.

4. VP-DELETION

4.1. General properties

The phenomenon of VP-deletion can occur with or across sentences, but not with a non-linguistic antecedent.[16]

(12) (a) John left before Bill did.
(b) John left. Bill won't.
(c) [Scene: John leaves] *Bill won't.

As mentioned above in Section 1, it has long been observed that the identity conditions governing VP-deletion are semantic rather than syntactic; the so-called "sloppy identity" illustrated in (13) (sloppy because "himself" ≠ "myself") should really be viewed as a case of strict semantic identity.

(13) John enjoyed himself. I did too.

SAG (1976) states the condition as follows:

"With respect to a sentence S, VPD can delete any VP in S whose representation at the level of logical form is a λ-expression that is an alphabetic variant of another λ-expression present in the logical form of S or in the logical form of some other sentence S', which precedes S in discourse" (pp. 105-6)

Sag does not give explicit rules defining the level of logical form nor a full set of rules for mapping syntactic structures onto logical form, but he gives enough examples and discussion to make it clear that his "logical form" is somewhere in between English syntactic structure ("shallow") and a logical language like IL. If we try to apply Sag's formulation to IL translations of sentences as in PTQ or in our first fragment, we find that it would incorrectly predict the possibility of deleting a VP semantically equivalent to some antecedent which is not itself a syntactic VP.

Consider (14), for instance.

(14) A paper was submitted by almost every student. *But Bill didn't.

If the first sentence of (14) is interpreted with *almost every student* having wide scope, the most natural reading, then there is a *semantic* constituent roughly paraphrasable as "λx[a paper was submitted by x]", which is semantically equivalent to "submitted a paper". But that VP-type meaning cannot serve as the antecedent for the missing VP, even though it would yield a most plausible reading, and the explanation seems to be that it is not the meaning *of* any syntactic VP in that sentence. So the basic identity condition involved seems to be semantic identity between the missing VP and some antecedent syntactic VP.[17]

4.2. Initial hypotheses: an interpretive principle

There are three basic approaches to VP-deletion available:

(i) "quantifying in" VP's (BACH 1977, 1979a);

(ii) deletion of a generated full VP (e.g. SAG 1976);

(iii) interpretation of an empty or "pro" VP (e.g. WILLIAMS 1977).

In the present framework the interpretive approach is most natural. Against quantifying in, there is the fact that VP deletion occurs across as well as within sentences[18], together with the absence of independent evidence of VP scope ambiguity. Such an approach in this framework would further require relaxing the constraints on QST to allow VP meanings as well as NP meanings to be stored. The deletion approach is incompatible with the context-free syntax of the present framework.

So as our first hypothesis we will assume that the lexicon includes an empty VP, which translates as P_i, a free property variable.[19] Somewhere (it is not clear where) a condition on the interpretation of the variable must be stated, to the effect that the possible values for P_i are the intensions of any of the VP's in the same or previous sentences. (This is basically like the interpretation of free pronoun meanings, except that possible salient values available in the non-linguistic context are excluded in the VP case.)

4.3. Examples

Our first fragment makes just the right VP-meaning available in the case of reflexives, as in (15).

(15) John admires himself. Bill does too.

The translation of *admires himself* comes out as (15') via the reflexive rule R20; a simplified form with the PTQ IV-type is (15")

(15') $\lambda\mathcal{P}[\mathcal{P}\{{}^\wedge\lambda x[\text{admire}'\ ({}^\wedge\lambda P[P\{x\}])\ {}^\wedge\lambda P[P\{x\}])]\}]$

(15") $\lambda x[\text{admire}'\ ({}^\wedge\lambda P[P\{x\}])(x)]$.

But for the sentence (16), there are two non-equivalent interpretations for the verb phrase even on the assumption that *he* is John; these are given as (16a) and (16b).

(16) *John* thinks *he* is sick. Bill does too.
(a) bound *he*: $\lambda x_0[\text{think}'\ ({}^\wedge\text{sick}'(x_0))(x_0)]$
(b) free *he*: $\lambda x_1[\text{think}'\ ({}^\wedge\text{sick}'(x_0))(x_1)] \equiv \text{think}'({}^\wedge\text{sick}'(x_0))$.

In our fragment so far, only one VP meaning is assigned to a VP constituent in generating the first sentence of (16), namely (17 (=16b).

(17) $\text{think}'\ ({}^\wedge\text{sick}'(x_0))$.

Whether x_0 is free or bound isn't determined until the VP is combined with the subject NP, either directly (thus leaving x_0 free), or via storing <John', 0> in QST, leaving $\lambda P[P\{x_0\}]$ in the sentence translation, and removing *John'* from QST at the sentence level via R17 (binding x_0). Since the latter binding occurs at the S-level and not the VP-level, it does not provide a VP-constituent to serve as a possible antecedent, even though the λ-abstract formed in R17 has an appropriate meaning (i.e. 16a). (Sentence (14) above shows that we can't in general take such abstracts as antecedents for VP's.) In the next two sections we will explore two alternative extensions to the grammar that will serve to generate both (16a) and (16b) as meanings for the VP of (16). In both extensions a problem arises which requires an unpleasantly *ad hoc* restriction involving SPST. A more fundamental problem concerning meanings of variables is raised in Section 4.6.

4.4. First option: fragment plus Derived Verb Phrase rule

Both Williams and Sag employ versions of the Derived Verb Phrase rule (DVP) of PARTEE (1973) in their treatments of VP-deletion; the version we will add here will be one which maps VP's onto VP's, syntactically doing nothing to the VP, but semantically abstracting on some variable free in the VP. We will illustrate the effect of the rule before stating it. Suppose we have generated the VP in (18), with the translation and stores as in (18'):

(18) believe that she loves him

(18') $\text{believe}'\ ({}^\wedge\text{love}'({}^\wedge\lambda P[P\{x_1\}])({}^\wedge\lambda P[P\{x_0\}]))$

$QST = \emptyset \quad LPST = \emptyset \quad SPST = \{0,1\}.$

With the DVP rule added, we will have three relevant options:

(i) If we don't apply the DVP rule at all, we will have the meaning given in (18'), with both variables free.

(ii) If we apply DVP to x_0, we will obtain a meaning in which *she* is bound by λ-abstraction, i.e. (simplified) (18").

(18") $\lambda x_0[\text{believe}'\ ({}^\wedge\text{love}'({}^\wedge\lambda P[P\{x_1\}])({}^\wedge\lambda P[P\{x_0\}]))(x_0)].$

(iii) If we apply DVP to x_1, we get a similar meaning but with *him* bound.

This gives us the right set of meanings; if we applied DVP to any variable not free in the verb phrase, the result would simply be equivalent to (18'). The rule is stated below.[20]

R22. (DVP Rule)

VP = VP	$i \in SPST(1) - LPST(1)$
$0' = \lambda x_i[1'(x_i)]$	$LPST(0) = LPST(1) \cup SPST(1)$
$QST(0) = QST(1)$	$SPST(0) = SPST(1) - \{i\}$

This rule provides appropriate translations of VP's for all the basic cases of VP-deletion and gives essentially the same results as Williams' and Sag's treatments. (But see Section 4.6 for problem cases.)

4.4.1. <u>The need for SPST</u>

The condition on LPST in the rule above is unnatural; the expected condition would be LPST(0) = LPST(1): the local arguments of the verb shouldn't be affected by the rule, and SPST (the set of *all* variables free in the expression) has been completely inessential in the fragment up to this point (unless we wanted to incorporate the "leftmost" condition on quantifier-pronoun binding). (Removing SPST from the first condition on the DVP rule would simply permit some innocuous cases of vacuous λ-abstraction.) Why do we need the condition $LPST(0) = LPST(1) \cup SPST(1)$? Consider sentence (19) with the indicated choices of variables.

(19) Every man $_{VP}$[says that he believes that he loves him].
$\qquad\qquad\qquad\qquad\qquad x_0 \qquad\qquad\quad x_0 \qquad x_1$

Suppose that in the derivation, the DVP rule were applied to the VP with respect to x_0, yielding (20):

(20) $\lambda x_0[\text{say}'(x_0, {}^\wedge\text{believe}'(x_0, {}^\wedge\text{love}'(x_0, x_1)))]$.

If we had the expected condition, LPST(0) = LPST(1), nothing would block the use of x_1 for storing and quantifying in the subject, since $LPST(1) = \emptyset$. Such a derivation would yield (21):

(21) $\Lambda x_1[\text{man}'(x_1) \rightarrow \text{say}'(x_1, {}^\wedge\text{believe}'(x_1, {}^\wedge\text{love}'(x_1, x_1)))]$.

But (21) is not a possible reading of (19), since the last *him* in (19) is not reflexive. The choice of x_1 as subject of the (derived) verb phrase (20) results in overriding the distinctness of x_0 and x_1 imposed on the lowest clause by the earlier use of LPST.

Putting all of SPST into LPST in the DVP rule prevents such "overriding" of earlier LPST restrictions, but the solution is at best *ad hoc*, and it turns out that the same problem arises in another quarter, discussed in the Appendix.

4.5. Second option: pronouns as optional "distant reflexives"

One interesting feature of the DVP rule is that its translation is logically equivalent to the translation that would have been gotten by the subject-control reflexive rule (R20) if the pronoun abstracted on had been a reflexive: $\lambda P[P\{{}^\wedge\text{SELF1}(\lambda x_i 1')\}]$.[21] This suggests an alternative to the DVP, in which non-reflexive pronouns are optionally given reflexive meanings which can be carried along indefinitely in QST, to be retrieved from store at any higher VP level. The new pronoun entries will be as in (22).

$$(22)\qquad \langle \textit{he}, \text{N}, \lambda P[P\{x_i\}], \overset{\text{QST}}{\begin{Bmatrix} \emptyset \\ \{\langle \text{LAMB}, i\rangle\} \end{Bmatrix}}, \overset{\text{LPST}}{\{i\}}, \overset{\text{SPST}}{\{i\}}\rangle.$$

(Similarly for *him, she,* etc., and analogously for *his* etc.)
The stored meaning "LAMB" is semantically identical to SELF1; the difference is that SELF1 must be cleared from the store at the first possible VP (via the restriction $\text{QST}_R = \emptyset$ on R6, R19), while no such restriction is placed on LAMB. (The conditions on LPST will in fact prevent clearing LAMB out at the *first* VP level if it corresponds to e.g. a direct object *him*. The DVP rule in this version will be conditioned by the presence of <LAMB,i> in QST, rather than purely optional as in the first option.

R22'. (Alternative DVP rule)[22]

VP = VP	$\langle \text{LAMB}, i\rangle \in \text{QST}(1)$
$0' = \lambda x_i[1'(x_i)]$	$\text{LPST}(1) \cap \{i\} = \emptyset$
$\text{QST}(0) = \text{QST}(1) - \{\langle\text{LAMB}, i\rangle\}$	$\text{LPST}(0) = \text{LPST}(1) \cup \text{SPST}(1)$
	$\text{SPST}(0) = \text{SPST}(1)$

(Note that this rule requires the same SPST condition as the previous version, and for the same reason.) For English, this option is exactly equivalent to the previous. We speculate that an option like this one might be preferable for languages with a distinction between reflexive and non-reflexive possessive pronouns.

4.6. The problem of "open properties"

So far we have been treating bound and free pronouns alike, as discussed in Section 1.1. This leads to a situation in which the notion of semantic identity between VP-meanings must sometimes involve closed properties and sometimes open ones, both within and across sentences. Consider example (16) again.

(16) John thinks he is sick. Bill does too.

We assume that the second sentence of (16) is translated as (16') (using simplified types):

(16') $\mathrm{Bill}'(P_0)$.

The two available values for P_0 are the intensions of the two expressions (16a) and (16b), repeated below.

(16a) $\lambda x_0[\mathrm{think}'(^\wedge\mathrm{sick}'(x_0))(x_0)]$ (closed)

(16b) $\lambda x_1[\mathrm{think}'(^\wedge\mathrm{sick}'(x_0))(x_1)]$ (open).

In case (16b), we depend on the assumption that the context includes a choice of variable assignment g (call it g_c), such that the same assignment g_c applies in both sentences. In the case of discourse (16) on the interpretation we have been considering, $g_c(x_0)$ is John. We can therefore look at the semantic identity in question either as involving open properties or as involving closed properties obtained by evaluation with respect to g_c.

The only cases which *crucially* involve identity of *open* properties are those in which a pronoun within the VP is bound from outside, as in one reading of (23).

(23) *Every man* believes that Sally loves *him* and Mary doesn't.

The verb phrase *loves him* has an interpretation (23').

(23') $\mathrm{love}'(x_0)$ (equivalent to $\lambda x_1[\mathrm{love}'(x_0)(x_1)]$).

We want to let the semantic value of (23') be the value of the variable P_0 of the missing VP, and we need to allow both "occurrences" of the free x_0 to be bound by *every man,* so we must *not* require g_c to supply values for all variables free in the VP meaning, when determining values for P_0.

So in order to account properly for VP-deletion in cases like (23), it appears that we must let the semantic value of the VP involve crucially what COOPER (1979) calls the *meaning* of the variable x_0 (cf. MONTAGUE, UG), which is a function from contexts (assignments) to individual concepts. This contradicts our starting assumption (Section 4.2) that it could be simply the *intension* of the antecedent VP which provides the value for the missing VP (intensions are functions only from worlds and times; we need a function from worlds, times, and assignments).[23]

At this point we seem to have established that the relevant semantic value of a VP containing a free variable should be an open property, which we can now characterize more precisely as a function from worlds, times, and assignments to VP-extensions. Supplying such a value for P_0 correctly allows the "copied" variable to be bound by a quantifier in whose scope it falls, as in (23), or to remain free and receive its value from the context assignments g_c, as in (16b).

But now a fundamental difficulty arises. What can prevent the incorrect prediction that a variable free within the VP could be bound in its antecedent occurrence and remain free in the deletion site, or vice versa, or be bound in both but by different variable-binders? These cases are illustrated in (24), (25), and (26) respectively.

(24) *No man* believes that Mary $\underbrace{\text{loves } \textit{him}}_{{}^\wedge love'(x_0)\ \text{(bound)}}$. #But she does.

↑ Bind with x_0

does ↓ P_0 ↓ ${}^\wedge love'(x_0)$ (free)

(25) Mary $\underbrace{\text{loves him}}_{{}^\wedge love'(x_0)\ \text{(free)}}$. #*Every boy* assumes that Sally does.

Every boy ↑ bind with x_0

does ↓ P_0 ↓ ${}^\wedge love'(x_0)$ (bound)

(26) #*Bill* believes that Sally will $\underbrace{\text{marry him}}_{{}^\wedge marry'(x_0)\ \text{(bound)}}$, but *everyone* knows that

Bill ↑ Bind with x_0

everyone ↑ bind with x_0

she won't_______.

↓ P_0 ↓ ${}^\wedge marry'(x_0)$ (bound)

Our rules permit all of these (bad) cases. In (24), the x_0 in the antecedent is bound by *no man,* while the x_0 of the missing VP receives a value from the context assignment g_c, giving the second sentence an interpretation such as "But she does love Fred", clearly impossible. In (25) the x_0 of the antecedent gets a value from g_c, say Fred, but the x_0 of the missing VP is bound by *every boy,* rather than being forced as it should be to also refer to Fred. In (26), both occurrences of x_0 are bound, but the first by *Bill* and the second by *everyone,* again yielding an impossible interpretation.

The problem is a fundamental one, in our view, because it leads to the conclusion that there is no semantic value that can be assigned to VP's such that VP-deletion can be characterized in terms of semantic identity. It is not difficult to state a restriction on the notion of semantic identity for open VP's that will exclude all of the above cases, but the restriction crucially involves global properties of the IL representation. The restriction is stated in (27) in terms of an example; generalizing it is cumbersome but straightforward.

(27) Two occurrences of "love'(x_0)" are semantically identical iff either (i) both occurrences of x_0 are free, or (ii) both occurrences are bound by the same (token) variable-binder.[24]

The restriction is global in that there is no limit to how far away from the VP a variable-binder might occur which binds a variable free within the VP. It crucially involves the syntax of IL because of the essential reference to tokens of variable-binders in (ii). In effect, it is as if the meaning of a variable free *within* a VP is determined by whether and where it is bound *outside* that VP; if this is so, then there is a strong sense in which meanings cannot after all be considered compositional.[25]

Part of the problem could be solved if we adopted Janssen's treatment of pronouns rather than Cooper's (see Section 1.1), translating pronouns as either $\lambda P[P\{c_i\}]$ or $\lambda P[P\{x_i\}]$, with c_i acting as a context constant that cannot be bound and with an output filter that rules as ill-formed any expression with a free x_i in it. The cases of examples (24) and (25) would no longer arise, and VP-deletion *across* sentences could be simply required always to involve *closed* VP-meanings, i.e. simple identity of intensions.

But even with this modification, a global restriction like the second part of (27) would still be required for VP-deletion within sentences, since there are cases crucially involving open properties within sentences, e.g. (23), and thus the problem cases like (26) could still arise. If, as suggested in footnote 18, it should turn out that it would be best to treat within-sentence VP-deletion separately from across-sentence VP-deletion, we should reconsider the possibility of using a quantifying-in rule for the former (which would increase the similarity of types of VP-deletion to types of pronominal anaphora). It looks as though splitting the VP-deletion phenomena in this way and incorporating Janssen's treatment of pronouns might obviate the need for a global IL-level restriction like (27). But we cannot make any firm conclusions without a treatment of "antecedent-contained" VP-deletion (see footnote 17), so we leave that possibility for future exploration.

Our conclusions from this section can be summarized as follows:

(i) if we follow Cooper in assigning uniform interpretations to free and bound pronouns, we cannot handle VP-deletion without a level of translation into IL;

(ii) if we instead adopt Janssen's non-uniform treatment of free and bound pronouns, and if we treat VP-deletion differently within and across sentences, then it *might* be possible to continue to regard the level of IL as dispensable;

(iii) the widely accepted view that VP-deletion involves semantic identity offers an interesting and so far unmet challenge, since it cannot be made precise without resolving the problematic issue of the semantic value of free variables.

5. SUMMARY

In this paper we have tried to provide an account within a constrained version of Montague grammar of various phenomena involving quantification and anaphora which challenge the assumption of the dispensability of any level of "logical form" or translation into IL mediating between surface syntax and model-theoretic interpretation. What REST theories do with global indexing mechanisms we have attempted to do with the addition of one limited device, LPST, to the recursive rules which simultaneously construct syntactic structures and their semantic interpretations. The interaction between LPST and our version of Cooper's storage device, QST, permits us to provide a fairly unified account of the distribution of reflexive and non-reflexive pronouns, the conditions on permissible binding of pronouns by quantifiers, the parallels among reflexives, obligatory control, and the DVP rule, and the parallels between WH-movement and quantification. QST can be viewed as a limited weakening of the compositionality constraint which makes possible a strengthened constraint on the syntax, namely that it be context free. (But see ENGDAHL (1980) for a serious challenge to this claim.) LPST can be viewed as a device which is neither purely syntactic nor purely semantic, perhaps as a concession to the need for some properties of a level of "logical form" intermediate between syntax and semantics; but if LPST suffices, it is a much more limited addition than a whole level of representation.

It is not clear yet whether LPST does suffice, however; the condition on both versions of the Derived Verb Phrase rule requiring all variables free in the resulting verb phrase (i.e. all of SPST) to be added to LPST is an *ad hoc* and in some sense global device that is needed to prevent earlier (local) effects of LPST conditions to be destroyed by later lambda-abstraction on a deeply embedded variable. The Appendix shows that reflexivization creates a similar problem in the first fragment even without the DVP rule. Since SPST is itself a rather limited device (see Janssen's discussion of his analogous BAG device), this problem does not by itself show that a level of "logical form" is indispensable.

The main problem facing our attempt to constrain the theory as described above arose in the extension of our fragment to handle VP-deletion. The DVP rule (either version) provides a natural extension of the fragment which gives the right results for the most part, but we have not yet been able to account for cases in which there is a variable free within the antecedent VP without appeal to global restrictions on the level of the syntax of IL. We feel that a deeper understanding of the semantics of free and bound variables is a crucial prerequisite to further investigation of the important question of whether a linguistic theory needs a level intermediate between syntax and model-theoretic semantics.

6. APPENDIX: FURTHER CASES OF APPARENT GLOBALITY OF COINDEXING RESTRICTIONS

In Section 4.4.1 we discussed the need to prevent combining a subject interpreted as $\lambda P[P\{x_i\}]$ with a derived verb phrase interpreted as $\lambda x_j \phi$ if if x_i is free in ϕ. The reason was that the net effect would be to identify x_i with x_j, overriding any earlier restriction that x_i and x_j be distinct (see examples (19), (20), (21)). The needed restriction was achieved in both versions of the DVP rule, R22 and R22', by adding all of SPST to the LPST of the derived verb phrase. In this appendix, we present three additional places in the grammar where a similar problem arises: (i) double applications of DVP; (ii) application of DVP to a subject-controlled reflexive VP; and (iii) double application of the subject-control reflexive rule, R20. In the first two cases, the solution of adding all of SPST to LPST works all right, but in the third it does not, suggesting that a more powerful global device than SPST may be needed.

6.1. Double applications of DVP

For the same reason as above, we must prevent the DVP (either version) from applying twice in a row to the same VP. (To see that this is the same problem, note that the result would have the effect of identifying the first variable abstracted on with the second variable abstracted on, again overriding any earlier distinctness requirements.) The restriction already suggested for the original problem, adding all of SPST to LPST in the DVP rule, takes care of this problem as well. Note that in the second option, R22' of Section 4.5, both <LAMB,i> and <LAMB,j> can be in QST together; the restriction simply prevents bringing them out of store on the same VP.

6.2. DVP and subject-controlled reflexives

Similarly, we must prevent applying DVP to a subject-controlled reflexive, which on our treatment is always interpreted as a λ-abstract, as in (28).

(28) convince himself that he loves him

x_0 x_0 x_1

<SELF1,0>

$\lambda x_0[\text{convince}'(x_0, x_0, {}^\wedge\text{love}'(x_0, x_1))]$

In this case as well, applying the DVP rule with respect to a variable free in the VP, here x_1, could override the earlier restriction (via LPST conditions on the embedded clause) that x_0 and x_1 be distinct.

And for the same reason, a reflexive VP such as the above must not be combined with a subject $\lambda P[P\{x_i\}]$ if x_i is free in the VP.

We can solve this pair of problems by adding all of SPST to LPST in the subject-control reflexive rule R20 as we did in R22. This is a modification of the original fragment which forces a retraction of the claim that SPST is inessential to that fragment.

6.3. Double application of the subject-control reflexive rule

We need to be able to apply R20 twice in a row to generate (29):

(29) Mary talks to herself about herself.

$\lambda P[P\{x_2\}]$ $\lambda P[P\{x_3\}]$

<SELF1,2> <SELF1,3>

(Why aren't the two reflexives necessarily given the same translation in the first place, if they are both subject-controlled, i.e. SELF1, and in the same clause? Because in this bottom-to-top recursive syntax, each one is generated separately and is "blind" to what it will be combined with later. We assume here that both of these pronouns are arguments of the verb *talk*, which would be categorized as $(VP/PP)/PP_{[+to][+about]}$; but the problem raised here would still arise if either or both PP's were regarded as VP or TVP modifiers.)

We need to give reflexives a basic $\lambda P[P\{x_i\}]$ translation in addition to their stored reflexive meaning in order to permit coindexing of reflexives with lower plain pronouns, as in (28) above.[26] But these factors together create the possibility here too of overriding earlier distinctness restrictions, as in (30) below.

(30) #Every man talked to himself by himself about a book that he gave to him.

$\lambda P[P\{x_0\}]$ $\lambda P[P\{x_1\}]$ $\lambda P[P\{x_0\}]$ $\lambda P[P\{x_1\}]$

<SELF1,0> <SELF1,1>

This problem, unlike the previous ones, cannot be solved by requiring the reflexive rule R20 to apply only to variables not in LPST (and adding all of SPST to LPST, as we already did in 6.2 above), since the reflexivized variable itself is *always* in LPST at the time R20 applies. We see no solution to this problem using the limited mechanisms of our framework.

The problem may of course arise from some fault in our particular analyses, or from some aspect of our framework we have not considered in detail. What worries us is that it looks like a natural solution to the whole family of problems raised in 4.4.1 and this appendix is to reinstate the translation into IL as a level of "logical form" and employ a global book-keeping device on that level akin to Chomsky's "anaphoric index": a set of integers attached to each indexed pronoun meaning indicating explicitly which indices it must be distinct from. To implement such a solution appears to require having access to the entire "logical form" of the sentence. (Note that there is nothing *semantically* ill-formed about *love'*(x_i,x_i).) The apparent need for a device of such power is a serious challenge both to compositionality (even as weakened via the use of storage devices) and to the dispensability of a level of "logical form".

FOOTNOTES

We are grateful to our colleagues, students, and visitors to our department this past year for providing a most stimulating atmosphere for discussing problems in syntax and semantics from the varying perspectives of Montague grammar and current versions of the Revised Extended Standard Theory. Special thanks go to Rick Saenz and Ken Ross for their important role in developing the framework discussed here, and to Robin Cooper, Elisabet Engdahl, Irene Heim, Mats Rooth, Arnim von Stechow, Lars Hellan, and James Higginbotham for valuable discussion. Remaining inadequacies are our own.

We gratefully acknowledge support for this work from the National Science Foundation, grant #N.S.F. B.N.S. 78-19309 and indirectly from the Sloan Foundation, grant #78-4-14.

1. As Cooper does, we will employ expressions of Montague's intensional logic to represent the model-theoretic entities they denote; except where explicitly noted, we are using the logic as an inessential convenience and not as a genuine "linguistic level".
2. For example, a position like Cooper's is adopted in MONTAGUE (1970b) and in GROENENDIJK & STOKHOF (1976); HAUSSER (1979) also argues for a uniform treatment of pronouns but with all of them involving context variables; BARTSCH (1979) invokes an intermediate representational level in which pronouns are represented uniformly, but at a subsequent stage some are translated as variables and others as constants. BENNETT (1978) makes a clear distinction between bound variable pronouns and demonstrative pronouns (constants) but does not discuss non-demonstrative anaphoric pronouns.
3. More precisely, to a class of grammars whose generative capacity does not exceed that of context-free grammars.

4. See footnote 1. Also see the paper by VAN BENTHEM (this volume) for some discussion of this issue. See also LADUSAW (1979).
5. The main features of our framework were originally proposed by Rick Saenz and Ken Ross (see SAENZ (forthcoming) and ROSS (forthcoming)), neither of whom is necessarily in agreement with the details of our use of it or the analysis of English presented. Similar proposals are mentioned in Section 1.2 (iii), above.
6. We assume a treatment of syntactic features similar to that of GAZDAR (1979a). The only features mentioned explicitly in the fragment below are a few non-standard ones that are essential for making sense of the rules.
7. Montague suggested approximately this principle in footnote 12 of MONTAGUE (1970a: EFL); the recursive definition of "exposed to reflexivization" in THOMASON (1976) has a similar function. The problem of how best to describe such constraints has received much recent attention in the EST framework (REINHART 1977, CHOMSKY 1980, HIGGINBOTHAM 1980).
8. We treat only the simplest cases of relative clauses with relative pronoun "that"; serious difficulties for direct surface structure generation of WH-constructions are posed by the work of ENGDAHL (1980).
9. The inclusion of the "rightwrap" transformation (see BACH 1979b) makes the grammar not a strictly phrase-structure grammar; see GAZDAR & SAG (this volume) for a way of including TVP (which frequently appears to be a discontinuous constituent in surface structure) in a phrase-structure grammar via a metarule which has much the same effect as our rightwrap operation.
10. Our discussion here is limited mainly to issues relevant to the discussion of VP-deletion in Section 4; further discussion of this fragment is contained in BACH & PARTEE (1980).
11. Since there is nothing within the semantics proper that prevents assigning the same value to distinct variables, our conditions do not require that the two pronouns refer to distinct individuals when sentence (5) is used as a complete sentence. We believe that our restrictions give semantically appropriate results with respect to pronouns as bound variables, as in (6 a,b), and that a full account of the effect of the restrictions on nonbound-variable pronouns requires an explicit pragmatics, which we have not attempted to provide. See POSTAL (1971) and HIGGINBOTHAM (1980, footnote 1) for discussion of the problems of

specifying the linguistically appropriate notion of "non-coreference".

12. A major shortcoming of the present work is the absence of a full treatment of complex NP's involving various sorts of prepositional phrase adjuncts; see HIGGINBOTHAM (forthcoming) and FIENGO & HIGGINBOTHAM (ms.) for rich sources of relevant examples, problems, and hypotheses.

13. On the positive side, our fragment predicts correctly that (i) is all right, while the leftmost constraint would rule it out.

    (i) In *its* early years, *every organization* suffers some setbacks.

    And it also predicts the acceptability of (ii), which the c-command condition of REINHART (1977) would disallow.

    (ii) Every student claimed that *one of his professors* was a genius in order to influence *her*.

    As apparent counterexamples, our fragment predicts that (iii) is bad and (iv) good; note that c-command would rule both bad, while the leftmost contraint would allow (iii) and disallow (iv).

    (iii) *Every man's* mother loves *him*.

    (iv) *His* mother loves *every man*.

    In BACH & PARTEE (1980), we offer an explanation of the goodness of (iii) consistent with our fragment by arguing that the *him* is not a bound variable but a "Geach's donkey"-pronoun, to be treated by Cooper's complex pronoun translation schema as in (3) above. See also HIGGINBOTHAM (1980, forthcoming) for discussion of the problem status of these cases.

14. Some speakers of English do apparently allow reflexives to act as referential pronouns; for them, but not for us, sentence (i) is ambiguous with respect to whether Bill voted for himself or for John.

    (i) John voted for himself, and Bill did, too.

    It appears that we would have to posit a substantial difference in grammars for these two dialects, but we have not worked out an explicit treatment for the other dialect so far.

15. THOMASON (1976) captures these restrictions via a recursive syntactic characterization of the property "exposed to reflexivization", and BENNETT (1976) via an abstract syntactic marker "*" for subscripted pronouns in reflexivizable positions. This seems to be another case where some device of considerable power is needed, and by restricting our syntax to disallow either of the above solutions, we are forced to represent the "missing subject" explicitly in the semantics.

16. We use the term "VP-deletion" for historical continuity; the issues to

be discussed seem to be independent of whether the phenomenon is treated as deletion or as interpretation of an empty or "Pro" VP.

17. We are skirting the difficult issue of "antecedent-contained" VP-deletion, illustrated by (i) and (ii) below, not because we think it is irrelevant, but because we haven't found an adequate treatment of it and the problems raised by pronouns in "simple" VP-deletion cases are difficult enough.
    (i) Sam $_{VP}$[put everything that Mary told him to $_{VP}$[__] into the suitcase].
    (ii) Sam $_{VP}$[put more shirts than he wanted to $_{VP}$[__] into the suitcase].
    However, it must be noted that these examples seem to be incompatible with the basic identity condition stated above, so a good account of the antecedent-contained cases might radically change our view of the "ordinary" cases.
    Sag offers some additional examples in which the antecedent appears not to be a syntactic VP constituent; it remains to be seen whether it is possible to define a level of logical form which will predict the impossibility of examples like (14) without placing any syntactic constraints on the antecedent.
18. But it might turn out that VP-deletion should be split into two cases, a bound-variable type occurring only within sentences, and a non-bound-variable type occurring both within and across sentences, with a quantifying-in approach appropriate for the former. Our remarks here would then apply only to the latter.
19. The semantic type of VP's in our fragment is <<s,f(NP)>,t>, since we treat VP's as taking their subjects as arguments. Technically, the translation of an empty VP should be $\mathbb{P}_i$, where $\mathbb{P}$ is of type <s,f(VP)>. We need an intensional rather than extensional variable to account for cases like (i).
    (i) John wants to catch a fish. Bill wants to, too.
    In what follows, we will not be careful about the intension/extension distinction, and will use the simpler PTQ type for ease of exposition except when showing specific derivations from our fragment.
20. The translation is in simplified PTQ types. The official translation for our type system would be:
    $$0' = \lambda P[P\{{}^\wedge\lambda x_i[1'({}^\wedge\lambda P[P\{x_i\}])]\}].$$
21. This remark applies to the official translation given in footnote 20.
22. In the type theory of our fragment, we could either use the same translation as that given in footnote 20, or to make the semantic identification of LAMB with SELF1 explicit, use the translation of the reflexive rule R20.

23. It has been pointed out to us by Robin Cooper (personal communication) and is discussed in BIGELOW (1978), that there is a major difference between Montague's treatment of the meanings of variables and that of CRESSWELL (1973). Montague treats the meaning of variables on a par with the meanings of indexicals like *I* and *here* by making assignments of values to variables a part of the context; Cresswell treats variables as essentially syncategorematic and lets them take themselves as meanings. We have not yet tried to work out the consequences of Cresswell's approach for our family of problems.
24. Sag and Williams both build this restriction into their definitions of alphabetic variance.
25. We state these conclusions tentatively in hopes that some logican will be able to show us a way of construing meanings of variables, or of rethinking VP-deletion, for which these problems can be solved within a compositional and purely model-theoretic semantics.
26. Ewan Klein pointed out to BHP at the conference that on Gazdar's treatment of reflexives via a "distinguished variable", where coindexing of a reflexive and a non-reflexive is not possible, the same effect can be achieved indirectly, as in the following sketch of a derivation of (i):
   (i) Every man convinced himself that he loved Mary.
      (a) he loved Mary: $\text{love}'(x_0,m)$
      (b) convince himself that he loved Mary: $\lambda r(\text{convince}'(r,r\,{}^\wedge\text{love}'(x_0,m)))$
      (c) treat *every man* as quantified in with respect to x_0, which in our system means translating it as $\lambda P[P\{x_0\}]$ and storing $\langle\text{every man}',0\rangle$.

   We avoided the distinguished variable treatment of reflexives in our fragment because of problems which Ewan Klein has since shown us may not be insurmountable; see ENGDAHL (1980) for discussion. If distinguished variables are used for reflexives, the reflexive problems discussed in this Appendix do not arise, but there may be other problems about multiple reflexives. As far as we are concerned, the issue is still open.

REFERENCES

BACH, E., 1977, 'An extension of classical transformational grammar', in: R. Saenz (ed.), *Problems in Linguistic Metatheory*, Proceedings of the 1976 conference at Michigan State University, East Lansing, Michigan.

BACH, E., 1979a, 'Montague grammar and classical transformational grammar', in: S. Davis and M. Mithun (eds), *Proceedings of the Conference on Montague Grammar at SUNY Albany*, University of Texas Press, Austin, 1979.

BACH, E., 1979b, 'Control in Montague grammar', *Linguistic Inquiry*, 10, 515-531.

BACH, E. & B.H. PARTEE, 1980, 'Anaphora and semantic structures', *CLS*, 16.

BARTSCH, R., 1979, 'The syntax and semantics of subordinate clause constructions and pronominal coreference', in: HENY & SCHNELLE 1979, pp.23-59.

BENNETT, M., 1976, 'A variation and extension of a Montague fragment of English', in: PARTEE 1976, 119-163.

BENNETT, M., 1978, 'Demonstratives and indexicals in Montague grammar', *Synthese*, 39, 1-80.

BIGELOW, J., 1978, 'Semantics of thinking, speaking, and translation', in: F. Guenthner and M. Guenthner-Reutter (eds), *Meaning and Translation*, New York University Press, New York, 1978, pp.109-135.

CHOMSKY, N., 1973, 'Conditions on transformations', in: S.R. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York, 1973, pp.232-286.

CHOMSKY, N., 1980, 'On Binding', *Linguistic Inquiry*, 11, 1-46.

COOPER, R., 1975, *Montague's Semantic Theory and Transformational Syntax*, unpublished doctoral dissertation, University of Massachusetts, Amherst.

COOPER, R., 1979, 'The interpretation of pronouns', in: HENY & SCHNELLE 1979, pp.61-92.

CRESSWELL, M.J., 1973, *Logics and Languages*, Methuen, London.

DAHL, O., 1972, 'On so-called "sloppy identity"', *Gothenburg Papers in Theoretical Linguistics*, 11, 3-45.

ENGDAHL, E., 1980, *Questions in Swedish*, unpublished doctoral dissertation, University of Massachusetts, Amherst.

FIENGO, R. & J. HIGGINBOTHAM, (ms.), 'Opacity in NP', CUNY Graduate Center and Columbia University.

GEACH, P.T., 1962, *Reference and Generality*, University of Cornell Press, Ithaca (amended ed., 1968).

GAZDAR, G., 1979a, 'Constituent Structures', unpublished manuscript, School of Social Sciences, University of Sussex (Brighton, BN1 9QN, England).

GAZDAR, G., 1979b, 'English as a context-free language', unpublished manuscript, School of Social Sciences, University of Sussex (Brighton, BN1 9QN, England).

GAZDAR, G. & I. SAG (this volume), 'Phantom categories in phrase structure grammar'.

GROENENDIJK, J. & M. STOKHOF, 1976, 'Some notes on personal pronouns, reflexives, and sloppy identity in a Montague grammar', in: K. Braunmüller and W. Kürschner (eds), *Akten des 10. linguistischen Kolloquiums, Tübingen 1975*, Max Niemeyer Verlag, Tübingen, 1976, pp.301-315.

GROENENDIJK, J. & M. STOKHOF (this volume), 'Semantics of *wh*-complements'.

HAUSSER, R., 1979, 'How do pronouns denote?', in HENY & SCHNELLE 1979, pp.93-139.

HELKE, M., 1971, *The grammar of English reflexives*, unpublished doctoral dissertation, M.I.T.

HENY, F. & H.S. SCHNELLE (eds), 1979, *Syntax and Semantics, Vol. 10: Selections from the Third Groningen Round Table*, Academic Press, New York, 1979.

HIGGINBOTHAM, J., 1980, 'Anaphora and GB: Some preliminary remarks', *NELS*, 10.

HIGGINBOTHAM, J. (ms.), 'Reciprocal interpretation', Columbia University.

HIGGINBOTHAM, J. (forthcoming), 'Pronouns and bound variables', to appear in *Linguistic Inquiry*.

JACOBSON, P.I., 1977, *The syntax of crossing coreference sentences,* unpublished doctoral dissertation, University of California, Berkeley.

JANSSEN, T., 1980, 'On problems concerning the quantification rules in Montague grammar', in: C. Rohrer (ed.), *Time, Tense and Quantifiers: Proceedings of the Stuttgart Conference on the Logic of Tense and Quantification,* Max Niemeyer Verlag, Tübingen, 1980, pp.113-134.

KARTUNNEN, L. (this volume), 'CF-rules with CS-filtering for incomplete phrases'.

KEENAN, E., 1971, 'Names, quantifiers, and a solution to the sloppy identity problem', *Papers in Linguistics* 4.2.

KEENAN, E., 1974, 'The functional principle: generalizing the notion of subject-of', *CLS,* 10, 298-309.

KOSTER, J., 1978, *Locality principles in syntax,* Foris Publications, Dordrecht.

KOSTER, J., 1979, 'Anaphora: an introduction without footnotes', Report No. DA 01-79 Filosofisch Insituut, Nijmegen.

LADUSAW, W., 1979, *Polarity sensitivity as inherent scope relations,* unpublished doctoral dissertation, University of Texas, Austin.

LASNIK, H., 1976, 'Remarks on coreference', *Linguistic Analysis,* 2, 1-32.

McCAWLEY, J.D., 1967, 'Meaning and the description of languages', *Kotoba no uchu,* 2, No. 9 - No. 11.

MONTAGUE, R., 1970a, 'English as a formal language', reprinted in MONTAGUE 1974, 188-221.

MONTAGUE, R., 1970b, 'Universal grammar', reprinted in MONTAGUE 1974, 222-246.

MONTAGUE, R., 1973, 'The proper treatment of quantification in ordinary English', reprinted in MONTAGUE 1974, pp.247-270.

MONTAGUE, R., 1974, *Formal philosophy: Selected papers of Richard Montague,* edited and with an introduction by R. Thomason, Yale University Press, New Haven.

MORGAN, J. 1970, 'On the criterion of identity for noun phrase deletion', in: *CLS, 6*, 380-389.

PARTEE, B., 1973, 'Some transformational extensions of Montague grammar', *Journal of Philosophical Logic, 2*, 509-539, reprinted in PARTEE 1976.

PARTEE, B. (ed.), 1976, *Montague Grammar,* Academic Press, New York.

PARTEE, B.H., 1979, 'Montague grammar and the well-formedness constraint', in: HENY & SCHNELLE 1979, pp.275-313.

POTTS, T., 1979, 'A general theory of the meaning of anaphoric pronouns', in: HENY & SCHNELLE 1979, pp.141-198.

POSTAL, P., 1971, *Cross-over phenomena,* Transatlantic Series in Linguistics, Holt, Rinehart and Winston, New York.

REINHART, T., 1976, *The syntactic domain of anaphora,* unpublished doctoral dissertation, M.I.T.

REINHART, T., 1977, 'Restrictions on the anaphoric interpretation of indefinite nouns', in: A. Zampolli (ed.), *Proceedings of the 4th International Summer School,* Pisa, 1977.

REINHART, T., 1979, 'Syntactic domains for semantic rules', in: F. Guenthner & S.J. Schmidt (eds), *Formal Semantics and Pragmatics for Natural Languages,* Reidel, Dordrecht, 1979.

RODMAN, R., 1976, 'Scope phenomena, "Movement Transformations" and relative clauses', in: PARTEE 1976, pp.165-176.

ROSS, J.R., 1968, 'A note on the notion "command"', in: *Report No. NSF - 20,* Harvard Computatoral Laboratory, Cambridge, Massachusetts.

ROSS, J.R., 1969, 'Guess who?', in: R. Binnick et al. (eds), *Papers from the fifth Regional Meeting of the Chicago Linguistic Society,* Dept. of Linguistics, University of Chicago, Chicago, 1969, pp.252-286.

SAENZ, R. (forthcoming), *English phrase structure,* unpublished doctoral dissertation, University of Massachusetts, Amherst.

SAG, I., 1976, *Deletion and logical form,* unpublished doctoral dissertation, M.I.T.

THOMASON, R., 1976, 'Some extensions of Montague grammar', in: PARTEE 1976, pp.77-117.

VAN BENTHEM, J. (this volume), 'Why is semantics what?'

VERKUYL, H.J., 1972, *On the compositional nature of the aspects*, Reidel, Dordrecht.

WILLIAMS, E., 1977, 'Discourse and logical form', *Linguistic Inquiry*, 8, 101-139.

DISTRIBUTIVE, COLLECTIVE AND CUMULATIVE QUANTIFICATION

by

Remko J.H. Scha

1. INTRODUCTION

Theories which relate English sentences to logical formulas representing their truth conditions usually assume that for every noun phrase in the surface structure there is a quantifier in the logical formula, and that it depends on the determiner of the noun phrase what quantifier that is. The present paper points out some hitherto neglected phenomena which give rise to significant departures from this approach.

The first of these issues concerns the meaning of plural noun phrases with "the". Unless they are combined with "collective verbs", like "gather" or "disperse", such noun phrases are commonly taken to indicate universal quantification over the extension of the noun.[1] It is easy to see that this is not adequate, if we consider the combination of two plural noun phrases with "the" with transitive distributive verbs. For instance, (1a), (2a) and (3a) have readings that may be paraphrased as (1b), (2b) and (3b), respectively. (See Fig. 1 and Fig. 2.)

(1a) The squares contain the circles.

(1b) Every circle is contained in some square.

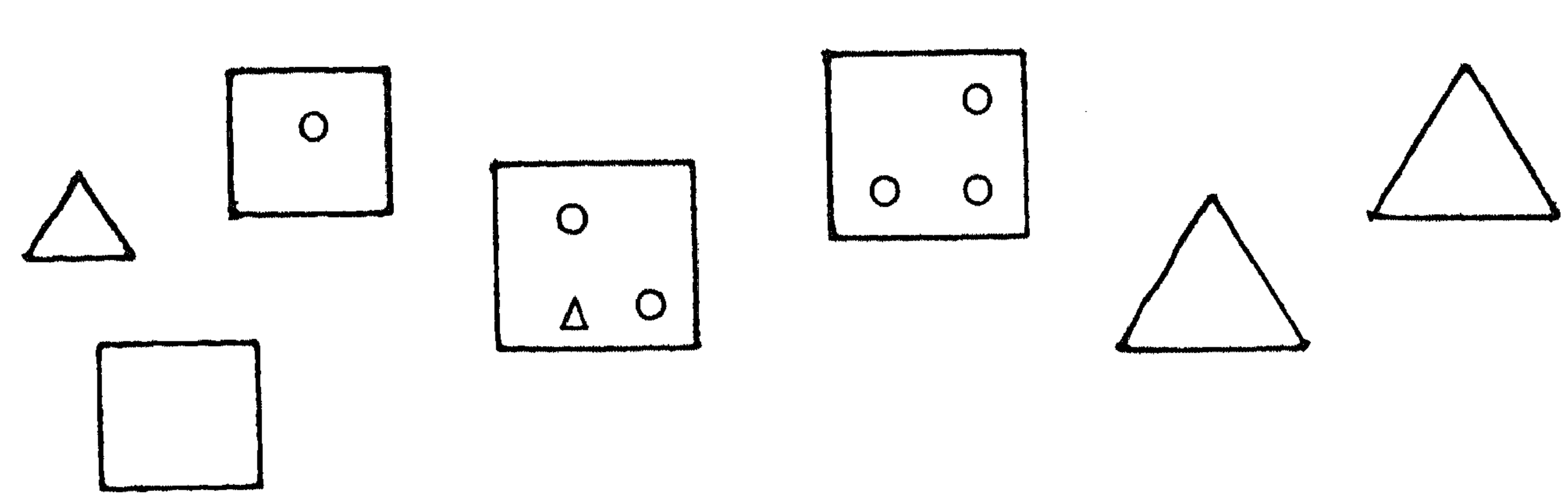


Figure 1.

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

(2b) Every side of rectangle 1 runs parallel to some side of rectangle 2 and every side of rectangle 2 runs parallel to some side of rectangle 1.

(3a) The sides of rectangle 1 cross the sides of rectangle 2.

(3b) Some sides of rectangle 1 cross some sides of rectangle 2.

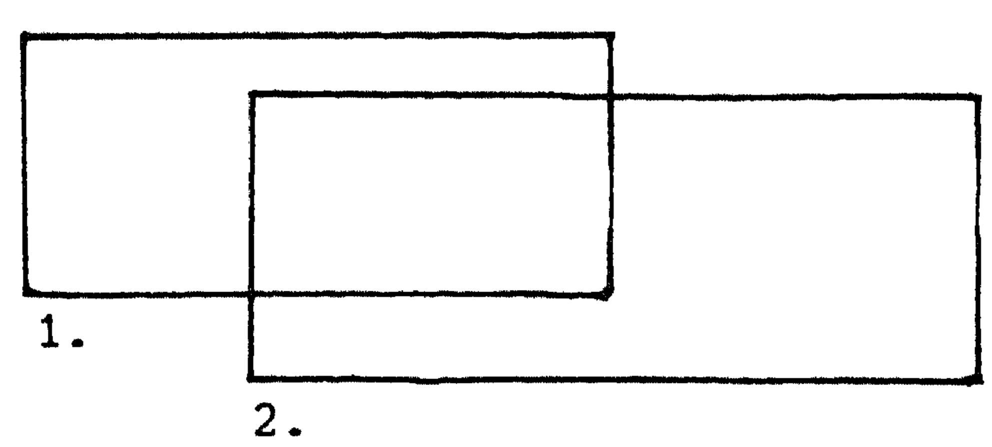


Figure 2.

The present paper describes a treatment of quantification which accounts for this variety of readings (see especially Section 5). This treatment abolishes the dichotomy between distributive and collective verbs that is usually assumed (cf. HAUSSER, 1974; BENNETT, 1975). It shows that the device of deriving distributive readings from collective ones by means of meaning postulates, first introduced by BARTSCH (1973), can be exploited so as to yield the desired results.

Section 7 discusses another phenomenon that has been ignored so far: sentences with indefinite noun phrases may have readings which cannot be represented by means of a formula which has one quantifier for every noun phrase - for instance, when (4a) is read as (4b).

(4a) 600 Dutch firms use 5000 American computers.

(4b) The total number of Dutch firms that use an American computer is 600 and the total number of American computers used by a Dutch firm is 5000.

This phenomenon has been called *cumulative quantification* (cf. SCHA, 1978). In order to generate cumulative readings, our grammar can translate a sequence of noun phrases into one single quantifier, ranging over the cartesian product of the extensions of the nouns.

A grammar for a fragment of English, which displays all the details of our treatment of quantification, is given at the end of the paper (Section 8). This grammar instantiates the framework of Universal Grammar (MONTAGUE, 1970): it specifies compositional syntactic rules which operate in parallel with compositional semantic rules.

Some limitations of our grammar should be mentioned at the outset. For the sake of simplicity, we have restricted ourselves to extensional verbs. (See MONTAGUE, 1973, for a treatment of intensional verbs). We also ignore intensional readings of noun phrases, such as generics. Mass terms, negation, tense, coordinated noun phrases, embedded noun phrases and topic/focus problems are ignored as well.

Because of some of these restrictions, we can use an extensional logical language to represent the truth conditions of the sentences of the fragment. On the other hand, our treatment of quantification involves some semantic notions which are not incorporated in the predicate calculus, the λ-calculus, or in languages such as Montague's Intensional Logic. Therefore, the first part of the paper (Section 2) introduces a new logical language which is used in the rest of the paper. This language has a rather rich type system. It allows variables ranging over n-tuples or sets, for example.

2. THE LOGICAL LANGUAGE

In this section we define the logical language that will be used in the paper to express the truth conditions of English sentences. Every expression of this logical language has a *semantic type*. Given an interpretation of the language, every semantic type has a set of entities as its *domain*. The denotation of an expression is necessarily an element of the domain of its type. The language is defined in two steps. First we define the semantic types. The type system is then used in the definition of the class of expressions of the language.

There are the following *primitive types*:
- the type *truthvalue*, which has the set {TRUE, FALSE} as its domain,
- the type *integer*, which has the set of integers as its domain,
- the type *individual*; an interpretation of the language assigns to this type a set of entities as its domain.

Compound types are constructed as follows:

- If α is a type, then $S(\alpha)$ is a type; its domain is the powerset of the domain of α.
- If $\alpha_1,\ldots,\alpha_n$ are types, then $T_n(\alpha_1,\ldots,\alpha_n)$ is a type; its domain is the cartesian product of the domains of $\alpha_1,\ldots,$ and α_n.
- If α and β are types, then $(\alpha \rightarrow \beta)$ is a type; its domain is the set of functions which map the domain of α into the domain of β.

The definition of the expressions of the language

The definition of the expressions of the language consists of two parts:

- a specification of the primitive expressions (the *terms*) of the language,
- a recursive definition of complex expressions in terms of simpler component expressions.

There are two kinds of terms: *constants* and *variables*. There are two kinds of constants: formal constants and referential constants. The *formal constants* stand for logical or mathematical notions, and receive the same standard denotation for every interpretation of the language. The formal constants are:

- TRUE and FALSE, both with type *truthvalue*
- the decimal representations of the integers, with type *integer*.

The denotation of the referential constants depends on the interpretation which is assigned to the language. The referential constants are translations of referential words of English (e.g. nouns and verbs). These constants are introduced in Section 8.2.

For every type α, there is a denumerably infinite number of variables: X_α, X'_α, X''_α, ... Furthermore, there are the following variables:

- for the type S(*individual*): u,v,w,x,y,z.
- for any type T_n(S(*individual*),...,S(*individual*)): $u_n, v_n, w_n, x_n, y_n, z_n$.
- for any type (T_n(S(*individual*),...,S(*individual*)) $\rightarrow$ *truthvalue*): P_n, Q_n, R_n.
- for the type S(S(*individual*)): U,V,W,X,Y,Z.
- for any type S(T_n(S(*individual*),...,S(*individual*))): $U_n, V_n, W_n, X_n, Y_n, Z_n$.
- for the type (S(*individual*) $\rightarrow$ *truthvalue*): P,Q,R.

Assigning an interpretation to the language consists of two steps:

1. A domain is assigned to the type *individual*. By virtue of the semantic rules of the type system, this defines the domain of *any* type.
2. To every term a *denotation* is assigned; this denotation must be an element of the domain of the type of the term.

Given an interpretation of the language, the denotation of any complex expression is recursively defined in terms of the denotations of the terms.

We now give the syntactic rules which define the compound expressions that belong to the language, and which assign a semantic type to each of those expressions. For the rules which introduce notations that are not widely used, we also describe the semantics.

1. If A is a variable of type α and B has type β, $(\lambda A{:}B)$ has type $(\alpha \longrightarrow \beta)$.
2. If F has type $(\alpha \longrightarrow \beta)$ and A has type α, then F(A) has type β. Instead of F(A) we may also write: (fun:F, arg:A).
3. If x is a variable of type α, A has type $S(\alpha)$, and B has type *truth-value*, then $(\forall x \in A : B)$ has type *truthvalue*.
4. If x is a variable of type α, A has type $S(\alpha)$, and B has type *truth-value*, then $(\exists x \in A : B)$ has type *truthvalue*.
5. If x is a variable of type α, A has type $S(\alpha)$, and B has type *truth-value*, , then $\{x \in A | B\}$ has type $S(\alpha)$.
6. If E and F are expressions, $(E = F)$ has type *truthvalue*.
7. If A_1 has type $\alpha_1, \ldots,$ A_n has type α_n, then $<A_1, \ldots, A_n>$ has type $T_n(\alpha_1, \ldots, \alpha_n)$.
8. If $A_1, \ldots, A_n$ have type α, $\{A_1, \ldots, A_n\}$ has type $S(\alpha)$.
9. If A and B have type *truthvalue*, then $(A \wedge B)$ has type *truthvalue*.
10. If A and B have type *truthvalue*, then $(A \vee B)$ has type *truthvalue*.
11. If A has type *truthvalue*, then $\neg A$ has type *truthvalue*.
12. If A_1 has type $S(\alpha_1), \ldots,$ A_n has type $S(\alpha_n)$, then $(A_1 \times \ldots \times A_n)$ has type $T_n(\alpha_1, \ldots, \alpha_n)$. It denotes the cartesian product of the denotation of A_1, the denotation of $A_2, \ldots,$ and the denotation of A_n.
13. If A has type $S(\alpha)$ then #(A) has type *integer*. It denotes the cardinality of the set denoted by A.
14. If T has type $T_n(\alpha_1, \ldots, \alpha_n)$, then, for any positive integer $i \leq n$: T[i] has type α_i. It denotes the i-th element of the n-tuple denoted by T.

15. For any integer $i \geq 1$, if A is an expression of type $S(T_n(\alpha_1,\ldots,\alpha_n))$ with $n \geq i$, then $\mathrm{proj}_i(A)$ has type $S(\alpha_i)$. It denotes the set containing precisely the i-th elements of the n-tuples in the set denoted by A.

16. If T has type $T_n(\alpha_1,\ldots,\alpha_n)$ and A has type $T_m(\delta_1,\ldots,\delta_m)$ then conc(T,A) has type $T_{m+n}(\alpha_1,\ldots,\alpha_n,\ \delta_1,\ldots,\delta_m)$. It denotes the m+n-tuple whose first n elements, in that order, constitute the n-tuple denoted by T and whose next m elements, in that order, constitute the m-tuple denoted by A.

17. If A has type *truthvalue* and E and F have type α, then (if : A, then : E, else : F) has type α. Its denotation is the denotation of E if A denotes TRUE; it is the denotation of F if A denotes FALSE.

18. If A has type $S(S(\alpha))$, then $\cup(A)$ has type $S(\alpha)$. It denotes the union of the sets in the denotation of A.

19. If A has type $S(\alpha)$, power(A) has type $S(S(\alpha))$. It denotes the powerset of the set denoted by A.

We now introduce some abbreviations of expressions.
We abbreviate:

(if : $\#(A) = 1$, then : $(\forall p \in A{:}B)$; else : FALSE)	as	$(\iota p \in A{:}B)$
$F(<A_1,\ldots,A_n>)$	as	$F[A_1,\ldots,A_n]$
$\{p \in \mathrm{power}(A) \mid B\}$	as	$\{p \subseteq A \mid B\}$
$\neg(A = B)$	as	$(A \neq B)$
$(\exists p \in \mathrm{power}(A){:}B)$	as	$(\exists p \subseteq A{:}B)$
$\{p \in \mathrm{power}(A) \mid \#(p) = 1\}$	as	A^*

If there is no possibility of confusion, parentheses may be left out.

3. DISTRIBUTIVE QUANTIFICATION

Our grammar translates nouns into expressions which denote sets. A quantification which ranges over the extension of a noun is called *distributive*. For example, "every boy" and "some girl" in (1a) lead to distributive quantifications in formula (1b):

(1a) Every boy kisses some girl.

(1b) $\forall x \in \mathrm{BOYS}^*: \exists y \in \mathrm{GIRLS}^*: \mathrm{KISS}[x,y]$.

To highlight some essential features of the structure of the grammar described in Section 8, we shall now sketch how it generates (1a) and how it generates, in parallel, a formula representing the truth conditions of (1a) which is equivalent to (1b).

In our grammar, noun phrases correspond to functions from one-place predicates to truthvalues. For example, (2a) means (2b):

(2a) every boy

(2b) $(\lambda P\colon \forall x \in BOYS^{*}\colon P(x))$

The meaning of (2a) is built up from the meaning of "every" and the meaning of "boy", i.e. (3a) and (3b):

(3a) $(\lambda X\colon (\lambda P\colon \forall x \in X\colon P(x)))$

(3b) $BOYS^{*}$

The expression (3b) is an abbreviation for (3c).

(3c) $\{x \in power(BOYS) \mid \#(x) = 1\}$

This illustrates that our system translates a noun into an expression which denotes a set of singleton sets, rather than a set of individuals. As we shall show in Section 5, this has advantages for the description of the relation between collective and distributive uses of verbs.

(3a) is a function which is applicable to sets of sets. The application of (3a) to (3b) yields (3d), which is equivalent to (2b).

(3d) $(fun\colon(\lambda X\colon(\lambda P\colon \forall x \in X\colon P(x))),$
$arg\colon BOYS^{*}).$

In a similar way, (4a) is constructed, with the meaning (4b):

(4a) some girl

(4b) $(\lambda P\colon \exists y \in GIRLS^{*}\colon P(y)).$

Noun phrases are joined by the grammar to form noun-phrase sequences. The meaning of a sequence of n noun phrases is a function from n-place relations to truthvalues (n-place relations are rendered as one-place predicates on n-tuples). For example, the meaning of (5a) is represented by (5b).

(5a) <every boy, some girl>

(5b) $(\lambda Q_2\colon \forall x \in \mathrm{BOYS}^*\colon \exists y \in \mathrm{GIRLS}^*\colon Q_2(\langle x,y\rangle))$

Verbs correspond to n-place relations. For instance, the verb "kiss" is represented as the constant KISS, which has the type

$$(T_2(\mathit{individual},\ \mathit{individual}) \rightarrow \mathit{truthvalue}).$$

Verbs are combined with noun-phrase sequences to form sentences. For instance, the noun-phrase sequence (5a) is combined with the verb form "kisses" to yield the sentence (1a), while in the semantics (5b) is combined with KISS to yield (6), which is equivalent to (1b).

(6) $(\mathrm{fun}\colon (\lambda Q_2\colon \forall x \in \mathrm{BOYS}^*\colon \exists y \in \mathrm{GIRLS}^*\colon Q_2(\langle x,y\rangle)),$
$\mathrm{arg}\colon \mathrm{KISS})$

We now give a list of some determiners which give rise to distributive quantifications:

(7a) "each", "every" and "all"

(7b) $(\lambda X\colon(\lambda P\colon \forall x \in X\colon P(x)))$

(8a) "a", "some", and the "empty determiner"

(8b) $(\lambda X\colon(\lambda P\colon \exists x \in X\colon P(x)))$

(9a) "less than three"

(9b) $(\lambda X\colon(\lambda P\colon \#(U(\{x \in X \mid P(x)\})) < 3))$

(10a) "three"

(10b) $(\lambda X\colon(\lambda P\colon \#(U(\{x \in X \mid P(x)\})) = 3))$

(11a) "the" (if combined with a singular noun phrase)

(11b) $(\lambda X\colon(\lambda P\colon \iota x \in X\colon P(x)))$

REMARKS.

1. "all", "some", the "empty determiner", "less than three" and "three" also have other readings, which are called *collective*. These are treated in Section 4.
2. Our grammar does not assign a distributive reading to plural noun phrases with "the". Such noun phrases get a collective reading only. This reading is described in Section 4. Our treatment of "the" + plural noun phrases is explained in Sections 5 and 6.

4. COLLECTIVE QUANTIFICATION

It has been widely noticed that distributive quantification is not sufficient for characterizing the truth conditions of English sentences.[2] For example, the most plausible reading of (1a) cannot be represented by (1b):

(1a) All boys gather

(1b) $\forall x \in BOYS^{*}: GATHER[x]$

(1a) does not express that every single boy gathers, but rather that the group of all boys gathers. In our grammar, "all boys" has therefore two readings: (2b), which was already introduced in the previous section, and (2c):

(2a) all boys

(2b) $(\lambda P: \forall x \in BOYS^{*}: P(x))$

(2c) $(\lambda P: P(U(BOYS^{*})))$

It may be noticed that (2c) is equivalent to (2d):

(2d) $(\lambda P: P(BOYS)))$.

If reading (2b) is assigned to "all boys", sentence (1a) gets the reading (1b); if reading (2c) is assigned to "all boys", (1a) gets the alternative reading (3):

(3) GATHER[BOYS]

HAUSSER (1974) and BENNETT (1975) treat "all" as ambiguous too. An important difference between their approach and the one presented here, consists in the fact that we treat individual things and groups of things as entities of the same kind: both are represented by sets of individuals. (In the case of an individual thing, this is a set with one element.) Therefore, any function which is applicable to individuals is also applicable to groups, and vice versa. Our grammar can thus allow distributive as well as collective noun phrases to combine with any verb.[3]
Hausser and Bennett make a syntactic distinction between distributive noun phrases and collective noun phrases, and subdivide the category of n-place verbs accordingly, distinguishing between verbs which require a distributive

subject and verbs which require a collective subject, between verbs which require a distributive direct object and verbs which require a collective direct object, etc. Since we do not make such subdivisions, our grammar is more simple. It is also more tolerant; for instance, it accepts (5a) as a grammatical English sentence, assigning it the meaning (5b):

(5a) Each boy gathers

(5b) $\forall x \in BOYS^{*}: GATHER[x]$

And, as we mentioned already, (1a) is treated as ambiguous in our grammar; it has both the implausible reading (1b) and the more likely reading (3).

HAUSSER (1974) or BENNETT (1975) would rule out (5a), and the (1b) reading of (1a), as ungrammatical. They would construe the verb "gather" as requiring a collective subject, while subcategorizing the noun phrase "each boy", as well as the (2b) reading of "all boys", as distributive.

We prefer to view the oddity of (5a) and of the (1b) reading of (1a) as a special case of "semantic anomaly", of the same kind as demonstrated in (6):

(6) Colourless green ideas sleep furiously.

It may be possible to treat this phenomenon at the semantic level, independent of our choice of syntactic categories (cf. THOMASON, 1972; WALDO, 1979).

The tolerance of our treatment has the advantage that (7a) is analysed as (7b) without any ad hoc complications in the syntax. (BARTSCH (1973), BENNETT (1975), and HAUSSER (1974) must assign a special syntactic category to "collective nouns" like "group" and "committee".)

(7a) Each committee gathers

(7b) $\forall x \in COMMITTEES^{*}: GATHER[x]$

Similarly, (8a) is ambiguous between (8b) and (8c):

(8a) All committees gather

(8b) $\forall x \in COMMITTEES^{*}: GATHER[x]$

(8c) $GATHER[COMMITTEES]$

We now give a list of some collective readings of determiners.

(9a) "all" or "the" (if combined with a plural noun)

(9b) $(\lambda X: (\lambda P: P(U(X))))$

(10a) "six"

(10b) $(\lambda X: (\lambda P: \exists u \in \{v \subseteq U(X) \mid \#(v) = 6\} : P(u)))$

(10c) $(\lambda X: (\lambda P: \#(U(\{u \subseteq U(X) \mid P(u)\})) = 6))$

(11a) "less than six"

(11b) $(\lambda X: (\lambda P: \exists u \in \{v \subseteq U(X) \mid \#(v) < 6\} : P(u)))$

(11c) $(\lambda X: (\lambda P: \#(U(\{u \subseteq U(X) \mid P(u)\})) < 6))$

The first one of the collective readings of indefinite noun phrases (i.e. reading (10b) of (10a), reading (11b) of (11a)) is a familiar one, widely recognized as necessary to analyse a sentence like (12a) as (12b).

(12a) Six boys gather

(12b) $\exists u \in \{v \subseteq \text{BOYS} \mid \#(v) = 6\} : \text{GATHER}[u]$

However, (12a) has another reading, represented by (13):

(13) $\#(U(\{u \subseteq \text{BOYS} \mid \text{GATHER}[u]\})) = 6$

In this interpretation, (12a) would be true in a situation where, for example, two groups of three boys were each gathering. In order to yield such sentence readings, the indefinite noun phrases need their second collective reading. For instance, for reading (13) of (12a), we need reading (10c) of (10a).

5. MEANING POSTULATES

An unconventional feature of our grammar consists in the fact that plural noun phrases with "the" are treated as having only a collective reading: (1a) means (1b) and nothing else.

(1a) the boys

(1b) $(\lambda P: P(\text{BOYS}))$

Because we do not distinguish between verbs which require a collective subject and verbs which require a distributive subject, or between verbs which require a collective object and a distributive object, etc., the noun phrase "the boys" can nevertheless be combined with verbs which are normally viewed as distributive, such as "walk". For instance, (2a) is generated by our grammar, with the meaning (2b).

(2a) The boys walk

(2b) WALK[BOYS]

Thus, according to our theory, (2a) is *not* synonymous with

(3) Every boy walks

which is as it should be, since (2a) is more vague than (3) - though it certainly comprises (3) as one of its "instances". The relation between (2a) and (3) is similar to the relation between (4a) and (4b):

(4a) The committee walks

(4b) Every member of the committee walks.

If in a given domain of discourse (2a) has similar truth conditions as (3), this can be described by a *meaning postulate* like (5):

(5) $\mathrm{WALK}[x] = (\#(x) \neq 0 \wedge \forall y \in x^{*}: \mathrm{WALK}[y])$

Given meaning postulate (5), (2b) can be translated into (6):

(6) $(\forall y \in \mathrm{BOYS}^{*}: \mathrm{WALK}[y]) \wedge \#(\mathrm{BOYS}) \neq 0.$

Meaning postulates like (5) define the application of a predicate to a collection in terms of the application of that same predicate to the smallest parts of that collection. The use of such meaning postulates to derive distributive readings of a sentence from collective ones was first proposed by BARTSCH (1973). She does not exploit the full potential of this idea, however, since she only considers intransitive verbs. For all distributive intransitive verbs she postulates a semantic property similar to the one we expressed for "walk" in formula (5). Later on in this section we show that this approach yields its most interesting results in the analysis of transitive verbs.

Because of the use of meaning postulates, our grammar can treat collective verbs like "gather" and distributive verbs like "walk" in exactly the same way at the syntactic level. The difference only shows in the semantics. For the predicate which represents "walk" there is a meaning postulate like (5) which defines the walking of a group in terms of the walking of its members; for the predicate which represents "gather" there is no such meaning postulate.

As we mentioned in the previous section, HAUSSER (1974) and BENNETT (1975) make a syntactic distinction between verbs like "walk" which require a distributive subject and verbs like "gather" which require a collective subject. Correspondingly, they treat plural noun phrases with "the" as ambiguous between a collective reading and a distributive reading; a distributively interpreted "the" + plural noun phrase indicates universal quantification over the extension of the noun. If we were to follow that line, we would directly assign two readings to (1a): (1b) and (8):

(8) $(\lambda P: \forall x \in BOYS^*: P(x))$

For intransitive verbs, such a treatment may be possible. If we consider transitive verbs, however, the Hausser/Bennett approach breaks down. Consider, for example, the sentences (9a), (10a), (11a), and their intuitively plausible readings (9b), (10b), (11b) (cf. Figure 1 and Figure 2 in Section 1 of this paper):

(9a) The sides of R1 run parallel to the sides of R2

(9b) $(\forall x \in SR1^*: \exists y \in SR2^*: PAR[x,y]) \wedge$
$(\forall y \in SR2^*: \exists x \in SR1^*: PAR[x,y]) \wedge$
$\#(SR1) \neq 0 \;\; \wedge \;\; \#(SR2) \neq 0.$

(10a) The sides of R1 cross the sides of R2

(10b) $\exists x \in SR1^*: \exists y \in SR2^*: CROSS[x,y]$

(11a) The squares contain the circles

(11b) $(\forall y \in CIRCLES^*: \exists x \in SQUARES^*: CONTAIN[x,y])$
$\wedge \; \#(CIRCLES) \neq 0$

The Hausser/Bennett approach, which treats "the" + plural as indicating universal quantification in the case of distributive verbs, does not generate these readings. Making "the" + plural ambiguous between universal and existential quantification would not solve the problem; it would generate a lot of wrong readings, which would have to be filtered out somehow.

The grammar we present in this paper, on the other hand, treats plural noun phrases with "the" as referring to the set of all entities in the extension of the noun. Therefore, sentences like (9a), (10a), (11a) above

Syntactic tree	Logical translation
for β = *singular* or *plural*:	
determiner (no) [*form*:β]	$(\lambda X: (\lambda P: \neg\exists x \in X: P(x)))$
determiner (no) [*form*:*plural*]	$(\lambda X: (\lambda P: \neg\exists Y \subseteq U(X): P(Y)))$
for α = "the" or "the one":	
determiner (α) [*form*:*singular*]	$(\lambda X: (\lambda P: \iota x \in X: P(x)))$
for α = "both the" or "both":	
determiner (α) [*form*:*plural*]	$(\lambda X: (\lambda P: (\text{if}: \#(X) = 2,$ then: $(\forall x \in X: P(x))$ else: FALSE)))

Syntactic trees of the category *numeral* correspond to logical expressions of type (S(S(*individual*)) $\longrightarrow$ *truthvalue*). There is one primitive tree of this category:

numeral (ε) [*form*:*plural*]

which translates into

$(\lambda X: \#(U(X)) > 0)$

(ε stands for the empty string).

8.3. Grammar rules

A. Rules which construct a tree of the category *numeral*, given a tree α of the category *number*.

A1. $SYN_{A1}(\alpha)$ = *numeral* (α) [*form*:β], where $\beta = \alpha$.*form*

$SEM_{A1}(\alpha) = (\lambda X: \#(U(X)) = \alpha')$

A2. $SYN_{A2}(\alpha)$ = *numeral* (more, than, α) [*form*:β]

where $\beta = \alpha$.*form*

$SEM_{A2}(\alpha') = (\lambda X: \#(U(X)) > \alpha')$

A3. $SYN_{A3}(\alpha')$ = *numeral* (less, than, α) [*form*:β]

where $\beta = \alpha$.*form*

$SEM_{A3}(\alpha') = (\lambda X: \#(U(X)) < \alpha')$.

B. Rule schema which constructs a tree of category *compnum* given n trees $\alpha_1,\ldots,\alpha_n$ of category *numeral*.

$SYN_{Bn}(\alpha_1,\ldots,\alpha_n) = compnum_n(\alpha_1,\ldots,\alpha_n)$

$SEM_{Bn}(\alpha'_1,\ldots,\alpha'_n) = (\lambda X_n: \alpha'_1(proj_1(X_n)) \wedge \ldots \wedge \alpha'_n(proj_n(X_n)))$.

are analysed as (12a,b,c):

(12a) $PAR[SR1, SR2]$

(12b) $CROSS[SR1, SR2]$

(12c) $CONTAIN[CIRCLES, SQUARES]$

In this analysis, plural noun phrases with "the" do not lead directly to any quantificational structure in the logical formula. Therefore, the differences between the meanings of (9a), (10a) and (11a), as shown in (9b), (10b) and (11b), can be explained by postulates about the meanings of the different verbs. For example, about the meanings of "run parallel to", "cross", and "contain", respectively, we may formulate the following meaning postulates.

(13a) $$PAR[u,v] = ((\forall x \in u^*: \exists y \in v^*: PAR[x,y]) \wedge (\forall y \in v^*: \exists x \in u^*: PAR[x,y]) \wedge \#(u) \neq 0 \;\wedge\; \#(v) \neq 0)$$

(13b) $$CROSS[u,v] = (\exists x \in u^*: \exists y \in v^*: CROSS[x,y])$$

(13c) $$CONTAIN[u,v] = ((\forall y \in v^*: \exists x \in u^*: CONTAIN[x,y]) \wedge \#(v) \neq 0)$$

These meaning postulates express some elementary semantic properties of the verbs. For instance, a compound entity y can be said to be contained in a compound entity x if every part of x is contained in some part of y. On the other hand, a compound entity x may be said to cross a compound entity y if some part of x crosses some part of y. We represent compound entities as sets; the part-whole relation is represented as the subset-relation. (Therefore, the individuals, being the smallest parts, are represented as singleton sets.)

Given the meaning postulates (13a,b,c), we can derive from the collective readings (12a,b,c) the readings (9b), (10b), (11b). If we disregard the non-emptiness conditions on the sets in formula (9b) above, it displays the quantificational structure that LANGENDOEN (1978) proposes to assign to every distributively interpreted transitive verb with two "the" + plural noun phrases. Our sentences (10a) and (11a) are counterexamples to his proposal.

So far we have considered sentences with only "the" + plural noun phrases. The same meaning postulates can be applied, however, yielding equally desirable results, in the case of sentences where "the" + plural noun phrases are combined with distributively interpreted noun phrases. Assume, for example, for the collective use of "date" the meaning postulate (14), analoguous to (13b) above.

(14) $\quad \mathrm{DATE}[u,v] = (\exists x \in u^{*}: \exists y \in v^{*}: \mathrm{DATE}[x,y])$.

Consider now sentence (15a), with the distributive interpretation of the noun phrase "600 girls". Primarily, this sentence would be analysed as (15b); if we take meaning postulate (14) into account, it is equivalent to (15c).

(15a) The boys date 600 girls

(15b) $\#(\{y \in \mathrm{GIRLS}^{*} \mid \mathrm{DATE}[\mathrm{BOYS},y]\}) = 600$

(15c) $\#(\{y \in \mathrm{GIRLS}^{*} \mid \exists x \in \mathrm{BOYS}^{*}: \mathrm{DATE}[x,y]\}) = 600$.

Note, also in this case, the discrepancy between "the" + plural noun phrases and noun phrases with "every"; (15a) is by no means synonymous to (16a), which gets the analysis (16b).

(16a) Every boy dates 600 girls

(16b) $\forall x \in \mathrm{BOYS}^{*}: \#(\{y \in \mathrm{GIRLS}^{*} \mid \mathrm{DATE}[x,y]\}) = 600$.

On the other hand, sentence (17) is ambiguous between a reading synonymous with (15a) and a reading synonymous with (16a).

(17) All boys date 600 girls.

Next, we consider the application of the meaning postulates for indefinite collective quantifiers. In this case, they yield interesting results as well. Consider, for instance, reading (18b) of sentence (18a), in combination with meaning postulate (18c). (This meaning postulate was introduced earlier in this section as meaning postulate (5).)

(18a) Ten boys walk

(18b) $\exists x \in \{y \subseteq \mathrm{BOYS} \mid \#(y) = 10\} : \mathrm{WALK}[x]$

(18c) $\mathrm{WALK}[u] = (\#(u) \neq 0 \wedge \forall y \in u^{*}: \mathrm{WALK}[y])$.

Given (18c), (18b) is equivalent to (19):

(19) $\#(\{y \in \mathrm{BOYS}^{*} \mid \mathrm{WALK}[y]\}) \geq 10$.

Because of this, we could treat, in our discussion of distributive quantification (Section 3), an indefinite noun phrase like (20a) as having (20b) as its only distributive reading. We do not need reading (20c).

(20a) ten boys

(20b) $(\lambda P: \#(\{x \in BOYS^* \mid P(x)\}) = 10)$

(20c) $(\lambda P: \#(\{x \in BOYS^* \mid P(x)\}) \geq 10)$.

6. MATTERS OF SCOPE

It has been noted before that the order of the quantifiers in the formula which represents the most plausible reading of a sentence often coincides with the order of the corresponding noun phrases in the sentence.[4] For instance, (1a) means (1b), whereas (2a) means (2b).

(1a) Some circle contains every square

(1b) $\exists x \in CIRCLES^*: \forall y \in SQUARES^*: CONTAIN[x,y]$

(2a) Every square is contained in some circle

(2b) $\forall y \in SQUARES^*: \exists x \in CIRCLES^*: CONTAIN[x,y]$.

Many formal treatments of quantification have ignored this phenomenon, and have simply allowed all possible quantifier orders - because all these orders correspond to *possible*, though perhaps implausible, readings of the sentence.

PARTEE (1973) introduced the distinction between a "loose" version of the syntax (generating all possible readings) and a "strict" version (generating preferred readings only). In the present paper, we also use this idea. The strict version of our syntax is the version without the rule F4, which in Section 8.3 is designated as "optional". The loose version is obtained by adding this rule. In the strict version, scope of nested quantifiers corresponds to the linear order of the noun phrases in the sentence. In the loose version of the syntax, all possible permutations of quantifiers are generated.

A consideration of the effects of word order on preferred quantifier scope gives further support to our thesis, presented in the previous section, that distributive interpretations of sentences involving plural noun phrases with "the" are best constructed indirectly, on the basis of an originally collective interpretation.

With respect to scope, noun phrases with "the" behave differently from noun phrases with "every", "some" or "less than five"; the position of noun phrases with "the" does not seem to influence quantifier scope. For instance, (3a) and (3b) both have the same meaning, (3c).

(3a) The squares contain the circles

(3b) The circles are contained in the squares

(3c) $(\forall y \in \text{CIRCLES}^*: \exists x \in \text{SQUARES}^*: \text{CONTAIN}[x,y]) \wedge \#(\text{CIRCLES}) \neq 0.$

This is explained by the fact that (3a) and (3b) are both analysed as (4):

(4) CONTAIN[SQUARES, CIRCLES]

Application of meaning postulate (5) (introduced in section 5, as meaning postulate (13c)) leads, for both sentences, to (3c).

(5) $\text{CONTAIN}[u,v] = ((\forall y \in v^*: \exists x \in u^*: \text{CONTAIN}[x,y]) \wedge \#(v) \neq 0).$

Another example: (6a) means (6c), and (6b) means the same:

(6a) Some square contains the circles

(6b) The circles are contained in some square

(6c) $(\exists x \in \text{SQUARES}^*: \forall y \in \text{CIRCLES}^*: \text{CONTAIN}[x,y]) \wedge \#(\text{CIRCLES}) \neq 0.$

This is explained by the fact that both (6a) and (6b) have only one "real" quantifying noun phrase: "some square"; they are both analysed as (7):

(7) $\exists x \in \text{SQUARES}^*: \text{CONTAIN}[x, \text{CIRCLES}].$

Using meaning postulate (5) about the collective use of CONTAIN, (7) can be shown to be equivalent to (6c). This example shows how our theory predicts the fact that "the" + plural noun phrases tend to have narrow scope readings, regardless of their position in the surface order: Quantifiers which appear as a result of a meaning postulate are always inside the scope of quantifiers which are explicitly mentioned in the sentence.

7. CUMULATIVE QUANTIFICATION

Consider sentence (1):

(1) 600 Dutch firms have 5000 American computers.

Given the treatment of indefinite noun phrases presented in Sections 3 and 4, this sentence has nine different readings in the strict version of the syntax: each of the two noun phrases has a distributive reading and two collective readings. The sentence has other readings, however, which are distinct from these, and not derivable from any of them - nor from any of the extra readings produced by the loose version of the syntax. For instance, (1) has a reading which can be paraphrased as:

(2) The number of Dutch firms which have an American computer is 600, and the number of American computers possessed by a Dutch firm is 5000.

We call this kind of quantification *cumulative quantification* (cf. SCHA, 1978). It cannot be expressed in a formula containing quantifiers with a one-to-one correspondence to the noun phrases in the sentence. In our logical language, reading (2) of (1) may be expressed as (3), for instance:

(3) $$\#(\{x \in \mathrm{DUFIS}^* \mid \exists y \in \mathrm{AMCOS}^*: \mathrm{HAVE}[x,y]\}) = 600 \;\wedge$$
$$\#(\{y \in \mathrm{AMCOS}^* \mid \exists x \in \mathrm{DUFIS}^*: \mathrm{HAVE}[x,y]\}) = 5000.$$

A problem with this expression is that its structure does not immediately suggest a way to derive it from the surface structure of (1). Using quantification over the cartesian product of DUFIS and AMCOS, however, we can construct an equivalent formula with a quite different structure:

(4) $$(\mathrm{fun}: (\lambda X_2: \#(\mathrm{proj}_1(X_2)) = 600 \wedge \#(\mathrm{proj}_2(X_2)) = 5000),$$
$$\mathrm{arg}: \{u_2 \in \mathrm{DUFIS}^* \times \mathrm{AMCOS}^* \mid \mathrm{HAVE}(u_2)\}).$$

Our grammar generates a reading of (1) which is equivalent to (4). It contains a rule which constructs "compound numericals" out of a sequence of numerical expressions. For instance, this rule combines "600" and "5000" into a compound numerical with the meaning (5):

(5) $$(\lambda X_2: \#(\mathrm{proj}_1(X_2)) = 600 \wedge \#(\mathrm{proj}_2(X_2)) = 5000).$$

Similarly, nouns are combined into "compound nouns". For instance, "Dutch firms" and "American computers" are combined into a compound noun with the meaning (6):

(6) $\qquad DUFIS^* \times AMCOS^*$.

Compound numericals and compound nouns are combined to form noun phrase sequences. For instance, the compound numerical consisting of "600" and "5000" is combined with the compound noun consisting of "Dutch firms" and "American computers", to form a noun phrase sequence consisting of "600 Dutch firms" and "5000 American computers". Semantically, this means that (5) and (6) are combined to form the function from 2-place relations to truth values (7):

(7)
$$(\lambda R_2\text{: (fun: } (\lambda X_2\text{: } \#(proj_1(X_2)) = 600 \wedge \#(proj_2(X_2)) = 5000), \text{ arg: } \{u_2 \in DUFIS^* \times AMCOS^* \mid R_2(u_2)\})).$$

A noun-phrase sequence is combined with a verb to form a sentence. For instance, the noun-phrase sequence consisting of "600 Dutch firms" and "5000 American computers" is combined with the verb "have" to form (1). Semantically, this means that (7) is combined with the two-place relation HAVE to form (8), which is equivalent to (4).

(8)
$$\text{(fun: } (\lambda R_2\text{: (fun: } (\lambda X_2\text{: } \#(proj_1(X_2)) = 600 \wedge \#(proj_2(X_2)) = 5000), \text{ arg: } \{u_2 \in DUFIS^* \times AMCOS^* \mid R_2(u_2)\})), \text{ arg: HAVE).}$$

8. A GRAMMAR FOR A FRAGMENT OF ENGLISH

8.1. Introduction

In this section we present a formal grammar for a fragment of English. This grammar implements our ideas about quantification as described in Sections 3 - 7. The framework we use is a variant of Montague grammar, of the kind proposed by PARTEE (1973). The grammar defines a class of ordered trees; the sequences of terminals of these trees are the sentences of our fragment of English. The grammar also defines, for each of these trees, a formula representing the truth conditions of the sentence.

The *syntactic rules* of the grammar specify how syntactic trees may be built from smaller syntactic trees. A *lexicon* specifies the primitive syntactic trees, and logical formulas representing their meanings. Corresponding to every syntactic rule SYN_i there is a *semantic rule* SEM_i which specifies how the meaning of a constituent produced by SYN_i is defined in terms of the meanings of the arguments of SYN_i. A syntactic rule operates on trees belonging to particular syntactic categories, and it may specify conditions further constraining its domain.

The nodes of the syntactic trees are labelled with syntactic categories, and may specify the value of a syntactic attribute. We use a "flat" notation for the trees. For instance, an expression of the form *cat*(TREE1, TREE2)[*att : val*] stands for a tree with a top node labelled *cat*, with subtrees TREE1 and TREE2, and with the value *val* for the syntactic attribute *att*.

Some notations that we shall use in describing the syntactic rules: If X is a syntactic tree and A is a syntactic attribute applicable to X,

X.A indicates the value of attributive A for tree X;
#daughternodes (X) indicates the number of daughternodes of X;
X[i] indicates the i-th subtree of X.

For instance, if X is a tree of the form *cat*(T_1,...,T_n)[*att:val*]:

X.*att*=*val*
#daughternodes (X)=n;
X[i]=T_i.

The arguments of a syntactic rule are indicated by Greek letters: $\alpha,\beta,\ldots$. The corresponding semantic translations are indicated by primed Greek letters: $\alpha',\beta',\ldots$.

8.2. Lexicon

Intransitive verbs have category $verb_1$; they translate into expressions of type (T_1(S(*individual*)) $\rightarrow$ *truthvalue*). For instance:

Syntactic tree	Logical translation
$verb_1$ (gather)[*form:plural*]	GATHER
$verb_1$ (gathers)[*form:singular*]	GATHER
$verb_1$ (walk)[*form:plural*]	WALK
$verb_1$ (walks)[*form:singular*]	WALK

Two-place verbs have category $verb_2$; they translate into expressions of type (T_2(S(*individual*), S(*individual*)) $\rightarrow$ *truthvalue*). For instance:

Syntactic tree	Logical translation
$verb_2$ (contains) [*form:singular*]	CONTAIN
$verb_2$ (is contained in) [*form:singular*]	$(\lambda x_2\colon \mathrm{CONTAIN}[x_2[2], x_2[1]])$

Three-place verbs have category $verb_3$; they translate into expressions of type (T_3(S(*individual*), S(*individual*), S(*individual*)) $\longrightarrow$ *truthvalue*). For instance:

Syntactic tree	Logical translation
$verb_3$ (give) [*form:plural*]	GIVE
$verb_3$ (gives) [*form:singular*]	GIVE

Nouns are translated into expressions of type S(S(*individual*)). For instance:

Syntactic tree	Logical translation
noun (circle) [*form:singular*]	CIRCLES*
noun (circles) [*form: plural*]	CIRCLES*
noun (committee) [*form:singular*]	COMMITTEES*
noun (committees) [*form:plural*]	COMMITTEES*

Numbers translate into expressions of type *integer*. For instance:

Syntactic tree	Logical translation
number (one) [*form:singular*]	1
number (two) [*form:plural*]	2

Determiners translate into expressions of type (S(S(*individual*)) $\longrightarrow$ ((S(*individual*) $\longrightarrow$ *truthvalue*) $\longrightarrow$ *truthvalue*). For instance:

Syntactic tree	Logical translation
for α = "each" or "every"	
determiner (α) [*form:singular*]	$(\lambda X\colon (\lambda P\colon \forall x \in X\colon P(x)))$
for α = "a" or "some"	
determiner (α) [*form:singular*]	$(\lambda X\colon (\lambda P\colon \exists x \in X\colon P(x)))$
determiner (some) [*form:plural*]	$(\lambda X\colon (\lambda P\colon \exists Y \subseteq U(X)\colon P(Y)))$

C. Rules which construct a tree of the category *determiner*, given a tree α of the category *numeral*.

β stands for the value of α.*form*.

C1. $SYN_{C1}(\alpha) = \textit{determiner}\ (\alpha)\ [\textit{form}:\beta]$

$SEM_{C1}(\alpha') = (\lambda X: (\lambda P: \alpha'(\{x \in X \mid P(x)\})))$

C2. $SYN_{C2}(\alpha) = \textit{determiner}\ (\alpha)\ [\textit{form}:\beta]$

$SEM_{C2}(\alpha') = (\lambda X: (\lambda P: \exists u \in \{v \subseteq \cup(X) \mid \alpha'(v)\} : P(u)))$

C3. $SYN_{C3}(\alpha) = \textit{determiner}\ (\alpha)\ [\textit{form}:\beta]$

$SEM_{C3}(\alpha') = (\lambda X: (\lambda P: \alpha'(\{u \subseteq \cup(X) \mid P(u)\})))$

C4. Rule schema: For γ = "the", "all the" or "all"

Condition: α.*form* = *plural*

$SYN_{C4}(\alpha) = \textit{determiner}\ (\gamma,\alpha)\ [\textit{form:plural}]$

$SEM_{C4}(\alpha') = (\lambda X: (\lambda P: (\text{if: } \alpha'(X), \text{ then: } P(\cup(X)) \text{ else: FALSE})))$

C5. Rule schema: For γ = "all the" or "all"

Condition: α.*form* = *plural*

$SYN_{C5}(\alpha) = \textit{determiner}\ (\gamma,\alpha)\ [\textit{form:plural}]$

$SEM_{C5}(\alpha') = (\lambda X: (\lambda P: (\text{if: } \alpha'(X) \text{ then: } (\forall x \in X: P(x)) \text{ else: FALSE})))$

D. Rule schema which constructs a tree of the category *compnoun* out of n trees, $\alpha_1,\ldots,\alpha_n$, of the category *noun*.

$SYN_{Dn}(\alpha_1,\ldots,\alpha_n) = \textit{compnoun}\ (\alpha_1,\ldots,\alpha_n)$

$SEM_{Dn}(\alpha'_1,\ldots,\alpha'_n) = \alpha'_1 \times \ldots \times \alpha'_n$

E. Rule which constructs a tree of the category *np* out of a tree α of the category *noun* and a tree β of the category *determiner*.

Condition: α.*form* = β.*form*

$SYN_E(\alpha) = np(\beta,\alpha)\ [\textit{form}:\gamma]$, where $\gamma = \alpha$.*form*

$SEM_E(\alpha') = \beta'(\alpha')$

F. Rules which construct trees of the category *nps*

F1. Given a tree α of category *np:*

$SYN_{F1}(\alpha) = nps(\alpha)$

$SEM_{F1}(\alpha') = (\lambda P_1: \alpha'((\lambda x: P_1(<x>))))$

F2. For any n > 1: Given a tree α of category $\textit{compnum}_n$ and a tree β of category $\textit{compnoun}_n$:

Condition $\forall i \leq n$: $\alpha[i]$.*form* = $\beta[i]$.*form*

$SYN_{F2.n}(\alpha,\beta) = nps\ (np(\alpha[1],\beta[1]),\ldots,np(\alpha[n],\beta[n]))$

$SEM_{F2.n}(\alpha',\beta') = (\lambda R_n\colon \alpha'(\{x_n \in \beta' \mid R_n(x_n)\}))$

F3. For any $m \geq 1$ and $n \geq 1$: Given a tree α of category *nps* and a tree β of category *nps*:

Condition: #daughternodes $(\alpha) = m \wedge$ #daughternodes $(\beta) = n$.

$SYN_{F3.m.n.}(\alpha,\beta) = nps\ (\alpha[1],\ldots,\alpha[m],\beta[1],\ldots,\beta[m])$

$SEM_{F3.m.n.}(\alpha',\beta') = (\lambda R_p\colon \alpha'(\lambda x_m\colon \beta'(\lambda y_n\colon R_p(conc(x_m,y_n))))),$
where $p = m+n$.

F4. (Optional).[5] For any $m \geq 1$ and $n \geq 1$: Given a tree α of category *nps* and a tree β of category *nps:*

Condition: #daughternodes $(\alpha) = m \wedge$ #daughternodes $(\beta) = n$

$SYN_{F4.m.n.}(\alpha,\beta) = nps(\alpha[1],\ldots,\alpha[m],\beta[1],\ldots,\beta[n])$

$SEM_{F4.m.n.}(\alpha',\beta') = (\lambda R_p\colon \beta'(\lambda y_n\colon \alpha'(\lambda x_m\colon R_p(conc(x_m,y_n))))),$
where $p = m+n$.

G. For any $n \geq 1$: Given a tree α of category *nps* and a tree β of category $verb_n$:

Condition: #daughternodes $(\alpha) = n \wedge \alpha[1].form = \beta.form$

$SYN_G(\alpha,\beta) = sentence(\alpha[1],\beta,\alpha[2],\ldots,\alpha[n])$

$SEM_G(\alpha',\beta') = \alpha'(\beta')$.

8.4. Examples

In this section we show the operation of the rules of the grammar, by describing some derivations. First, we show the derivation of a reading of (1) where both noun phrases are taken to indicate distributive quantification.

(1) Less than five boys date more than six girls.

The word "five" is a terminal in the elementary syntactic tree (2a), specified in the lexicon with the meaning (2b).

(2a) *number* (five) [*form:plural*]

(2b) 5

Rule A3 constructs on the basis of (2ab) the tree (3a)[6] with meaning (3b).

(3a) *numeral* (less, than, *number* (five)) [*form:plural*]

(3b) $(\lambda X\colon \#(U(X)) < 5)$.

Rule C1 constructs (4ab) on the basis of (3ab).

(4a) *determiner* (*numeral* (less, than, *number* (five))) [*form:plural*]

(4b) $(\lambda X\colon (\lambda P\colon (\text{fun}\colon (\lambda Y\colon \#(U(Y)) < 5),$
$\quad \text{arg}\colon \{x \in X \mid P(x)\})))$

(4b) is equivalent to (4c).

(4c) $(\lambda X\colon (\lambda P\colon \#(U(\{x \in X \mid P(x)\})) < 5)$

The word "boys" is a terminal in the elementary syntactic tree (5a), specified in the lexicon with meaning (5b).

(5a) *noun* (boys) [*form:plural*]

(5b) BOYS^{*}

Rule E constructs (6a) from (4a) and (5a), and constructs (6b) from (4c) and (5b).

(6a) *np* (*determiner* (*numeral* (less, than, *number* (five))),
noun (boys)) [*form:plural*]

(6b) $(\text{fun}\colon (\lambda X\colon (\lambda P\colon \#(U(\{x \in X \mid P(x)\})) < 5)),$
$\text{arg}\colon \text{BOYS}^{*})$

(6b) is equivalent to (6c)

(6c) $(\lambda P\colon \#(U(\{x \in \text{BOYS}^{*} \mid P(x)\})) < 5)$.

From (6a), rule F1 constructs (7a); from (6c) it constructs (7b), which is equivalent to (7c).

(7a) *nps* (*np* (*determiner* (*numeral* (less, than, *number* (five))),
noun (boys)) [*form:plural*])

(7b) $(\lambda P_1\colon (\text{fun}\colon (\lambda P\colon \#(U(\{x \in \text{BOYS}^{*} \mid P(x)\})) < 5),$
$\quad \text{arg}\colon (\lambda y\colon P_1(<y>))))$

(7c) $(\lambda P_1\colon \#(U(\{x \in \text{BOYS}^{*} \mid P_1(<x>)\})) < 5)$.

Analogously, syntactic tree (8a) is constructed, with meaning (8b).

(8a) *nps* (*np* (*determiner* (*numeral* (more, than, *number* (six))),
noun (girls)) [*form:plural*])

(8b) $(\lambda Q_1\colon \#(U(\{y \in \text{GIRLS}^{*} \mid Q_1(<y>)\})) > 6)$.

Rule F3 constructs (9a) from (7a) and (8a); it constructs (9b) from (7c) and (8b). (9b) is equivalent to (9c).

(9a) *nps* (*np* (*determiner* (*numeral* (less, than, *number* (five))),
noun (boys)) [*form:plural*],
np (*determiner* (*numeral* (more, than, *number* (six))),
noun (girls)) [*form:plural*])

(9b) $(\lambda R_2\text{: (fun: }(\lambda P_1\text{: }\#(U(\{x \in BOYS^* \mid P_1(<x>)\})) < 5),$
$\text{arg: }(\lambda x_1\text{: (fun: }(\lambda Q_1\text{: }\#(U(y \in GIRLS^* \mid Q_1(<y>)\})) > 6),$
$\text{arg: }(\lambda y_1\text{: }R_2(conc(x_1,y_1)))))))$

(9c) $(\lambda R_2\text{: }\#(U(\{x \in BOYS^* \mid \#(U(\{y \in GIRLS^* \mid R_2(<x,y>)\})) > 6\})) < 5).$

Note that rule F4 would have constructed the same syntactic tree as F3, but a logical formula with the alternative quantifier order.

Next, we find in the lexicon (10a) with meaning (10b).

(10a) *verb*$_2$(date)[*form:plural*]

(10b) DATE

Rule G combines (9ac) with (10ab) into (11ab). (11b) is equivalent to (11c).

(11a) *sentence* (*np* (*determiner* (*numeral* (less, than, *number*(five))),
noun (boys)) [*form:plural*],
verb$_2$ (date) [*form:plural*],
np (*determiner* (*numeral* (more, than, *number* (six))),
noun (girls) [*form:plural*])

(11b) $\text{(fun: }(\lambda R_2\text{: }\#(U(\{x \in BOYS^* \mid \#(U(\{y \in GIRLS^* \mid R_2(<x,y>)\})) > 6\})) < 5),$
arg: DATE)

(11c) $\#(U(\{x \in BOYS^* \mid \#(U(\{y \in GIRLS^* \mid DATE(<x,y>)\})) > 6\})) < 5.$

As another example, we now show a different derivation of sentence (1), which assigns it the "cumulative quantification" reading.

In this case, (12ab) is derived exactly as (3ab) above, and (13ab) is derived analogously.

(12a) *numeral* (less, than, *number* (five)) [*form:plural*]

(12b) $(\lambda X\text{: }\#(U(X)) < 5)$

(13a) *numeral* (more, than, *number* (six)) [*form:plural*]

(13b) $(\lambda Y\text{: }\#(U(Y)) > 6).$

Rule B2 combines (12ab) and (13ab) into (14ab). (14b) is equivalent to (14c).

(14a) *compnum*$_2$ (*numeral* (less, than, *number* (five))[*form:plural*],
numeral (more, than, *number* (six)) [*form:plural*],

(14b) $(\lambda X_2\text{: (fun: }(\lambda X\text{: }\#(U(X)) < 5)\text{, arg: }proj_1(X_2)) \wedge$
$\text{(fun: }(\lambda Y\text{: }\#(U(Y)) > 6)\text{, arg: }proj_2(X_2)))$

(14c) $(\lambda X_2\text{: }\#(U(proj_1(X_2))) < 5 \wedge \#(U(proj_2(X_2))) > 6)$.

In the lexicon we find (15a) with meaning (15b), and (16a) with meaning (16b).

(15a) *noun* (boys) [*form:plural*]

(15b) $BOYS^*$

(16a) *noun* (girls) [*form:plural*]

(16b) $GIRLS^*$

Rule D2 combines (15ab) and (16ab) into (17ab).

(17a) *compnoun*$_2$ (*noun* (boys) [*form:plural*],
noun (girls) [*form:plural*])

(17b) $BOYS^* \times GIRLS^*$

Rule F2.2 combines (14ac) and (17ab) into (18ab).

(18a) *nps* (*np* (*numeral* (less, than, *number* (five)),
noun (boys)) [*form:plural*],
np (*numeral* (more, than, *number* (six)),
noun (girls)) [*form:plural*])

(18b) $(\lambda R_2\text{: (fun: }(\lambda X_2\text{: }\#(U(proj_1(X_2))) < 5 \wedge \#(U(proj_2(X_2))) > 6)\text{,}$
$\text{arg: }\{y_2 \in BOYS^* \times GIRLS^* \mid R_2(y_2)\}))$

Rule G combines the lexical entry for "date", i.e. (10ab), with (18ab). This results in a syntactic tree which is identical to the one specified in (11a). Semantically, the function (18b) is applied to the two-place relation DATE. The result is equivalent to (19).

(19) $\text{(fun: }(\lambda X_2\text{: }\#(U(proj_1(X_2))) < 5 \wedge \#(U(proj_2(X_2))) > 6)\text{,}$
$\text{arg: }\{y_2 \in BOYS^* \times GIRLS^* \mid DATE(y_2)\})$.

FOOTNOTES

1. For instance: WOODS (1967), BARTSCH (1973), HAUSSER (1974), BENNETT (1975), KROCH (1975), CUSHING (1977), VAN LEHN (1978), BARWISE & COOPER (1980).
2. Cf. VENDLER (1962), WOODS (1967), DIK (1972; 1974), BARTSCH (1973), HAUSSER (1974), BENNETT (1975).
3. This possibility may also be achieved in a different way: by using a logical language which includes variables ranging over the union of the set of individuals and the set of sets of individuals. Such languages were defined by SCHA (1976), LANDSBERGEN & SCHA (1979) and BRONNENBERG et al. (1980).
4. Proposals to the contrary at least agree about one important case: the fact that in active sentences the subject noun phrase tends to have wide scope over the object noun phrase (cf. IOUP, 1975; REINHART, 1978).
5. See Section 5.
6. To simplify the syntactic trees, we sometimes leave out the specification of a syntactic attribute value, if it does not play a role in the remainder of the derivation.
7. To simplify the logical formulas, we apply equivalence transformations during the derivation process. Thus, the finally derived formula is not the one actually specified by the grammar rules, but is equivalent to it.

REFERENCES

ALLÉN, S. & J.S. PETŐFI (eds), 1979, *Aspects of automatized text processing,* Papers in text linguistics, vol. 17, Buske, Hamburg.

BARTSCH, R., 1973, 'The semantics and syntax of number and numbers', in: KIMBALL (1973).

BARWISE, J. & R. COOPER, 1980, *Generalized quantifiers and natural language,* Mimeograph.

BENNETT, M.R., 1975, *Some extensions of a Montague fragment of English,* Indiana University Linguistics Club.

BOLC, L. (ed.), 1980, *Natural language question answering systems,* Hanser, München & Wien; MacMillan, London, Basingstoke.

BRONNENBERG, W.J.H.J., BUNT, H.C., LANDSBERGEN, S.P.J., SCHA, R.J.H., SCHOENMAKERS, W.J. & E.P.C. VAN UTTEREN, 1980, 'The question-answering system PHLIQA1', <u>in</u>: BOLC (1980).

CUSHING, S., 1977, *The formal semantics of quantification*, Indiana University Linguistics Club.

DAVIS, S. & M. MITHUN (eds), 1979, *Linguistics, philosophy and Montague grammar*, University of Texas Press, Austin & London.

DIK, S.C., 1972, *Sets in semantic structure*, Publikaties van het Instituut voor Algemene Taalwetenschap, nr 4, University of Amsterdam.

DIK, S.C., 1974, *Universal quantifiers in Dutch*, Publikaties van het Instituur voor Algemene Taalwetenschap, nr 6, University of Amsterdam.

GUENTHNER, F. & S.J. SCHMIDT (eds), 1978, *Formal semantics and pragmatics for natural languages*, Reidel, Dordrecht.

HAUSSER, R., 1974, *Quantification in an extended Montague grammar*, dissertation, University of Texas at Austin.

HINTIKKA, J., MORAVCSIK, J. & P. SUPPES (eds), 1973, *Approaches to natural language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, Reidel, Dordrecht.

IOUP, G., 1975, 'Some universals for quantifier scope', <u>in</u>: KIMBALL (1975).

KIMBALL, J.P. (ed.), 1973, *Syntax and semantics*, Vol. 2, Seminar Press, New York, London.

KIMBALL, J.P. (ed.), 1975, *Syntax and semantics*, Vol. 4, Academic Press, New York, San Francisco & London.

KROCH, A.S., 1975, *The semantics of scope in English*, Indiana University Linguistics Club.

LANDSBERGEN, S.P.J. & R.J.H. SCHA, 1979, 'Formal languages for semantic representation', <u>in</u>: ALLÉN & PETÖFI (1979).

LANGENDOEN, D.T., 1978, 'The logic of reciprocity', *Linguistic Inquiry*, <u>9</u>, 177-197.

MONTAGUE, R., 1970, 'Universal grammar', *Theoria* <u>36</u>, 373-398. Reprinted in Montague (1974).

MONTAGUE, R., 1973, 'The proper treatment of quantification in ordinary English', in: HINTIKKA et al. (1974), reprinted in MONTAGUE (1974).

MONTAGUE, R., 1974, *Formal Philosophy,* Yale Univ. Press, New Haven & London.

PARTEE, B., 1973, 'Some transformational extensions of Montague grammar', *Journal of Philosophical Logic* 2, reprinted in PARTEE (1976).

PARTEE, B. (ed.), 1976. *Montague grammar,* Academic Press, New York, San Francisco & London.

REINHART, T., 1978, 'Syntactic domains for semantic rules', in: GUENTHNER & SCHMIDT (1978).

SCHA, R.J.H., 1976, 'Semantic types in PHLIQA1', *COLING 76 Preprints,* Ottawa.

SCHA, R.J.H., 1978, 'A formal treatment of some aspects of quantification in English' (Abstract), *Preprints of the 7th International Conference on Computational Linguistics,* Bergen, Norway.

THOMASON, R., 1972, 'A semantic theory of sortal incorrectness', *Journal of Philosophical Logic* 1, 209-258.

VAN LEHN, K.A., 1978, *Determining the scope of English quantifiers,* MIT AI Lab., AI-TR-483.

VENDLER, Z., 1962, 'Each and every, any and all', *Mind* 71, reprinted in VENDLER (1967).

VENDLER, Z., 1967, *Linguistics in philosophy,* Ithaca, NY.

WALDO, J., 1979, 'A PTQ semantics for sortal incorrectness', in: DAVIS & MITHUN (1979).

WOODS, W.A., 1967, *Semantics for a question-answering-system,* Ph.D. Thesis, Harvard University.

THE SEMANTICS OF TOPIC-FOCUS ARTICULATION

by

Anna Szabolcsi

ABSTRACT

This paper proposes a purely grammatical approach to topic-focus articulation. Section 1 indicates why this approach seems preferable to communicative ones. Section 2 gives an outline of the syntax of word order and intonation in Hungarian and suggests how these phenomena can be treated within Montague Grammar. Section 3 is concerned with term phrases in Focus position. Section 4 examines scope phenomena related to Focus. Section 5 adds a few informal comments about the behaviour of the finite verb.

1. METHODOLOGICAL SUGGESTIONS

It is commonly agreed that for communication to be effective, one's intended message needs to be adapted to the situation in which it is to be conveyed. There are two large areas of linguistic research particularly involved in the implementation of this idea, namely, the theories of speech acts and the theories of communicative articulation, the latter being concerned with the more specific claim that our sentences fulfill both an 'anchoring' and a 'furthering' function in the flow of communication.

I am using the ad hoc cover terms 'anchoring function' and 'furthering function' for the following reason: there exist a vast number of notions (topic-comment, theme-rheme, background-focus etc.), which seem to stem from some common intuitive basis but whose actual contents tend to vary almost from author to author although each appears to be useful in explaining some interesting facts of syntax, or semantics, or pragmatics. This proliferation of notions is both promising and frightening. It is promising because it seems to indicate the heuristic value of recognizing some articulation of sentences beyond that of, say, subject and predicate; and it is frightening because one might expect that if the essence of the phenomenon were captured, we would not need this many of them.

Without wishing to give a critical review of the field, let me briefly point out why it is so difficult to compare rival claims about this kind of sentential articulation. We seem to have the following (overlapping) possibilities for divergence:

(Ai) some authors set up their definitions of the two functions and from those definitions they try to deduce which segments of sentences will fulfill them;

(Aii) some others pick out some grammatical phenomenon and identify its properties, so to say by definition, with the properties of one of those functions;

(B) since those functions (or, associated notions) tend to come in pairs, the one an author gives conceptual prominence will also create a complement with some rather uninspiring definition;

(C) authors vary as to whether they attribute only a pragmatic or also a semantic significance to their notions.

These treatments seem to rest on the (sometimes tacit) postulation of a rather *direct* correlation between the alleged communicative principles and certain grammatical processes. Given however that the existence of such a direct correlation is dubious and the notions in current use are fairly vague (so vague in fact that even authors giving radically different analyses for the same sentences might well agree in them), a conceivable way out from this diversity would be to forget about the adaptation idea for a while and look for some independent grammatical evidence.

It may sound absurd to seek independent grammatical evidence for something I refuse to clearly identify in advance but it is actually not that absurd. There is the indisputable fact that the 'same' sentence may have various word order and/or intonation variants, with possibly different use conditions. Now, by virtue of their mere existence, any grammar can be expected to generate those variants and given that not all permutations are grammatical, by first restricting our attention to this side of the matter we can arrive at a set of distinctions (rules, categories etc.) with self-contained formal motivation. The next step is to interpret sentences in observance of those distinctions - that is, to proceed in the very same fashion as we do in the case of any other grammatical phenomenon. It may turn out that certain 'word order rules' affect the truth conditions of the sentences (i.e. that some of the differences in use conditions are simply due to differences in meanings proper) while the operation of other becomes relevant only at a text, or discourse, level indeed.

Apart from being justified on its own, this grammar-minded approach seems useful for the following reason as well. Out intuitive concept and mutual understanding of the anchoring versus furthering division is presumably based on the above mentioned distinctions, whose unbiased and thorough examination has been hindered, however, by giving a too ready rationale to their existence. The examination I propose is very likely to verify many of the usual claims about this kind of sentential articulation but will hopefully also give more substance to them and will save us from premature generalizations.

2. SYNTACTIC MOTIVATION AND OUTSET

2.1. In this paper I will attempt to account for the semantic significance of some Hungarian word order rules. Besides being my mother tongue, Hungarian has the following advantages: it is a 'free word order' language (but, as opposed to Russian and Czech, for instance, the definite-indefinite distinction is marked with articles and not with linear position) and a syntactic treatment of its word order has been proposed by É. KISS (1980). Since this work provides the fundamentals of those *self-contained formal distinctions* that I required in the previous section, I begin with a semiformal summary of it.

E. Kiss proposes the following base rules:

(I) a. $S'' \rightarrow X^{n*}\ S'$
b. $S' \rightarrow X^{n}\ S^{o}$
c. $S^{o} \rightarrow V\ \ X^{n*}$.

The set of maximal major categories X^{n*} immediately dominated by S" and the single maximal major category X^{n} immediately dominated by S' are called T and F respectively. These mnemotechnic names are reminiscent of 'topic' and 'focus' but the introduction of the corresponding positions is motivated on purely formal grounds. Their nodes are generated empty and can be filled by the optional transformation *move* α from S^{o}, leaving a trace behind.[1]

Now, empirically speaking, what is the motivation for the introduction of the T and F positions?

(II) (i) The MAIN STRESS of the sentence falls on the first major category in F or on the finite verb in case F is left empty;

(ii) The X^{n}s within T and the X^{n}s behind the verb (i.e. sister nodes) are interchangeable preserving gramaticality. E.g. taking a sentence with a finite verb *v* and any two 'mobile constituents' *a* and *b* with no specific restrictions:

V a b	* v A b	* v a B
V b a	* v B a	* v b A
A v b	a V b	* a v B
B v a	b V a	* b v A
* A b v	a B v	a b V
* B a v	b A v	b a V

(iii) In certain cases F must be filled in a specific way:
- an X^n modified by a negative (interrogative, optative etc.) operator must occupy the F position:

(1) a. [$_T$e] [$_F$ Nem PÉTER] sétál Máriával
not Peter walks Mary-with
b.* [$_T$ Nem Péter] [$_F$ MÁRIÁVAL] sétál
c.* [$_T$ Máriával] [$_F$e] SÉTÁL nem Péter

- in the presence of a so-called reduced complement (e.g. the *converb*) the F position may only be left empty if the verb is modified by an operator of the above mentioned kind:

(2) a. [$_T$ Péter] [$_F$ *BE*] szaladt
Peter in ran
b. [$_T$e] [$_F$ PÉTER] szaladt *be*
c. [$_T$ Péter] [$_F$e] nem SZALADT *be*
Peter not ran in
d.* [$_T$ *Be* Péter] [$_F$e] SZALADT
e.* [$_T$ Péter] [$_F$e] SZALADT *be*

(iv) If the X^n in F position consists of a head plus an embedded sentence, the embedded sentence must be moved to the end or the beginning of the matrix sentence (the same is possible, but not obligatory, in other positions):

(3) a.* [$_{S_1}$ [$_F$ AZT [$_{S_2}$ hogy [$_F$ PÉTER] győzött]] hallottam]
it+acc that Peter won-he heard-I
b. [$_{S_2}$ Hogy [$_F$ PÉTER] győzött] [$_{S_1}$[$_F$ AZT] hallottam]
c. [$_{S_1}$ [$_F$ AZT] hallottam] [$_{S_2}$ hogy [$_F$ PÉTER] győzött]

(v) Through certain bridge verbs the F of the *that*-clause may be raised to the F position of the matrix sentence (and similarly for Ts):

(4)=(3) $[_{S_1}$ $[_F$ PÉTERT$_i$] hallottam $[_{S_2}$ hogy $[_F$ t_i] GYŐZÖTT]]
Peter+acc heard-I that won-he

Although this proposal does not give a full account of all subtleties of Hungarian word order and intonation, it is both comprehensive and reliable enough to motivate that whatever interpretation and raison d'être should be attributed to these variants, the positions T and F must be distinguishd in any grammar of the language for purely formal reasons. Moreover, I believe that Hungarian speakers' intuition about 'communicative articulation' must be based on these very distinctions and therefore any reasonable interpretive notion must be definable in these terms or must be possible to give a similarly strong formal motivation. Note by the way that É. Kiss's findings do *not* support the assumption that Hungarian sentences are best characterized by an inherently bipartite structure as neither T nor F has a syntactically significant 'complement'.

2.2. I have already noted that É. Kiss's rules are set up without keeping an eye on interpretation (which I do not regard as a virtue in general but in this particular case it has its advantages). Nevertheless, the version of Extended Standard Theory she uses does not even have a sophisticated interpretive component as yet. For this reason I will take her claims as empirical facts and try to formulate my results in terms of *Montague Grammar*.

Note first that if we want to produce all these variants we can no longer expect function and argument expressions to combine in a uniform surface order. A first approximation may be to assume that the relevant rules of functional application have *three versions*: one for inserting the 'nominal expression' in front of the finite verb (=F position), a second for inserting it 'somewhere' to the left of the verb but not to the right of F (=T position), and a third for inserting it 'somewhere' to the right of the verb (=neutral position).

Provided that the truth/use conditions of the sentence are indeed dependent on which versions of the rules are applied in its derivation, we can be prepared for the following three main possibilities:

(a) It may turn out that whatever fills the F (or T, or neutral) position, the sentence will gain the same kind of interpretational surplus compared to what is predicted in PTQ.

(b) It may turn out that the effects of the same kind of rule version vary with the nature of the inputs.

If (a) or (b) obtains we will have to provide each rule version with a specific translation, in addition to what it carries over from PTQ.

(c) It may turn out that no interpretational surplus arises in fact and only the distribution of PTQ-predicted readings is constrained by the way we filled those positions.

I will argue that - at least as far as Hungarian is concerned - it is a combination of (b) and (c) that obtains. That is, 'word order rules' may add to the literal meaning of the sentence, although not in a uniform fashion and, further, interpretation options are sometimes also constrained (e.g. in connection with quantifier scopes). It is in fact not very surprising that (a) does not obtain; among others, this may be a reason why many attempts to treat this kind of sentential articulation on the assumption of a direct correlation between communication and grammar turn out to be inconsistent or impossible to check against new examples.

I am far from claiming that I can give an exhaustive treatment of the issue. Here my attention will be restricted primarily to *the behaviour of term phrases in F position and related problems, supplemented with a few remarks on the verb and the converb*. Nevertheless, I hope that even within these limitations I can motivate the claims I made in the paragraph above and that my considerations will illustrate the advantages of a syntax-based approach.

Note that my approach also implies that the results may be more or less language specific, i.e. the significant syntactic distinctions and their respective interpretations may vary from language to language. Apart from the theoretical consequences of this fact, let me warn the reader of the practical consequence that the English 'equivalents' I can give for my examples may happen to be only near-equivalents.

In connection with interpretation, I will refer to the constituent in F position as Focus and to a constituent in T position as Topic.

3. TERMS IN FOCUS

3.0. In rather informal terms we can say that the common feature distinguishing T and F from neutral constituents is that only the former may be *contrastive*. This statement of course needs to be made more precise in various respects. First, although both are put under the same label, Topic-contrast and Focus-contrast are two different matters, in force as well as in content. The characteristic difference is that by using a sentence with a contrastive Topic, one suggests (or, implicates[2]) that the claim he is making need not be true of something else, whereas by using a contrastive Focus one asserts that the claim he is making is in fact not true of anything else[3]. An additional difference is that whether the Topic of a sentence is contrastive or not usually depends on whether it receives an extra intonation contour while most Foci (i.e. most expressions in F position) are necessarily contrastive in the above sense. On the other hand, the assumption that this kind of general interpretational surplus may only be attributed to T and F is corroborated by the fact that maximally elliptical (one-X^n) sentences must follow either the T-pattern or the F-pattern and can only be conjoined with non-elliptical sentences if those have the same kind of X^ns in their respective positions.

3.1. Exhaustive listing

It will have become clear that I regard exhaustive listing as the predominant semantic characteristic of Focus, as opposed, for instance, to those who argue that a sentence like (6)

(6) $[_F$ Mária] látta Pétert 'MARY saw Peter'
Mary saw Peter+acc

presupposes (in one of the many senses of this term) that the set of those who saw Peter is not empty and asserts that Mary is contained in that set. Needless to say, the postulation of such a presupposition is in itself not incompatible with exhaustive listing and therefore I will return to it in 3.3. As for the other parts of the two claims, the choice between them may seem like a matter of simple intuition as long as we only consider one individual denoting expressions in F. Note however that while the proposal I am arguing against predicts that from (7) we can infer (6), this is not the case: from (7) we are only entitled to infer (8):

(7) $[_F$ Mária és Éva] látta Pétert 'MARY AND EVE saw Peter'[4]

(8) $[_F$e] Látta Pétert Mária 'Mary saw Peter'

This suggests that co-ordinate NPs in F position may not be derived via conjunction reduction (whether it be a syntactic or a logical application of the idea). For illustrative purposes we might say that Focus has something like an invariably collective reading but, of course, in view of (8) being a logical consequence of (7), this may only be metaphorical. The same situation obtains with plural quantifiers: (10) is not a logical consequence of (9):

(9) [$_F$ Három lány] látta Pétert 'THREE GIRLS saw Peter'

(10) [$_F$ Két lány] látta Pétert 'TWO GIRLS saw Peter'

It might be argued that the reason why this letter inference is unjustified is that natural numbers are to be interpreted as numerically definite quantifiers. Apart from missing a generalization, this would not be a good argument, however, since on the one hand, these quantifiers get a probably numerically definite interpretation in F position only and, on the other hand, in other positions we cannot even get on without the 'at least' meaning. E.g., on one reading (11) undoubtedly means that at most two girls may have seen Peter, which would be impossible if *three* meant exactly three:

(11) [$_T$ Háron lány] nem látta Pétert 'Three girls, didn't see Peter'

And finally, for those who may not trust the juggling with inferences in such communicatively delicate cases: without exhaustive listing we cannot explain why the biconditional is normally expressed in Hungarian by mere Focusing (for the syntax of (12), see (II.iv) above):

(12) [$_F$ Akkor] megyek veled ha cilindert veszel
then go-I you-with if tophat+acc take-you
'I'll go with you only if you put on a tophat'

These observations seem to motivate the taking of exhaustive listing to be a property of the F position that must be directly reflected in truth conditions.

3.2. A first extension of PTQ

Although exhaustive listing appears to be a logically very unsophisticated notion, the appropriate formulation of the translation rule corresponding to F-filling turns out to be rather complicated, due to the fact that logic lacks the comfortable *and none/nothing else* idiom that can be suffixed to just everything. To make discussion simpler, I begin by sketching a few tentative extensions of the PTQ grammar for Hungarian.

I will retain English lexical items for derivations to be easier to decipher. Also, I will state syntactic rules almost as loosely as I did in *2.2*; to develop the full marking technique would be a routine job but its explication would make the rules overcomplicated here. I will ignore problems of pronominalization throughout the rules.)

B_T = {*Mary*, *Peter*,...,he_1,he_2,...}
where he_{2k} translates as $\lambda P^{\vee}P(x_{2k})$,
he_{2k+1} translates as $^{\vee}\mathcal{P}_{2k+1}$.
B_{TV} = {[]*find*, []*lose*, []*seek*,...}
B_{IV} = {[]*run*, []*sing*,...}

S4a. (Focus) If $\alpha \in P_T$ and has the form he_n and $\delta \in P_{IV}$ and its main verb is prefixed with [, then $F_{4a}(\alpha,\delta) \in P_t$ and is obtained by replacing [in δ with /α

S4b. (Topic) If $\alpha \in P_T$ and does not have the form he_{2k+1} and $\delta \in P_{IV}$ then $F_{4b}(\alpha,\delta) \in P_t$ and is obtained by inserting α somewhere to the left of [or / in δ.

S4c. (Neutr) If $\alpha \in P_T$ and does not have the form he_{2k+1} and $\delta \in P_{IV}$ then $F_{4c}(\alpha,\delta) \in P_t$ and is obtained by inserting α somewhere to the right of the main verb in δ.

T4. If $\alpha \in P_T$ and $\delta \in P_{IV}$ and α,δ translate as α', δ' respectively, then $F_{4a}(\alpha,\delta)$, $F_{4b}(\alpha,\delta)$, $F_{4c}(\alpha,\delta)$ translate as $\alpha'(^{\wedge}\delta')$.

And similarly for S5.

S14'. The same as the PTQ quantification rule, with the difference that only he_{2k} pronouns may be replaced and if $\alpha = he_{2k+1}$ it may only replace a focused pronoun.

S17'. The same as in PTQ, with the difference that its operations also have three versions and, in particular, in $F_{11}(\alpha,\delta)$ *not* replaces] in δ.

S21. If $\phi \in P_t$ and ϕ contains an occurrence of he_{2k+1} or him_{2k+1} then $F_{21}(\phi) \in P_{t/T}$ and the pronoun in ϕ is 'capitalized' (= marked to receive sentence stress).

T21. If $\phi \in P_t$ and ϕ translates as ϕ', then $F_{21}(\phi)$ translates as $\lambda\mathcal{P}_{2k+1}\phi'$.

S22. If $\alpha \in P_T$ but $\alpha \neq he_n$ and $\beta \in P_{t/T}$ then $F_{22}(\alpha,\beta) \in P_t$ and is obtained by replacing HE_{2k+1} or HIM_{2k+1} in β by α.[5]

T22. If $\alpha \in P_T$ and $\beta \in P_{t/T}$ and α,β translate as α',β' respectively, then $F_{22}(\alpha,\beta)$ translates as $\beta' = \lambda\mathcal{P}\forall R[\alpha'(^{\wedge}\lambda x[\beta'(^{\wedge}\lambda P^{\vee}P(x)) \wedge {}^{\vee}R(x)]) \rightarrow {}^{\vee}\mathcal{P}(R)]$.
For instance

(6) [$_F$ Mária] látja Pétert

/MARY]sees Peter+acc, *22*
Mary   /HE_1]sees Peter+acc, *21*
/he_1]sees Peter+acc, *4a*
he_1 [ ]see Peter+acc, *5c*
[ ]see   Peter

(6') $\lambda\mathcal{P}_1[{}^\vee\mathcal{P}_1({}^\wedge see({}^\wedge\lambda P^\vee P({}^\wedge p)))] = \lambda\mathcal{P}\forall R[\lambda P^\vee P({}^\wedge m)({}^\wedge\lambda x[\lambda\mathcal{P}_1[{}^\vee\mathcal{P}_1({}^\wedge see ({}^\wedge\lambda P^\vee P({}^\wedge p)))]({}^\wedge\lambda P^\vee P(x)) \wedge {}^\vee R(x)]) \rightarrow {}^\vee\mathcal{P}(R)]$

Now, considering that this monstrous formula says the same as (6"):

(6") $\forall v[see_*(p)(v) \leftrightarrow v = m]$

one may ask why it is needed. Notice that the intention behind the whole procedure is to retain the original translations of the term phrases (since it would be somewhat strange to claim that *Mary*, for instance, means something different in one position than in another). Consequently, T22. must be applicable to term phrases with all kinds of internal structures. What happens now if the term to be focused is somewhat more complicated than a proper name? First, consider a conjunction like *Mary and Eve*. Assuming that it translates as $\lambda P[{}^\vee P({}^\wedge m) \wedge {}^\vee P({}^\wedge e)]$ and realizing that (7) above means that someone saw Peter iff he/she is identical to either Mary or Eve, a uniform translation rule to this effect must require that someone saw Peter iff he/she is contained in every set that contains the person(s) listed in F. Nevertheless, the situation is even worse with, say, *two girls* ($\lambda P\exists x\exists y[\neg(x=y) \wedge girl(x) \wedge girl(y) \wedge {}^\vee P(x) \wedge {}^\vee P(y)]$) in F position, since it would be far too much to require that for someone to have seen Peter, he/she must be contained in every set that contains at least two girls - we are only interested in those sets that contain at least two girls who saw Peter. (And that is also sufficient, for (a) if there are indeed at least two girls, say, Mary and Eve, who saw Peter, then by letting R be ${}^\wedge\lambda x[{}^\vee x=m \vee {}^\wedge x=e]$ we exclude the possibility that any third person, too saw Peter, and (b) if there are not at least two girls who saw Peter - that is, when the value of the function on the right hand side will be 1 for every $\mathcal{P}$, then by letting $\mathcal{P}_1$ be, say, ${}^\wedge\lambda P[\neg(a=a)]$ we have at least one argument for which the value of the left hand side function is 0 and therefore, just as we expect, the sentence will be false.)

This much does not explain everything about T22. or the other rules I introduced, it only serves to show that at least in intuitively simple cases this translation does give the correct results. The remaining problems will be discussed step by step.

3.3. Existential presupposition

Let us now turn to the problem whether there is an existential presupposition associated with Focus. This will also bring us to the treatment of so-called constituent negation and *wh*-questions.

It goes without saying that by claiming that a sentence like (13) presupposes that Peter is beating someone,

(13) $[_F$ Máriát] veri Péter 'Peter is beating MARY'
Mary+acc beats Peter

we are committed to the view that there is something wrong with the negation of (13) if Peter is not beating anyone. (14) and (15) are synonymous in Hungarian:

(14) $[_F$ Nem Máriát] veri Péter 'Peter is not beating MARY'
not Mary+acc beats Peter

(15) Nem igaz az, hogy $[_F$ Máriát] veri Péter
'It is not the case that Peter is beating MARY'

Nevertheless - and in this respect Hungarian may be different from English - it seems that neither semantic nor pragmatic conventions are violated if (14) is continued in either of the following ways:

(14) a. ... hanem $[_F$ Évát].
but Eve+acc

b. ... hanem $[_F$ a gyerekkel] játszik.
but the kid-with plays

c. ... hanem $[_F$ az ajtó] csapódott be.
but the door banged.

On the contrary, such conjunctions sound perfectly natural and are not infrequent to occur. In view of these facts it would seem unwarranted to assume that (13) and (14) presuppose that Peter is beating someone. (The correspondence between such sentences and *wh*-questions (see 3.5) will support this conclusion, too.)

Shall we say, however, that (14) is *ambiguous* with respect to the scope of negation? That is, (14) might be said to assert either (a) that Peter is beating someone but not Mary, or (b) that Peter is involved in some activity but not in beating Mary, or (c) that something happened but not that Peter is beating Mary.

Note first the difficulties arising in connection with such a proposal. On the one hand, the variation of the scope of negation in the above fashion would not be too easy to build into our grammar, in particular if we consider that subjects, too may fill the F position. On the other hand, notice that in the paragraph above I neglected exhaustive listing on the whole.

Taking that into account as well, we ought to vary, not only the scope of negation but also the 'scope of Focus'. Nevertheless, aside from formal difficulties again, that would be equivalent to the abandoning of all the significant syntactic generalizations the whole approach is based on.

I suggest that (14) is in fact *unambiguous* and being the negation of (13), it simply asserts that it is not the case that Peter is beating someone if and only if that person is Mary. From a logical (semantic) point of view this is just a very unspecific claim and thus it is compatible with all the continuations required. From a communicative point of view, this unspecificity may be regarded as *vagueness* (cf. KEMPSON (1975)). In view of the Gricean principles of conversation, we can predict that in case the speaker of (14) is relevant and does not add anything to this vague negation, he probably means that Peter is beating someone else than Mary. This is, however, only a special case of the working of those principles and is to be accounted for by a unified theory of language and its use, quite independently of the fact that it arises in connection with 'word order rules'.

In virtue of these considerations I propose to add the following rule to those in 3.2:

S22neg. If $\alpha \in P_T$ but $\alpha \neq he_n$ and $\beta \in P_{t/T}$ then $F_{22neg}(\alpha,\beta) \in P_t$ and is obtained by replacing HE_{2k+1} or HIM_{2k+1} in β by *not* α.

T22neg. If $\alpha \in P_T$ and $\beta \in P_{t/T}$ and α,β translate as α',β' respectively, then $F_{22neg}(\alpha,\beta)$ translates as
$\neg(\beta' = \lambda\mathcal{P}\forall R[\alpha'(^\wedge\lambda x[\beta'(^\wedge\lambda P\,^\vee P(x)) \wedge {}^\vee R(x)]) \rightarrow {}^\vee\mathcal{P}(R)])$.

F_{22neg} bears the very same relation to F_{22} as Montague's F_{11} to F_4. This is in keeping with the intuition that a Hungarian sentence can be 'negative' in just two cases: (i) if F_{11} applied in its derivation and it has no Focus, or (ii) if F_{22neg} applied, regardless of whether F_{11} applied or not. These observations can be utilized in the treatment of yes-no questions and scope restrictions.

Notice that the above formulation also makes it unnecessary to introduce some quasi-filter (cf. É. KISS (1980) and (II.iii) above) for ensuring that the negated X^n occupies the F position since *not* and the term in question are not regarded as forming a single mobile constituent.[6] (Incidentally, a closer examination of syntactic data also shows that the same treatment would be necessary in the case of the other 'focusing operators' as well. Its demonstration would go beyond the scope of the present paper, however.)

3.4. Qualitative contrast in F

Before turning to *wh*-questions, let us consider another interesting aspect of Focus-negation. I will use English examples since they seem to work as well as the corresponding Hungarian ones.

Compare the following sentences:

(16) a. My friend, I invited, the minister, I didn't.

b. I invited MY FRIEND and not THE MINISTER.

c. I invited my friend. I didn't invite the minister.

It is clear that for (16)a. and (16)c. to be true, my friend and the minister must be two different persons. This is not so with (16)b. however: it has a reading on which the two descriptions may well apply to the same person and therefore the F position provides for qualitative contrast. To make it more explicit: the sentence *I didn't invite THE MINISTER* does not necessarily license the inference that I did not invite the person who happens to be the minister; rather, it says that the person invited must be intensionally different from the minister.

In order to avoid the temptation to attribute this peculiarity of (16)b. to the highly sophisticated social nature of inviting, let us consider the following examples as well:

(17) I married a NICE GIRL, not a RICH GIRL

(18) I am living in AN ANCIENT MANOR and not in A RAMSHACKLE COTTAGE

(19) This game was not won by PETER, it was lost by MARY

In (17), the girl may be rich, too and in that case the sentence suggests that I did not marry her for her fortune. (18) can be a fine expression of snobbery. The intention behind (19) may be spelled out by pointing out that, say, the winner was necessarily identical to whoever played against Mary.

(19_1) This game was not won by PETER

(19_1') $\neg(\lambda P_1[{}^{\vee}P_1({}^{\wedge}\text{won this game})] = \lambda P[P = {}^{\wedge}\lambda p\,{}^{\vee}P({}^{\wedge}p)])$

Where should we get from, however, the rules for producing these meanings? Notice that T22. has a serious limitation: due to the quantification on the right hand side, it cannot be sensibly applied when HIM_{2k+1} is an intensional object. Inelegant as it is, it seems that a separate rule is needed for such cases. Let us assume that F_{5a} marks its second argument with, say, $+i$ if it is not of the form he_{2k} and the verb is intensional. Then:

S23. If $\alpha \in P_T$ and $\alpha \neq he_n$ and $\beta \in P_{t/T}$ and contains an occurrence of

HIM^{+i}_{2k+1} then $F_{23}(\alpha,\beta) \in P_t$ and is obtained by replacing HIM^{+i}_{2k+1} in β by α.

T23. If $\alpha \in P_T$ and $\beta \in P_{t/T}$ and α,β translate as α',β' respectively then $F_{23}(\alpha,\beta)$ translates as $\beta' = \lambda P[P = {}^{\wedge}\alpha']$.

This gives us a chance to formulate the negative version of this rule without the *+i* restriction, i.e. so that it may apply to cases like (16)-(19) as well:

S23neg. If $\alpha \in P_T$ and $\alpha \neq he_n$ and $\beta \in P_{t/T}$ then $F_{23neg}(\alpha,\beta) \in P_t$ and is obtained by replacing HE_{2k+1} or HIM_{2k+1} in β by *not* α.

T23neg. If $\alpha \in P_T$ and $\beta \in P_{t/T}$ and α,β translate as α',β' respectively, then F_{23neg} translates as $\neg(\beta' = \lambda P[P = {}^{\wedge}\alpha'])$.

3.5. Wh-questions

For determining the focus of a sentence, many authors in topic-comment literature use the *wh*-question test, e.g., claim that (21), but not (22) being an appropriate answer to (20), *John* is the focus of (21) but not of (22):

(20) Who kissed Mary?

(21) JOHN kissed Mary.

(22) John kissed MARY.

Now, apart from the problem whether focus is to be determined in that way or not (I believe it will have become clear that in my opinion, not), how does the present proposal account for the intuitively correct results of this test?

HAUSSER (1978) put forth a very convincing proposal for the treatment of question-answer pairs in Montague Grammar. In his formulation, an interrogative denotes a function from points of reference into sets of corresponding non-redundant answer constituent denotations. E.g.:

(20) Who kissed Mary? (23) John.

(20') $\lambda P_1[{}^{\vee}P_1({}^{\wedge}kiss({}^{\wedge}\lambda P{}^{\vee}P({}^{\wedge}m))]$ (23') $\Gamma({}^{\wedge}\lambda P{}^{\vee}P({}^{\wedge}j))$

where Γ is a context-variable ranging over translations of t/T interrogatives provided by the context and thus by lambda-conversion, *John.* as an answer to (20) will be equivalent to $kiss_*(m)(j)$.

Notice however that (23) is in fact equivalent to (21), that is, such an answer may only be meant to assert that [among the people relevant to be considered] John and only John kissed Mary - otherwise one should have said something like *John, did..* This should be recognized in the translation of the question-answer pair.

Fortunately enough, (20) is only grammatical in Hungarian if the question word *(ki)* occupies the F position:

(20H) a. $[_F$ Ki$]$ csókolta meg Máriát?
b. $[_T$ Máriát$]$ $[_F$ ki$]$ csókolta meg?

and given that the rules for Focus were possible to formulate in such a way that (20') happens to be a subexpression in the translation of (21H), it is easy to imagine that the intended correspondence must be possible to capture in this framework. The only open question is whether the interrogative or the elliptic answer should be made responsible for exhaustive listing.

I suggest that it is the translation of *John.* that should bring exhaustive listing into the picture. That is, I would retain Hausser's treatment of the interrogative and propose to translate (23) in analogy to T22. for Focus:

(23") $\quad \Gamma = \lambda\mathcal{P}\forall R[{}^\wedge\lambda P\,{}^\vee P({}^\wedge j)({}^\wedge\lambda x[\Gamma({}^\wedge\lambda P\,{}^\vee P(x)) \wedge {}^\vee R(x)]) \rightarrow {}^\vee\mathcal{P}(R)]$

This choice can be given the following motivation:

(i) As I mentioned above, (20) in itself does not compel the hearer to give an exhaustive answer - he may use a more redundant answer in which *John* occupies the T position.

(ii) As I will point out in 4.1, there are terms that may not occupy the F position (e.g. *mindenki* 'everybody') and thus may not be associated with exhaustive listing but can nevertheless be given as non-redundant answers. For those I want to retain the $\Gamma({}^\wedge\alpha')$ type translation originally offered by Hausser. (And similarly for any kind of expression which may fill F but turns out not to express exhaustive listing.)

(iii) Note that elliptical sentences like *John.* (or, *Not John.*) do not only function as non-redundant answers but can also be conjoined with appropriate non-elliptical sentences, e.g.:

(24)	$[_T$ Máriát$]$	$[_F$ Péter$]$	csókolta meg	és	nem	$[_F$ János$]$
	Mary+acc	Peter	kissed	and	not	John

I assume that the derivation and translation of the second conjunct in (24) must be in all the relevant respects similar to that of (23) - with the difference that a co-text variable Δ might be used and the syntactic rules would be somewhat more complicated - which is again a reason for attributing the property of exhaustive listing to the answer (i.e. the elliptic sentence consisting of an F) rather than to the interrogative.

4. SCOPE PHENOMENA

Given that scope phenomena are often claimed to be dependent on linear order and/or dominance relations, it is natural to ask if the operation of 'word order rules' has any particular constraining effect on interpretation options in Hungarian. It seems it has indeed, as I will point out below. Nevertheless, heretic as it may sound, I will not try to draw final conclusions and will restrict my attention to a few quantifiers only. Apart from my work being far from a final stage, I have the following reason for making this reservation. Although one is often tempted to be sure that word order or intonation disambiguates sentences in a particular way, there has hardly ever been any everyday speaker to conform to one's expectations consistently. Whether deviations should simply refute the claims or are due to dialectal variations or performance factors is very difficult to decide. On the other hand, if the plain *ungrammaticality* of certain sentences can be traced back to the joint effects of independently stipulated restrictions, this may indicate that those restrictions are justified on their own. I will therefore try to avoid making claims about cases in which I do not (as yet) have this kind of justification.

I will argue that the most perspicuous scope restrictions *are associated with F* and are of two kinds:

(i) restrictions on the quantifier in F position,

(ii) restrictions on the scopes of non-F quantifiers with respect to Focus (i.e. the binding of he_{2k+1}).

By (i) and (ii) I also mean to suggest that there seem to be no specific restrictions on non-F quantifiers and on their relative scopes unless they follow from (ii) or from the inherent restrictions of those quantifiers with respect to, say, negation. (The restrictions I will point out might be formalized in a fasion similar to HAUSSER (1976)).

4.1. First, it appears that the quantifier in F position must have *wider scope than preverbal* (i.e. F_{11}) *negation*. For instance, it may not be a logical consequence of (25) that there are not any two boys who are running (unless, in view of exhaustive listing, we also add that there are not at least four boys in the universe to be considered). Derivations like (25)b. are therefore to be excluded:

(25) $[_F$ Két fiú] nem fut 'TWO BOYS aren't running'
two boy not run

(25') $$\lambda P_1[{}^{\vee}P_1({}^{\wedge}\lambda x_2[\neg run(x_2)])] = \lambda P\forall R[\exists x\exists y[\neg(x=y) \wedge boy(x) \wedge boy(y) \wedge \neg run(x) \wedge \neg run(y) \wedge {}^{\vee}R(x) \wedge {}^{\vee}R(y)] \rightarrow P(R)]$$

(25)a.
TWO BOYS not run, *22*
two boys — HE_1 nor run, *21*
he_1 not run, *14*
he_1 — he_2 not run, *11a*

(25)b.*
TWO BOYS not run, *22*
two boys — HE_1 not run, *21*
he_1 not run, *11a*
he_1 — run

Fortunately enough, the validity of this claim can be tested, not only againsts a set of random examples but also against a rather crucial one. Namely, HAUSSER (1976) observes that presupposing quantifiers have scope restrictions with respect to negation and, in particular, *every* α is bound to have more narrow scope than negation. In Hausser's notation:

every^{-W} translates as $\lambda Q\lambda P\forall x \in[{}^{\vee}Q(x)]\ {}^{\vee}P(x)$

Given that its Hungarian equaivalent *minden* α is subject to the same restriction, its behaviour, that is, whether or not it may occur in F position, may quite straightforwardly qualify my claim.

Now, in case we had to rely on vague intuition or loose syntactic observations in determining the 'communicative articulation' of Hungarian sentences, we would probably expect that *minden* will refute the claim. After all, it may occur in front of finite verbs, it may receive a fairly high pitch and why should it be excluded from the role of, say, being the most important piece of new information? The reliance on rigorous syntactic criteria will turn out to be useful here, however. Remember that Hungarian verbs may have converbs (which are mobile but form a single lexical item together with the verb, e.g. *be-rúg* 'get drunk' lit. 'in kick') and É. Kiss's rules imply that in case F is filled, the converb may not be prefixed to the finite verb (cf. (II.iii) above). In view of these facts the ungrammaticality of (26) - as opposed to (27) - indicates that *minden* α may not fill the F position,

(26)* $[_F$ Minden fiú] rúgott be 'EVERY BOY got drunk'

(27) $[_F$ Két fiú] rúgott be 'TWO BOYS got drunk'

which in turn evidences that F is inherently related to having wider scope than preverbal negation and that is why a quantifier with the opposite restriction may not fill it, even in affirmative sentences. (This restriction might be formalized by assigning a +*w* feature to he_{2k+1} in the lexicon.)

4.2. Let us now turn to the *interaction* of Focus and other quantifiers. Given that Focus may express exhaustive listing, it is easy to see that it makes a big difference whether a non-F quantifier is introduced before or after the binding of he_{2k+1}. For instance, if my intuition about English is correct, (28) says that for everybody individually, he loves Mary and only Mary, whereas (29) makes the weaker claim that Mary is the only person unanimously loved:

(28) Everybody loves MARY

(28') $\forall u_2 \in [human_*(u_2)]\ \forall v[love_*(v)(u_2) \leftrightarrow v = m]$

(29) MARY is loved by everybody

(29') $\forall v[\forall u_2 \in [human_*(u_2)]\ love_*(v)(u_2) \leftrightarrow v = m]$

In any case, the distribution of these readings in Hungarian is as follows:

(30)=(28') $[_T$ Mindenki] $[_F$ Máriát] szereti
Everybody MARY+ACC loves, *14*
everybody he$_2$ MARY+ACC loves, *22*

(31)=(29') $[_{Tc}$ Mindenki] $[_F$ Máriát] szereti
Everybody MARY+ACC loves, *22*
Mary *everybody* Him$_1$ loves, *21*

(32)=(29') $[_F$ Máriát] szereti mindenki
MARY+ACC loves everybody, *22*
Mary HIM$_1$ loves everybody, *21*

That is, if *mindenki* precedes the Focus but does not receive a contrastive topic intonation (as in (30)), it is bound to have *wider scope than Focus*.

On the other hand, when it receives a contrastive intonation in T position (as in (31)) or is in neutral position (as in (32)) it must have *more narrow scope than Focus.*[7] (This latter fact suggests that, besides conceivable conventional implicatures associated with Topic-contrast, the intonation of T deserves special attention.)

Once again, the validity of these claims can be tested againsts negation. Remember that I treated Focus-negation as a special kind of sentential negation, the F-filling rule having both an 'affirmative' and a negative version. It follows then that a *-w* quantifier may not be introduced after 'negative Focusing', and indeed, (30n) is ungrammatical while (31n) and (32n) are not:

(30n)* $[_T$ Mindenki$]$ nem $[_F$ Máriát$]$ szereti

(30n') $\forall u_2 \in [human_*(u_2)] \neg \forall v[love_*(v)(u_2) \leftrightarrow v = m]$

(31n) $[_{Tc}$ Mindenki$]$ nem $[_F$ Máriát$]$ szereti

(31n') $\neg\forall v[\forall u_2 \in [human_*(u_2)]\ love_*(v)(u_2) \leftrightarrow v = m]$

(32n) Nem $[_F$ Máriát$]$ szereti mindenki

(32n') = (31n')

Nevertheless, (30n') i.e. that nobody loves MARY being a perfectly good meaning to express, one may ask how it is expressed then in Hungarian. With this we have to make a short excursus.

Hungarian has so-called 'multiple negation', e.g.:

(33) Nem ment senki sehová semmikor
not went noone nowhere no time

Senki is to be translated as $\lambda P\forall x[human(x) \rightarrow {}^{\vee}P(x)]$ (and similarly for its brothers). Its behaviour is in many respects similar to that of *anyone*; note however the following qualifications: (i) *se(m)* α may only occur in negative sentences, and (ii) it is not only bound to have wider scope than negation but must also be introduced immediately after the negation which triggers it (of course, in case there is more than one *se(m)* α in the sentence, this latter property is inherited). That (ii) is the case can be easily demonstrated on a Focus-free example: (34) is only two, rather than six, ways ambiguous:

(34) Nem látott két fiút senki
not saw two boy+acc noone

(34) a.' $\forall v[\ldots \rightarrow \neg\exists_2 u[\ldots \wedge saw_*(u)(v)]]$
$\exists_2 u[\ldots \wedge \forall v[\ldots \rightarrow \neg\ saw_*(u)(v)]]$

Coming back to interaction with Focus: it appears that *senki* is subject to the same restrictions with respect to Focus as *mindenki*[8], the differences between their behaviour being accountable for by their opposite scope restrictions with respect to negation (and by (ii) for *senki* above). The 'grammatical version' of (30n) is (35):

(35) [$_T$ Senki] nem [$_F$ Mariat] szereti

Noone not MARY+ACC loves, *14*

noone he$_2$ not MARY+ACC loves, *22neg*

Mary he$_2$ HIM$_1$ loves, *21*

On the other hand, both (36) and (37) are ungrammatical since *senki* cannot at the same time have more narrow scope than Focus and wider scope than negation:

(36)* [$_{Tc}$ Senki] nem [$_F$ Máriát] szereti

(37)* Nem [$_F$ Máriát] szereti senki

Cases with preverbal negation also conform to predictions:

(38)* [$_T$ Senki] [$_F$ Máriát] nem szereti

Noone MARY+ACC not loves, *14*

noone he$_2$ MARY+ACC not loves, *22*

(39) [$_{Tc}$ Senki] [$_F$ Máriát] nem szereti

Noone MARY+ACC not loves, *14*

Mary *noone* him$_1$ not loves, *21*

he$_1$ noone him$_2$ not loves, *14*

noone he$_4$ him$_2$ not loves, *11b*

(40)=(39) [$_F$ Máriát] nem szereti senki

4.3. As I mentioned at the beginning of this section, it is easy to demonstrate the validity of positional scope restriction claims in the case of quantifiers with lexically given restrictions. With natural numbers in the place of *minden* or *se(m)* I can only suggest that, at least on the preferred readings, they seem to obey the same principles. More precisely: it appears that (41) is unambiguous indeed, with *two boys* having more narrow scope than Focus:

(41) [$_F$ Máriát] szereti két fiú 'MARY is loved by two boys'

On the other hand, whether the intonation of T is contrastive does not seem to be so decisive here:

(42) $[_{T}$ Két fiú] $[_{F}$ Máriát] szereti
(43) $[_{Tc}$ Két fiú] $[_{F}$ Máriát] szereti } perhaps both ambiguous

These differences between the roles of T or Tc in the case of natural numbers and universal quantifiers may be due to the marked properties of the latter (cf. fn.7).

At the beginning of this section I also suggested that apparently there are no specific restrictions for non-F quantifiers unless they follow from the restrictions on interaction with Focus (or negation). Although the following claims would be difficult to prove in the fashion I adhered to so far, I suggest that while (44) is ambiguous with respect to the relative scopes of ∀ and ∃, (45) is not:

(44) $[_{T}$ Mindenki] $[_{F}e]$ látott egy filmet Máriával
everybody saw a film+acc Mary-with

(45) $[_{T}$ Mindenki] $[_{F}$ Máriával] látott egy filmet
everybody Mary-with saw a film+acc

Given that the relative positions of *mindenki* and *egy filmet* are the same in (44) and (45) - regardless of whether it be stated in linear or dominance terms - this can only be explained by the absence of an exhaustive Focus from (44). That is, there is nothing in (44) to 'order' the non-F quantifiers while in (45) F_{22} does the job.

4.4. To conclude this section, let me discuss two scope problems arising in connection with the translation rule I gave for F-filling:

T22. If $\alpha \in P_{T}$ and $\beta \in P_{t/T}$ then $F_{22}(\alpha,\beta)$ translates as
$$\beta' = \lambda \mathcal{P} \forall R[\alpha'({}^{\wedge}\lambda x[\beta'({}^{\wedge}\lambda P\, {}^{\vee}P(x)) \wedge {}^{\vee}R(x)]) \rightarrow {}^{\vee}\mathcal{P}(R)]$$

In 3.2 it has already been shown that the right hand side of the equation is not in fact unnecessarily overcomplicated. Nevertheless, there are still two disastrous looking properties of this formula:

(i) It is easy to see that - as it is stated in S22. - we may not allow α to have the form he_{2k}. Imagine, for instance, the following derivation:

(46) $[_{F}$ Két fiu] fut

*TWO BOYS run
two boys HE_2 run
he_2 HE_1 run, *21*

(46') $\exists x \exists y[\neg(x{=}y) \wedge boy(x) \wedge boy(y) \wedge \lambda \mathcal{P}_1[{}^{\vee}\mathcal{P}_1({}^{\wedge}run)] = \lambda \mathcal{P} \forall R[run(x)$
$\wedge {}^{\vee}R(x) \rightarrow {}^{\vee}\mathcal{P}(R)] \wedge \lambda \mathcal{P}_1[{}^{\vee}\mathcal{P}_1({}^{\wedge}run)] = \lambda \mathcal{P} \forall R[run(y) \wedge {}^{\vee}R(y) \rightarrow {}^{\vee}\mathcal{P}(R)]]$

that is, (46') would say that there are two distinct boys such that only the one of them runs and only the other of them runs.

(ii) It is also easy to see that, regardless of how β is derived, α will get the widest scope in the ϕ of β, for instance:

let α' be $\lambda P \exists y[boy(y) \wedge {}^{\vee}P(y)]$ and
let ϕ' be $\forall z \epsilon[girl(z)]\ love(\mathcal{P}_1)(z)$.

Then T22. will yield

(47') $\lambda \mathcal{P}_1[\forall z \epsilon[girl(z)]\ love(\mathcal{P}_1)(z)] =$
$\lambda \mathcal{P} \forall R[\exists y[boy(y) \wedge \forall z \epsilon[girl(z)]\ love_*({}^{\vee}y)({}^{\vee}z) \wedge {}^{\vee}R(y)] \rightarrow {}^{\vee}\mathcal{P}(R)]$

and similarly if ϕ' is the value of F_{11}.

While not wishing to pretend that T22. is intended to have these properties, let me point out that their consequences are in fact not so disastrous as one might imagine.

The negation problem is the easiest to explain away: in 4.1 I argued that the quantifier in F must have wider scope than preverbal negation and therefore ϕ' will never be the value of F_{11}. This, together with the restriction in S22. ensures that this translation will never lead to contradictions and it is actually possible to show that T22. is in keeping with the algebraic requirements of *Universal Grammar*.

There remains the question of what are the interpretation options that are unfortunately excluded in view of (i) and (ii). For instance, we might want both (48) and (49) to be two ways ambiguous:

(48) $[_T$ Minden lány] $[_F$ egy fiút] szeret
every girl a boy+acc loves

a. For every girl, she loves only one person, namely, a boy; but the boys may vary with the girls. Permitted reading.
b. There is a boy (say, Peter) such that for every girl, she loves him and only him. Excluded in view of (i).

(49) $[_{TC}$ Minden lány] $[_F$ egy fiút] szeret

a. There is a boy (say, Peter) such that he is the only person unanimously loved by the girls. Permitted reading.
b. Same as *a.* but with possibly different boys. Excluded in view of (ii), cf. (47').

It is clear that (48)b. is a very nice meaning; nevertheless, it being a special case of (48)a. it is tolerable if it cannot be expressed directly. As for (49): it seems that its strongly preferred reading is the permitted one. (49)b. is actually terribly vague and although I can imagine situations in which one would use (49) with this meaning, some circumscribed version is a lot more plausible to occur. It is possible that there are cases in which the shortcomings of T22. lead to more counterintuitive consequences; in virtue of those I have discovered so far, however, it seems it can be accepted at least as a preliminary formulation.

5. VERBS, CONVERBS, AND CONCLUSION

Although my attention in this paper is centered around terms in F position, the background of my investigations seems to require some indication of what other problems should be taken account of in this framework. Therefore I add a few informal notes about the verb and the converb.

Intuitively, these two constituents may be interesting for the following reasons: (i) It will have become clear from 2.1 that É. Kiss's syntax does not allow the *finite verb* to occupy either of the distinguished positions. One may ask, however, whether it may really not be associated with those kinds of interpretational surplus that Topic and Focus may carry. (ii) The *converb* is a so-called reduced complement of the verb. Its function is somewhat similar to that of verbal prefixes in Russian in that it may make the verb perfective and/or may change its lexical meaning. Nevertheless, the converb is a mobile constituent and, moreover, it 'tends to occupy the F position' (II.iii)). Now, what does it mean for a sentence to have a converb in its T or F position?

It seems expedient to begin with (ii). Consider (50)-(54):

(50) [$_T$ Péter] [$_F$ be] rúgott 'Peter got drunk'
Peter in kicked

(51) [$_T$ Péter] [$_F$*e*] nem rúgott be
Peter not kicked in

(52) [$_T$ Péter] [$_F$ be] nem rúgott
Peter in not kicked

(53) [$_T$ Péter] nem [$_F$ be] rúgott
Peter not in kicked

(54) [$_T$ Péter be] [$_F$*e*] nem rúgott
Peter in not kicked

(50) can be an entirely neutral sounding sentence. The feeling that its F being filled with the converb may not have much semantic significance is straightforwardly verified by the fact that its plain negation is (51), in which the converb switches to neutral position. This is not so with the other examples, however. (52) is emphatic in the sense 'Peter did not get drunk by any means'. This kind of emphasis does not seem to affect truth conditions. (53) is different:

(53)a. [$_T$ Péter] nem [$_F$ be] rúgott hanem [$_F$ le] feküdt
Peter not in kicked but down lay

b. [$_T$ Péter] nem [$_F$ be] rúgott hanem [$_F$ meg] ittasult
Peter not in kicked but *perf*. got intoxicated

That is, we encounter a phenomenon similar to the case of Focus-negation with terms. (53)a. tells what Peter did instead of getting drunk and (53)b. can be true even if, extensionally speaking, getting drunk and getting intoxicated are no different but intensionally, they are. And finally:

(54) [$_T$ Péter be] [$_F$ *e*] nem rúgott de [$_T$ énekelni] [$_F$ *e*] énekelt
Peter in not kicked but to sing sang
'Get drunk he didn't but sing, he did'

There are two interesting points about these examples. On the one hand, notice that although only the converb moves around, it is the content of the verb+converb unit that gets Focused or Topicalized. This is most evident with verbs like *berúg* whose meaning has nothing to do with kicking and therefore a contrast with, say, kicking *out* would make no sense but it can also be verified with converbs of true directional meaning. This suggests that even if the position of the verb is fixed, it may send a messenger into the marked positions and thus obtain the required interpretational surplus. On the other hand, the 'obligatory' focusing of the converb in (50) may be suspicious. Given that É. Kiss had rather technical reasons for not letting the finite verb fill F, one may wonder whether this solution may not be revised (also considering that contrasts like (53)b. are possible with bare finite verbs as well).

There is a third notable point about the word order role of the converb. Although É. Kiss claims that (2)e. is ungrammatical, this is not the case:

(2)e. [$_T$ Péter] [$_F$ *e*] szaladt be
Peter ran in

It *is* grammatical, only its tense/aspect interpretation is different from that of (2)a. (or (50)). Hungarian has nothing like an overtly marked progressive or perfect. Nevertheless, (2)e. should be translated into English either as 'Peter has run in' or as 'Peter was running in (when...)'. That this distinction is not mere speculation can be verified by pointing out that although the string (55) is usually claimed to be plainly ungrammatical in literature.

(55) Mindenki szaladt be

it is only ungrammatical if *mindenki* is pronounced as Focus and it is as good as any sentence if the main stress falls on *szaladt* and is interpreted as 'present perfect' or 'past progressive'.

The significance of these last scattered remarks about tense and aspect is as follows. It is a commonplace that word order and intonation may serve to mark such grammatical distinctions in one language that are marked with overt morphenes in others. This is the case with the definite/indefinite distinction in Slavic languages. Nevertheless, since this latter distinction can also be given a 'communicative interpretation' in terms of given and new, one might get the impression that the choice of word order in 'free word order languages' may always be associated with such communicative notions. Given that the above mentioned tense and aspect phenomena do not seem to allow such an interpretation, at least not directly, they may also serve as a warning to approach this kind of sentential articulation *first* on a purely *grammatical* basis.

FOOTNOTES

* I am grateful to my colleague Miklós Sántha for his hard and helpful criticism throughout the writing of this paper.

1. I will not mark the traces of preposed X^ns in the examples; given that Hungarian has case markers and verb agreement, this will not give rise to ambiguities.
2. Formally speaking, the characteristics of 'Topic-suggestion' are the same as those of conventional implicatures (cf. KARTTUNEN & PETERS (1979)). Nevertheless, although conventional implicatures can be handled in an exact fashion their theoretical status is entirely puzzling to me and therefore I will not operate with them in this paper (see also ibid. p.15).

3. 'Of anything else', of a so-called relevant universe of discourse. This latter notion might be formalized by using restricted quantification; for the sake of simplicity, however, I will ignore it in my formalism.
   That Focus-contrast, as opposed to Topic-contrast, has the force of assertion, rather than that of conventional implicature seems intuitively very clear for Hungarian. This might be demonstrated by showing that it passes the crucial tests, too. It is possible that focus in English, as described in JACKENDOFF (1972), for instance, does not have the same property, which may be due to the fact that it is syntactically much less marked than its Hungarian 'counterpart'.
4. In the English translations I will always use contrastive stress, rather than clefting, simply in order to maintain the 'simple sentence atmosphere'. It is possible that in certain cases clefts would be more illustrative; for instance, it seems highly improbable to me that from 'It was Mary and Eve who saw Peter' one may infer 'It was Mary who saw Peter'. Note however that by translating my examples I do not mean to make any claims about English.
5. See also 3.4 and 4.4 below.
6. For this reason I will also abandon É. Kiss's bracketing strategy and represent sentences like (14) as *Nem* [$_F$ *Máriát*] *veri Péter*.
7. Following É. Kiss I labelled *mindenki* in (30) and (31) as T although syntactically this is not unproblematic since, as opposed to well-behaved Ts, *mindenki* may not undergo 'scrambling' here:

   A bortól mindenki [$_F$ be] rúgott
   the wine-from everybody in kicked
   *Mindenki a bortól [$_F$ be] rúgott
   everybody the wine-from in kicked

   (The same restriction applies to a number of other X^ns, too, e.g. non-F *wh*-words in multiple questions.) Further, *mindenki* may only receive Topic-contrastive intonation if the sentence has a Focus or else if F_{11} applies in the derivation. These facts may make one wonder if it is to be maintained that everything that precedes F is in T or some other solution should be chosen. Certain considerations that would be far too lengthy to elaborate here suggest, however, that it is preferable to retain T here and constrain the interchangeability claim for a specified subset of X^ns. (Note that the / notation I used in 3.2 may also be used for marking where these whimsical Ts must get.)
8. *Se(m)* α also resembles *minden* α in that it may not fill F either. For

proving this the converb-test cannot be used since F_{11} switches the converb to neutral position; nevertheless, it happens that the relative clause in α need not be extracted or extraposed (cf. (II.iv)), which substantiates this claim.

REFERENCES

CHAFE, W., 1976, 'Givenness, contrastiveness, definiteness, subjects, topics, and point of view', in: Ch. Li (ed.), *Subject and topic,* Academic Press, New York, 1976.

DAHL, Ö. (ed.), 1974, *Topic and comment, contextual boundness and focus,* Helmut Buske Verlag, Hamburg.

HAUSSER, R., 1976, 'Scope ambiguity and scope restrictions in Montague grammar', in: J. Groenendijk & M. Stokhof (eds.), *Proceedings of the Amsterdam colloquium on Montague grammar and related topics,* Amsterdam papers in formal grammar 1, Centrale interfaculteit, University of Amsterdam, 1976.

HAUSSER, R., 1978, 'Surface compositionality and the semantics of mood', in: J. Groenendijk & M. Stokhof (eds.), *Proceedings of the second Amsterdam colloquium on Montague grammar and related topics,* Amsterdam papers in formal grammar 2, Centrale Interfaculteit, University of Amsterdam, 1978.

JACKENDOFF, R., 1972, *Semantic interpretation in generative grammar,* MIT Press, Cambridge.

KARTTUNEN, L. & S. PETERS, 1979, Conventional implicature, in: Oh and Dinneen (eds.), *Syntax and semantics vol.* 11, Academic Press, New York.

KEMPSON, R., 1975, *Presupposition and the delimitation of semantics,* Cambridge Univ. Press, Cambridge.

KIEFER, F., 1967, *On emphasis and word order in Hungarian,* Mouton.

KISS, É.K., 1980, 'Structural relations in Hungarian, a "free word order" language', *Linguistic Inquiry* (to appear).

KUNO, S., 1973, *The structure of the Japanese language,* MIT Press, Cambridge.

MONTAGUE, R., 1974, 'The proper treatment of quantification in ordinary English', in: R. Thomason (ed.), *Formal philosophy,* Yale Univ. Press, New Haven and London.

SGALL, P., E. HAJIČOVÁ & E. BENEŠOVÁ, 1973, *Topic, focus and generative semantics,* Kronberg, Taunus.

VENNEMANN, TH., 1975, 'Topics, sentence accent, ellipsis', in: E. Keenan (ed.), *Formal semantics of natural language,* Cambridge Univ. Press, Cambridge.

DATA SEMANTICS

by

Frank Veltman

0. INTRODUCTION

The usual semantical explication of (logical) validity runs as follows: an argument is valid iff it is impossible for the premises of the argument to all be true while the conclusion of the argument is not true. Compare this account of validity with the following one: an argument is valid iff it cannot possibly occur that its premises are all true *on the basis of the set of data available* while its conclusion is not true *on the basis of that set*. Do the principles of classical logic retain their validity when one changes over from the first explication of validity to the second one?

In what follows, my principal concern will be to answer this question for the case of propositional logic. Of course, the ultimate interest of the answer hangs on the claim that the second explication of validity conforms better than the first to what goes on in actual reasoning. I shall not try to support this claim in its full generality, but will restrict myself to showing that a number of problems which arise if one tries to analyse the logical behaviour of "if ... then", "must" and "may" in terms of the first explication of validity simply vanish if one uses the second explication.

It will be clear that the question at issue cannot be answered until two other questions have been settled: (i) what is a set of data, and (ii) what does it mean for a sentence to be true on the basis of a set of data? Section 1 deals with the first of these questions, and Section 2 with the second. The final section is devoted to a discussion of some of the more salient features of the resulting logic.

1. DATA SETS

The semantic system developed here differs in various respects from the kind of systems developed within the framework of possible worlds semantics. From an ontological point of view, the most important difference is that the models for a given language are built not on 'the set of possible worlds', but on 'the set of possible facts'.

I do not intend to say a great deal about the nature of facts. Yet I do want to maintain a few assumptions about them. To begin with, I trust that there is no harm in talking about *possible* facts. I shall take this notion in such a way that it is a truism to say 'All possible facts are such that it is possible for them to hold, though some of them will never actually do so'. Certain philosophers, following Quine[1], would claim that this way of speaking commits one to assuming that possible facts exist. I am not, in this respect, an unreserved follower of Quine. I doubt whether it makes much sense to speak of existence in connection with possible facts, even in the case of those possible facts which obtain here and now. However, this is not a crucial issue. I shall certainly quantify over possible facts, and if this commits me to assuming that they somehow 'exist', then I am ready to do so.

Second, I shall hold that the totality of all possible facts can be treated as a *set* (in the mathematical sense of the word). As far as I can see, the only conceivable objection to this might be that set theory does not admit sets with the properties we shall ascribe to the set of all possible facts. But in fact set theory does admit such sets.

Third, suppose we have two possible facts f and g. I shall assume that if f and g can obtain simultaneously, this simultaneous occurrence of f and g qualifies as another possible fact. This fact is called *the combination of* f *and* g. Since we would like to talk of the combination of f and g even if f and g cannot possibly hold together, we introduce as a technical convenience the so-called *improper fact,* and we stipulate that if f and g cannot obtain simultaneously, the combination of f and g amounts to this improper fact.

These considerations taken jointly give the set of possible facts the structure of a semi-lattice:

<u>DEFINITION 1</u>. A *data lattice* is a triple $\langle F,\circ,0\rangle$ with the following properties:

(i) $0 \in F$, $F \sim \{0\} \neq \emptyset$;

(ii) $\circ$ is a binary operation of F such that

(a) $f \circ f = f$

(b) $f \circ g = g \circ f$

(c) $(f \circ g) \circ h = f \circ (g \circ h)$

$0 \circ f = 0$;

(iii) . . .

EXPLANATION. The members of $F \sim \{0\}$ are to be conceived of as the possible facts. '$f \circ g$' is to be read as 'the combination of f and g'. 0 is to be thought of as the improper fact. Given our informal remarks, it will be clear that the $\circ$-operation should have the properties laid down in (a) - (d). When $f \circ g = 0$, we shal often say that f *and* g *are incompatible*[2], and when $f \circ g = f$, we shall say that f *incorporates* g.

Definition 1 is not finished, since clause (iii) remains to be completed. This clause can be expressed informally as follows: Let E be a subset of the set of possible facts and let f be some possible fact not contained in E. Suppose that it is possible for there to be a situation in which all the facts $g \in E$ obtain while f does not. Then it is also possible for there to be a situation in which all the facts $g \in E$ obtain as well as some fact h incompatible with f.

If we want to formalize this idea, then we must determine which subsets of the set of possible facts can play the part of *possible situations*. I think the right candidates for this role are the proper filters.

DEFINITION 2. Let $\langle F, \circ, 0 \rangle$ be a triple with the properties i) and ii). A *filter* in $\langle F, \circ, 0 \rangle$ is a subset S of F such that

(i) if $f, g \in S$, then $f \circ g \in S$;

(ii) if $f \in S$ and $f \circ g = f$, then $g \in S$.

A filter S is *proper* if $0 \notin S$.

Henceforth, every set of possible facts which is closed under (i) combination and (ii) incorporation will count as a possible situation, and no other set will do so.

We can now express clause (iii) in formal terms. The shortest way to do so is this:

(iii) if S is a proper filter in $\langle F, \circ, 0 \rangle$ and $f \in (F \sim S)$, then there is some proper filter S' and some $g \in S'$ such that $S \subseteq S'$ and $f \circ g = 0$.

PROPOSITION 1. Let $\langle \mathcal{F},\circ,0\rangle$ be a data lattice. Call a proper filter $\mathcal{S}$ in $\langle \mathcal{F},\circ,0\rangle$ *maximal* iff there is no proper filter $\mathcal{S}'$ such that $\mathcal{S} \subseteq \mathcal{S}'$ and $\mathcal{S} \neq \mathcal{S}'$.

(i) Every proper filter can be extended to some maximal proper filter;

(ii) a proper filter $\mathcal{F}$ is maximal iff for each $f \in \mathcal{F}$, either $f \in \mathcal{S}$ or there is some g such that $g \in \mathcal{S}$ and $f\circ g = 0$.

PROOF. (i) Omitted.[3]

(ii) follows from clause (iii) of Definition 1. □

Proposition 1(i) states that every possible situation is embedded in at least one possible situation that cannot be extended any further. Proposition 1(ii) adds that a possible situation of maximal extent is precisely a possible situation with the property that if a certain fact f does not obtain in it, then some fact incompatible with f obtains in it.

Considering (i) and (ii), it appears that the maximal proper filters in a data lattice would make excellent *possible worlds*; therefore, I shall from now on sometimes refer to them in this way[4].

We are ready now to explain the notion of a possible set of data. Informally: every set of facts that might be obtained by investigating some possible world is a possible set of data. Formally:

DEFINITION 3. Let $\langle \mathcal{F},\circ,0\rangle$ be a data lattice, and $\mathcal{D} \subseteq \mathcal{F}$. $\mathcal{D}$ is a (*possible*) *data set* in $\langle \mathcal{F},\circ,0\rangle$ iff for every finite subset $\{f_1,\ldots,f_n\}$ of $\mathcal{D}$, $f_1\circ f_2\ldots\circ f_n \neq 0$.

That this formalisation is adequate appears from Proposition 1 together with the next one.

PROPOSITION 2. Let $\langle \mathcal{F}_0,\circ 0\rangle$ be a data lattice, and $\mathcal{D} \subseteq \mathcal{F}$. $\mathcal{D}$ can be extended to a proper filter iff for every finite subset $\{f_1,\ldots,f_n\}$ of $\mathcal{D}$, $f_1\circ f_2\ldots f_n \neq 0$.

We omit the proof.[5]

One last observation before we pass on to questions of semantics: notice that the theory of facts put forward here does not carry the metaphysical burden of many other theories. It is not assumed, for example, that there are facts of minimal complexity: any fact may incorporate other facts. Neither is it assumed that there are facts of maximal complexity: any fact may be incorporated by other facts. And finally, it is not assumed that there are *negative* facts: if a certain possible fact f does not obtain in a certain possible world, then some possible fact g incompatible with f obtains in it; but there does not have to be some particular fact g incompatible with f that obtains in every possible world in which f does not obtain[6].

2. DATA SEMANTICS

What does it mean for a sentence to be true on the basis of a certain set of data? As indicated in the introduction, we shall answer this question only for a particular class of sentences. To be more specific, the sentences in question all belong to a formal language L with

(i) a vocabulary consisting of countably many atomic sentences, two parentheses, three one place operators $\neg$, *must*, and *may*, and three two place operators $\wedge$, $\vee$ and $\rightarrow$; and

(ii) the formation rules that one would expect for a language with such a vocabulary.

The operators $\neg$, *must*, *may*, $\wedge$, and $\vee$ are meant as formal counterparts of "not", "must", "may", "and", and "or", respectively. The operator $\rightarrow$ should be read as "if ... then"; if ϕ and ψ are formal translations of the English sentences ϕ' and ψ', then $\phi \rightarrow \psi$ is meant to be a formal translation of the *indicative* conditional with antecedent ϕ' and consequent ψ'.

In presenting the semantics for this language L, I shall follow usual practice and first state how its non-logical symbols are to be understood.

DEFINITION 4. A *model* (*for* L) is a quadruple $\langle F,\circ,0,I\rangle$ such that $\langle F,\circ,0\rangle$ is a data lattice and I is a function assigning some element of F to each atomic sentence of L. I is called an *interpretation* (*of* L) *into* $\langle F,\circ,0\rangle$.

Intuitively, Definition 3 can be put as follows: each atomic sentence describes a possible fact (or the improper fact). Hence, in a way the definition offers a final clue to the question of what possible facts are; apparently, possible facts are things that can be described by the most elementary kind of sentences.[7]

Let $M = \langle F,\circ,0,I\rangle$ be a model, D a data set in $\langle F,\circ,0\rangle$ and ϕ a sentence of L. In the sequel, "$\mathcal{D} \Vdash_M \phi$" abbreviates "$\phi$ is true (in M) on the basis of $\mathcal{D}$" and "$\mathcal{D} \dashV_M \phi$" abbreviates "$\phi$ is false (in M) on the basis of $\mathcal{D}$".

DEFINITON 5. Let $M = \langle F,\circ,0,I\rangle$ be a model and $\mathcal{D}$ a data set in $\langle F,\circ,0\rangle$.

- If ϕ is atomic,
  $\mathcal{D} \Vdash_M \phi$ iff $I(\phi) \in \mathcal{D}$
  $\mathcal{D} \dashV_M \phi$ iff for some $f \in \mathcal{D}$, $f\circ I(\phi) = 0$;
- $\mathcal{D} \Vdash_M \neg\phi$ iff $\mathcal{D} \dashV_M \phi$;
  $\mathcal{D} \dashV_M \neg\phi$ iff $\mathcal{D} \Vdash_M \phi$;
- $\mathcal{D} \Vdash_M (\phi\wedge\psi)$ iff $\mathcal{D} \Vdash_M \phi$ and $\mathcal{D} \Vdash_M \psi$;
  $\mathcal{D} \dashV_M (\phi\wedge\psi)$ iff $\mathcal{D} \dashV_M \phi$ or $\mathcal{D} \dashV_M \psi$;
- $\mathcal{D} \Vdash_M (\phi\vee\psi)$ iff $\mathcal{D} \Vdash_M \phi$ or $\mathcal{D} \Vdash_M \psi$;
  $\mathcal{D} \dashV_M (\phi\vee\psi)$ iff $\mathcal{D} \dashV_M \phi$ and $\mathcal{D} \dashV_M \psi$;
- $\mathcal{D} \Vdash_M (\phi\rightarrow\psi)$ iff for every data set $\mathcal{D}' \supseteq \mathcal{D}$, if $\mathcal{D}' \Vdash_M \phi$ then $\mathcal{D}' \Vdash_M \psi$;
  $\mathcal{D} \dashV_M (\phi\rightarrow\psi)$ iff for some data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \Vdash_M \phi$ and $\mathcal{D}' \dashV_M \psi$;
- $\mathcal{D} \Vdash_M$ *may* ϕ iff for some data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \Vdash_M \phi$;
  $\mathcal{D} \dashV_M$ *may* ϕ iff for no data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \Vdash_M \phi$;
- $\mathcal{D} \Vdash_M$ *must* ϕ iff for no data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \dashV_M \phi$;
  $\mathcal{D} \dashV_M$ *must* ϕ iff for some data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \dashV_M \phi$.

The remainder of this section is devoted to a discussion of this definition. But first I need to introduce some concepts that will play a prominent part in that discussion as well as in the next section.

DEFINITION 6. Let ϕ be a sentence.
ϕ is T-stable iff for every model M and data set $\mathcal{D}$, if $\mathcal{D} \Vdash_M \phi$, then $\mathcal{D}' \Vdash_M \phi$ for every data set $\mathcal{D}' \supseteq \mathcal{D}$.
ϕ is F-stable iff for every model M and data set $\mathcal{D}$, if $\mathcal{D} \dashV_M \phi$, then $\mathcal{D}' \dashV_M \phi$ for every data set $\mathcal{D}' \supseteq \mathcal{D}$.
ϕ is stable iff ϕ is both T-stable and F-stable.

So, intuitively, a sentence ϕ is T-stable iff it has the following property: once ϕ has turned out to be true on the basis of some set of data, ϕ will remain true, whatever additional data may come to light. Likewise an F-stable sentence has the property that once its falsity has been established, there is no possibility that further investigations will yield a set of data on the basis of which it is not false.

It is not the case that every English sentence is T-stable and F-stable in this intuitive sense. We shall meet some examples of unstable sentences when we come to discuss the truth and falsity conditions of sentences of the form $(\phi\rightarrow\psi)$, *may* ϕ and *must* ϕ. But we shall discuss sentences of different forms first.

2.1. Atomic sentences

According to Definition 5, an atomic sentence ϕ is true on the basis of a certain set $\mathcal{D}$ of data iff the fact described by ϕ belongs to $\mathcal{D}$. And an atomic sentence ϕ is false on the basis of a certain set $\mathcal{D}$ of data iff the fact described by ϕ is incompatible with some element of $\mathcal{D}$.

Let M be a model and ϕ an atomic sentence. Notice:

- If $\mathcal{D} \Vdash_M \phi$, then $\mathcal{D}' \Vdash_M \phi$ for every data set $\mathcal{D}' \supseteq \mathcal{D}$.
  If $\mathcal{D} \dashv\vdash_M \phi$, then $\mathcal{D}' \dashv\vdash_M \phi$ for every data set $\mathcal{D}' \supseteq \mathcal{D}$.
- There are data sets $\mathcal{D}$ such that neither $\mathcal{D} \Vdash_M \phi$ nor $\mathcal{D} \dashv\vdash_M \phi$.
- There are no data sets $\mathcal{D}$ such that both $\mathcal{D} \Vdash_M \phi$ and $\mathcal{D} \dashv\vdash_M \phi$.
- If $\mathcal{D}$ is a maximal data set, then either $\mathcal{D} \Vdash_M \phi$ or $\mathcal{D} \dashv\vdash_M \phi$.

In other words, each atomic sentence ϕ is stable - once its truth or falsity has been established, it has been established for good. However, it is not always possible to decide on the basis of the data available whether ϕ is true or false. Of course, ϕ can never turn out to be both true and false. And ultimately ϕ must turn out to be either true or false.

The third and the fourth of the above observations apply to all sentences:

<u>PROPOSITION 3</u>. Let M be a model, $\mathcal{D}$ a data set (pertaining to M), and ϕ a sentence.

(i) It is not the case that both $\mathcal{D} \Vdash_M \phi$ and $\mathcal{D} \dashv\vdash_M \phi$;

(ii) if $\mathcal{D}$ is *maximal*, then either $\mathcal{D} \Vdash_M \phi$ or $\mathcal{D} \dashv\vdash_M \phi$.

<u>PROOF</u>. Induction on the complexity of ϕ. □

It may be very well be that a certain fact f does not occur in a certain set $\mathcal{D}$ of data, but does hold in any possible situation in which all facts in $\mathcal{D}$ hold. According to Definition 5, a sentence ϕ describing f is *not* true on the basis of $\mathcal{D}$ in such a case. Yet wouldn't it be plausible to call ϕ true on the basis of $\mathcal{D}$ here?[8]

I do not think so. Of course, if one keeps on adding more information to $\mathcal{D}$, then $\mathcal{D}$ will inevitably grow into a data set $\mathcal{D}'$ on the basis of which ϕ is true. Consequently, I would not object if one were to call the sentence *must* ϕ true on the basis of the set $\mathcal{D}$ of data. Nor would I object if one were to call the sentence ϕ just *true* - without an explicit reference to the evidence involved. I think, however, that it would blur an important distinction - that between *direct* and *indirect* evidence - if one were to maintain that it is simply and solely *on the basis of the set $\mathcal{D}$ of data* that the sentence ϕ is true.

2.2. Negation

I trust that the truth and falsity conditions for sentences of the form $\neg\phi$ do not need any further explanation. It may, however, be illuminating to compare these conditions with a few alternatives.

Presumably, it will not be difficult to convince the reader that the following stipulation would have been completely mistaken:

(*) $\mathcal{D} \Vdash_{M} \neg\phi$ iff $\mathcal{D} \nVdash_{M} \phi$.

If the few data presently at my disposal do not allow me to conclude that it is raining in Ipanema, this does not mean that they allow me to conclude that it is not raining there. Hence, (*) does not capture the meaning of English negation. Within the present framework, the equivalence expressed by (*) only holds in case $\mathcal{D}$ is a maximal data set, but that is a rather exceptional case.

Readers familiar with Kripke's semantic analysis of intuitionistic logic or with model theoretic forcing[9] will be attracted to the following alternative to the account of negation given in Definition 5:

(**) $\mathcal{D} \Vdash_{M} \neg\phi$ iff for every data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \nVdash_{M} \phi$.

I can hardly imagine that anyone would adhere to this (**)-definition and yet agree with the falsity conditions proposed in Definition 5; there seem to be no grounds for denying that the following two statements are equivalent:

(i) ϕ is false on the basis of the data;
(ii) the negation of ϕ is true on the basis of the data.

So I would expect the supporters of (**), if any, to completely reject our falsity conditions, rather than to reject the equivalence between (i) and (ii). The incorporation of (**) in Definition 5, therefore, would almost certainly bring a drastic revision of the entire system along with it.

At this moment, we are not yet in a position to explain in detail why

Definition 5 offers a better analysis of the meaning of negation in English than (**) does. I shall here briefly sketch the relevant argument trusting that the remainder of this paper will enable the reader to fill in the details for himself.

To begin with, it is worth noting that the negation described by (**) is expressible within the framework presented here, albeit not by means of the operator $\neg$. Still,

$$\mathcal{D}' \Vdash_{M} \phi \text{ for every data set } \mathcal{D}' \supseteq \mathcal{D} \text{ iff } \mathcal{D} \Vdash_{M} \textit{must } \neg\phi.$$

Hence, the easiest way to compare the (**)-negation and the negation of Definition 5, is to study the different properties attributed by Definition 5 to sentences of the form *must* $\neg\phi$ on the one hand, and sentences of the form $\neg\phi$ on the other. By doing so for different kinds of sentences, one will undoubtedly sooner or later arrive at the conclusion that 'not' has more in common with the operator $\neg$ than with the operator *must* $\neg$. The reader is invited to test this for himself - each of the following cases will settle the matter: (i) ϕ is atomic, (ii) ϕ is a sentence of the form $(\psi\rightarrow\chi)$, (iii) ϕ is a sentence of the form *must* ψ.

2.3. Disjunction and conjunction

English sentences of the form $\ulcorner\phi$ or $\psi\urcorner$ are often uttered in a context where the available data do not enable the speaker to decide which of the sentences ϕ and ψ *are* true, but only tell him that *at least one* of the sentences *has to be* true. Moreover, it would seem that sentences of the form $\ulcorner\phi$ or $\psi\urcorner$ are sometimes true, and indeed true on the basis of the data, when uttered in such a context. So it is quite possible, I think, that the police superintendent who says that either Mr. B. or Mr. C. killed Mrs. D. says something that is true on the basis of the available evidence, even though it may be weeks before the case of Mrs. D. is definitively solved.

If this observation is correct, then the only conclusion to be drawn is that in most contexts the operator $\vee$ cannot serve as the formal counterpart of 'or'. According to Definition 5, a sentence of the form $(\phi\vee\psi)$ is not true on the basis of the data unless it is possible to decide which of the sentences ϕ and ψ is true on that basis - and, on most occasions, this is a bit too much to ask.

Fortunately, the present system provides yet another possible analysis of disjunctive sentences: in place of a sentence of the form $(\phi\vee\psi)$, one can take a sentence $must(\phi \vee \psi)$ as their formal translation. $must(\phi \vee \psi)$ is true on the basis of the data set $\mathcal{D}$ iff for no extension $\mathcal{D}'$ of $\mathcal{D}$, both ϕ and ψ are false on the basis of $\mathcal{D}'$; in view of Proposition 3, this means that at least one of the sentences ϕ and ψ will eventually turn out to be true on the basis of the data if one continues to accumulate information.

At this point the reader may wonder why I did not assign to sentences of the form $(\phi \vee \psi)$ the truth and falsity conditions which are now associated with sentences of the form $must(\phi \vee \psi)$. Wouldn't that have been a more elegant procedure?

The reason that I did not proceed that way is this: sometimes disjunction *is* used in the manner formally captured by the truth and falsity conditions associated with the operator $\vee$. Here are a few examples:

- It is not the case that Mr. B. or Mr. C. killed Mrs. D.
- If Mr. E. or Mr. F. killed Mrs. D., then Mr. B. and Mr. C. are innocent.
- Maybe Mr. E. or Mr. F. killed Mrs. D.

Actually, from a syntactical point of view, there are only a few cases (the case where 'or' occurs as the main connective of the relevant sentence being the most obvious) in which the meaning of English disjunction does not seem to conform to the meaning of $\vee$. Yet I venture the hypothesis that even in these special cases the *literal* meaning of 'or' *can* be equated with the meaning of $\vee$, and that it is for *pragmatic* reasons that one is inclined to understand a statement of the form $\ulcorner\phi$ or $\psi\urcorner$ as $\ulcorner$it must be the case that ϕ or $\psi\urcorner$: to put it briefly, if one were to take such a statement literally, one would be forced to assume that its utterer is violating the conversational *maxim of quantity*.[10] If, on the other hand, the relevant disjunction is embedded in a more complex sentence, then this predicament is less likely to arise and therefore one can in general take the disjunction at its face value in such cases.

The truth and falsity conditions pertaining to conjunction need no further comment - if indeed the reader is not inclined to barter the falsity conditions of $\ulcorner\phi$ and $\psi\urcorner$ for the truth conditions of $\ulcorner$it cannot be that both ϕ and $\psi\urcorner$.

PROPOSITION 4. Suppose $\neg$, $\wedge$ and $\vee$ are the only operators occurring in ϕ. Then ϕ is stable.

In the sequel, I shall sometimes discriminate between the sentences in which $\neg$, $\wedge$ and $\vee$ are the only occurring operators and the other ones by calling the first *descriptive* and the last *nondescriptive*. Intuitively, the difference between these two kinds of sentences amounts to this: by uttering a descriptive sentence a speaker only informs his audience of the data he has gathered *so far*. By uttering a non-descriptive sentence he also gives words to his expectations about the outcome of *further* investigations.

2.4. Implication

According to Definition 5, a sentence of the form ⌜If ϕ then ψ⌝ is true on the basis of a set $\mathcal{D}$ of data iff there is no possibility of extending $\mathcal{D}$ into a data set $\mathcal{D}'$ on the basis of which ϕ is true and ψ is not true: if, by any chance, further investigations should reveal that ϕ is true, they will reveal that ψ is true too. Furthermore, it is stated that ⌜If ϕ then ψ⌝ is false on the basis of a set $\mathcal{D}$ of data iff, given $\mathcal{D}$, it is still possible that further investigations will yield an extension $\mathcal{D}'$ of $\mathcal{D}$ on the basis of which ϕ is true and ψ is false.

It will be clear that on this account a sentence of the form ⌜If ϕ then ψ⌝ is not necessarily F-stable. This, I hope, conforms to the reader's intuitions. Consider for instance the sentence 'If Mary went to the party, then John went there, too', and suppose that John's best friend is Peter. Peter happens not to know that John is head over heels in love with Mary, and, accordingly, his data allow for the possibility that Mary attended the party and John did not do so. So, on the basis of the limited set of data available to Peter, the sentence 'If Mary went to the party, then John went there, too' is false. On the other hand, it is very likely that Peter will be able to exclude this possibility - knowing John for what he is - as soon as he learns that John has fallen in love again. So, on the basis of this extension of Peter's data, the sentence 'If Mary went to the party, then John went there, too' will probably not be false[11] anymore. Hence, it is not F-stable.

Let ϕ be F-stable and suppose that ϕ is false on the basis of the data set $\mathcal{D}$. Then according to Definition 5, ⌜If ϕ then ψ⌝ is true on the basis of $\mathcal{D}$ for any sentence ψ. Likewise: let ψ be T-stable and suppose that ψ is true on the basis of $\mathcal{D}$. Then ⌜If ϕ then ψ⌝ is true on the basis of $\mathcal{D}$ for any sentence ϕ.

In other words, the present treatment of conditionals does not meet the requirement that a sentence of the form ⌜If ϕ then ψ⌝ should *never* be true unless the antecedent ϕ is somehow 'relevant' to the consequent ψ.[12]

Should we regret this?

There is, I think, no need to do so: pragmatic constraints ensure that a conditional will normally be uttered only in circumstances where the antecedent is somehow 'relevant' to the consequent. Hence, there is no need to incorporate relevance into the semantics.

Let me indicate why I think that relevance can be delegated to the pragmatics.

(i) The most natural context of utterance for an indicative conditional $\ulcorner$If ϕ then $\psi\urcorner$ - and here I restrict myself to the case where both ϕ and ψ are *descriptive*[13] - is one in which the following conditions are satisfied: (a) it is not the case that ψ is true on the basis of the data available, though (b) it is possible that ψ will on further investigation turn out to be true; (c) it is not the case that ϕ is true on the basis of the data available, though (d) it is possible that ϕ will on further investigation turn out to be true.[14] (If condition (a) is not satisfied, then by the maxims of quantity and manner ψ should be uttered rather than $\ulcorner$If ϕ then $\psi\urcorner$, for ψ is both stronger and less wordy than $\ulcorner$If ϕ then $\psi\urcorner$. Likewise, if condition (d) is not fulfilled, $\ulcorner$It must be the case that not $\phi\urcorner$ should be uttered rather than $\ulcorner$If ϕ then $\psi\urcorner$. Furthermore, if both (a) and (d) are satisfied and either (b) or (c) are not satisfied, then $\ulcorner$If ϕ then $\psi\urcorner$ is false on the basis of the data available. Thus, in view of the maxim of quality, it is forbidden to utter $\ulcorner$If ϕ then $\psi\urcorner$ in either of these cases.)

(ii) Now, if a sentence of the form $\ulcorner$If ϕ then $\psi\urcorner$ is uttered in the circumstances appropriate to its use, then the present truth condition by itself guarantees that this sentence cannot be true unless the antecedent ϕ is highly relevant to the consequent ψ: whenever the available data are extended in a way that results in ϕ being true on the basis of the new data set, ψ must be true on the basis of that extended data set too. It will be clear that there must be some positive connection between ϕ and ψ if this is to be so in circumstances where in particular the conditions (a) and (d) are satisfied.

2.5. *may* and *must*

The clearest examples of T-unstable sentences are found among sentences of the form $\ulcorner$it may be the case that $\phi\urcorner$. A sentence of this form - take 'it may be snowing' - will often at first (as you awake one winter morning) be true on the basis of the data available, and then (open the curtains and what do you see?) turn out false as soon as new data become available. In view of Definition 5, this should be a very common occurrence, for the definition states (i) that a sentence of the form ⌜it may be the case that ϕ⌝ is true on the basis of the data $\mathcal{D}$ as long as it is possible for $\mathcal{D}$ on further investigation to grow into a set of data on the basis of which ϕ is true; and (ii) that such a sentence is false on the basis of the data as soon as this possibility can be excluded.

In the previous pages I have hinted several times at the truth and falsity conditions associated with the operator *must*. According to Definition 5, a sentence of the form ⌜It must be the case that ϕ⌝ is true on the basis of the available data iff there is no possibility that this data set will on further investigation grow into a set of data on the basis of which ϕ is false. (Hence, as the investigation proceeds, the data will inevitably grow into a set on the basis of which ϕ is true.) However, as long as this possibility is not excluded, ⌜it must be the case that ϕ⌝ is false on the basis of the data.[15]

It is worth noting that this analysis predicts that in many cases, notably if ϕ is descriptive, a sentence of the form ⌜It must be the case that ϕ⌝ is weaker than the corresponding sentence ϕ. If a descriptive sentence ϕ is true on the basis of the data, then ⌜It must be the case that ϕ⌝ is true on that basis as well, but ⌜It must be the case that ϕ⌝ can be true on the basis of the data without ϕ being true on that basis.

That 'it must be the case that ϕ' is on most occasions weaker than ϕ itself, has been noticed by a number of authors. Lauri Karttunen[16] illustrates this phenomenon with the following examples:

(a) *John must have left*
(b) *John has left*

His informal explanation fits in neatly with out formal analysis: 'Intuitively, (a) makes a weaker claim than (b). In general, one would use (a) the epistemic *must* only in circumstances where it is not yet an established fact that John has left. In (a), the speaker indicates that he has no first-hand evidence about John's departure, and neither has it been reported to him by trustworthy sources. Instead (a) seems to say that the truth of *John has left* in some way logically follows from other facts the speaker knows and some reasonable assumptions that he is willing to entertain. A man who has actually seen John leave or has read about it in the newspaper would not ordinarily assert (a), since he is in the position to make the stronger claim in (b)'.

Similar remarks can be found in GROENENDIJK & STOKHOF (1975) and LYONS (1977). Yet despite this unanimity, so far no formal theory has been proposed which actually predicts that for descriptive sentences ϕ, $\ulcorner$It must be the case that $\phi\urcorner$ is a logical consequence of ϕ. Most theories treat *may* and *must* as epistemic modalities and depending on whether the underlying epistemic notion is either knowledge or belief, *must* ϕ turns out to be either stronger than ϕ or independent of it.

(Notice in passing that the present theory does not predict that ϕ is stronger than *must* ϕ for all sentences ϕ. Example: let $\phi = \neg(\psi\rightarrow\chi)$ and ψ,χ be descriptive. $\mathcal{D} \Vdash_{\overline{M}}$ *must* $\neg(\psi\rightarrow\chi)$ iff $\mathcal{D} \Vdash_{\overline{M}}$ *must*$(\psi\wedge\neg\chi)$, whereas $\mathcal{D} \Vdash_{\overline{M}} \neg(\psi\rightarrow\chi)$ iff $\mathcal{D} \Vdash_{\overline{M}}$ *may*$(\psi\wedge\neg\chi)$. Hence, *must* ϕ turns out stronger than ϕ here.)

The next proposition is an immediate consequence of Proposition 3.

<u>PROPOSITION 5</u>. Let M be a model and let $\mathcal{D}$ be a *maximal* data set (pertaining to M).

$\mathcal{D} \Vdash_{\overline{M}} \neg\phi$ iff $\mathcal{D} \not\Vdash_{\overline{M}} \phi$

$\mathcal{D} \Vdash_{\overline{M}} \phi\wedge\psi$ iff $\mathcal{D} \Vdash_{\overline{M}} \phi$ and $\mathcal{D} \Vdash_{\overline{M}} \psi$

$\mathcal{D} \Vdash_{\overline{M}} \phi\vee\psi$ iff $\mathcal{D} \Vdash_{\overline{M}} \phi$ or $\mathcal{D} \Vdash_{\overline{M}} \psi$

$\mathcal{D} \Vdash_{\overline{M}} \phi\rightarrow\psi$ iff $\mathcal{D} \not\Vdash_{\overline{M}} \phi$ or $\mathcal{D} \Vdash_{\overline{M}} \psi$

$\mathcal{D} \Vdash_{\overline{M}}$ *may* ϕ iff $\mathcal{D} \Vdash_{\overline{M}} \phi$

$\mathcal{D} \Vdash_{\overline{M}}$ *must* ϕ iff $\mathcal{D} \Vdash_{\overline{M}} \phi$.

In other words, it does not make much sense to use the phrases 'if ... then', 'must', and 'may' in a context where the data set is maximal: in such a context, 'if ... then' gets the meaning of the material conditional while both $\ulcorner$it must be the case that $\phi\urcorner$ and $\ulcorner$it may be the case that $\phi\urcorner$ turn out equivalent to ϕ. However, in such an ideal case there is no need to use nondescriptive sentences - the data set is *complete*; so, what could possibly be the good of speculations on the outcome of *further* investigations?

3. DATA LOGIC

DEFINITION 7. Let ϕ be a sentence and let Δ be a set of sentences. $\Delta \Vdash \phi$ iff there is no model $M = \langle F,\circ,0,I\rangle$ such that for some data set $\mathcal{D}$ in $\langle F,\circ,0,I\rangle$, $\mathcal{D} \Vdash_M \psi$ for every $\psi \in \Delta$ while $\mathcal{D} \nVdash_M \phi$.

'$\Delta \Vdash \phi$' abbreviates 'the argument Δ / ϕ (i.e. the argument with the set Δ of premises and conclusion ϕ) is valid'. We shall feel free to write '$\Vdash \phi$' instead of '$\emptyset \Vdash \phi$' and '$\Delta,\psi_1,\ldots,\psi_n \Vdash \phi$' instead of '$\Delta \cup \{\psi_1,\ldots,\psi_n\} \Vdash \phi$'. Read '$\Vdash \phi$' as '$\phi$ is valid'.

The following remarks should give the reader an idea of how the logic generated by the above definition works. A more systematic account will be given in a subsequent paper.[17]

3.1. Data logic and classical logic

Our first observations concern the initial question of this paper: In what respects does data logic differ from classical logic? The following list shows that many classical principles are valid in the sense of Definition 7 as well.

(i) $\Delta,\phi\wedge\psi \Vdash \phi$; $\Delta,\phi\wedge\psi \Vdash \psi$

(ii) $\Delta,\phi,\psi \Vdash \phi\wedge\psi$

(iii) $\Delta,\phi \Vdash \phi\vee\psi$; $\Delta,\psi \Vdash \phi\vee\psi$

(iv) If $\Delta,\phi \Vdash \chi$ and $\Delta,\psi \Vdash \chi$, then $\Delta,\phi\vee\psi \Vdash \chi$

(v) $\Gamma,\phi,\neg\phi \Vdash \psi$

(vi) $\Delta,\phi,\phi\rightarrow\psi \Vdash \psi$.

The reader will notice that this list is made up entirely of principles which underly the classical system of natural deduction. Actually, only two of the principles underlying that system are missing. These principles fail within the present context: It is not generally so that

If $\Delta,\phi \Vdash \psi$, then $\Delta \Vdash \phi\rightarrow\psi$.

Nor does it hold that

If $\Delta,\neg\phi \Vdash \phi$, then $\Delta \Vdash \phi$.

The next two principles partially make up for this:

(vii) If $\Delta,\phi \Vdash \psi$ and each $\chi \in \Delta$ is T-stable, then $\Delta \Vdash \phi\rightarrow\psi$;

(viii) If $\Delta,\neg\phi \Vdash \phi$ and each $\chi \in \Delta$ is T-stable, then $\Delta \Vdash$ *must* ϕ.

Illustrations

- It is easy to check that $\neg(\phi \vee \neg\phi) \Vdash \phi \vee \neg\phi$.
Yet if ϕ is a descriptive sentence, then $\nVdash \phi \vee \neg\phi$. (Actually, there are no valid descriptive sentences at all.)
must$(\phi \vee \neg\phi)$, on the other hand, is valid whether ϕ is descriptive or not.
In this connection, it is worth noting that also *must* $\phi \vee \neg$(*must* ϕ) and *must* $\phi \vee$ *must* $\neg\phi \vee$ (*may* $\phi \wedge$ *may* $\neg\phi$) are valid for any ϕ.

- Suppose ϕ is an atomic sentence.
Then we have that *may* ϕ, $\neg\phi \Vdash \phi$, whereas neither *may* $\phi \Vdash \phi$ nor *may* $\phi \Vdash$ *must* ϕ.

What is notable here, is not so much the invalidity of *may* ϕ / ϕ and *may* ϕ / *must* ϕ as the validity of *may* ϕ, $\neg\phi$ / ϕ. Actually, according to the present theory, any conclusion can be drawn from the premises *may* ϕ and $\neg\phi$. To put it differently, by the standards here applied, the sentence

(a) *It may be raining in Ipanema now and it isn't*

is just as contradictory as

(b) *It is raining in Ipanema now and it isn't.*

These examples show that the present theory of 'may' differs widely from the theories of 'may' developed within the framework of possible worlds semantics and pragmatics. According to the latter[18], a sentence like (a) can be perfectly true although no one can assert it without violating the *maxim of quality*; consequently, the argument *may* ϕ, $\neg\phi$ / ψ is considered not as *logically* valid, but at best as *pragmatically* valid.

(Is there any evidence in favour of the claim that arguments of the form *may* ϕ, $\neg\phi$ / ψ - with ϕ atomic, or at least F-stable - are pragmatically valid rather than logically valid? Clearly, this evidence should consist in an informal example which shows that the putative 'seeming' inconsistency of the premises of an argument of this form can, in principle, be cancelled. I am pretty sure, however, that no such example can be found[19].)

- Let ϕ and ψ be two distinct atomic sentences.
It goes without saying that $\phi,\psi \Vdash \psi$ and *may* $\phi,\psi \Vdash$ *may* ϕ.
Furthermore, $\phi \Vdash \psi \rightarrow \phi$ but *may* $\phi \nVdash \psi \rightarrow$ *may* ϕ.
Hence, the present theory labels the first of the following arguments as

valid and the second as invalid.

(a) *John's bicycle is red. Therefore, if John's bicycle is green, then it is red*

(b) *Maybe John's bicycle is red. Therefore, if John's bicycle is green, then it may be red.*

Perhaps the reader finds it difficult to accept the validity of (a). If so, he is invited to read the conclusion once more without losing sight of the premise - The conclusion does not say that John's bicycle would be red if it *had been* green.

If this does not help, then presumably the problem is that (i) the conclusion suggests that there is some positive connection between the supposed greeness of John's bicycle and its actual redness, whereas (ii) no such connection can possibly exist. However, that there is no positive connection between the antecedent and the consequent of the conclusion does not imply that the conclusion does not hold. That would only follow if the relevant conditional had been uttered in circumstances appropriate to its use (see Section 2.4). But this particular conditional is uttered in rather exceptional circumstances: given the premise of the argument, the antecedent of the conclusion is false on the basis of the data (and it will remain false if the data are extended), and its consequent is true on the basis of the data (and it will remain true if the data are extended). So we see that it is *pragmatically incorrect* to utter the conclusion in the circumstances described by the premise. However, we also see that if one does utter it anyway, then one can count the resulting statement as trivially true.

Notice in passing that the conclusion of (a) is trivially false if it is uttered in circumstances where neither the truth nor the falsity of either the antecedent or the consequent have been established. Assuming, then, that an addressee expects a speaker to observe all conversational rules, it is quite understandable that ones first reaction to (a) might be one of protest.

Let us now turn to argument (b). It is illuminating to compare this argument with the following.

(b') *Maybe John's bicycle is red. Therefore, if John's bicycle turns out green, then it is still true on the basis of the data presently at my disposal that John's bicycle may be red.*

Unlike (b), the argument (b') is valid. Roughly, the difference is this: in the consequent of the conclusion of (b') explicit reference is made to the data available at the time of utterance. The consequent of the conclusion of (b), on the other hand, implicitly refers to the potential sets of data available after John's bicycle has turned out to be green.

3.2. Substitution and Replacement

Let ϕ and ψ be two distinct atomic sentences. It is only one step from the validity of $\phi / \psi \rightarrow \phi$ to the validity of $\phi \rightarrow (\psi \rightarrow \phi)$ and from the invalidity of *may* $\phi / \psi \rightarrow$ *may* ϕ to the invalidity of *may* $\phi \rightarrow (\psi \rightarrow$ *may* $\phi)$. Still, it is worthwile to take these steps, for the resulting examples show that the *Principle of Substitution* cannot be carried over from classical logic to data logic without modification. In general, only uniform substitution of a *stable* sentence for an atomic sentence will transform a valid sentence into a valid one (Uniform substitution of an *instable* sentence may yield an invalid sentence.)

Also the *Principle of Replacement* needs to be treated with some care. Let us call the sentences ϕ and ψ *weakly equivalent* if both $\phi \Vdash \psi$ and $\psi \Vdash \phi$, and *strongly equivalent* iff $\phi \Vdash \psi$, $\psi \Vdash \phi$, $\neg\phi \Vdash \neg\psi$ and $\neg\psi \Vdash \neg\phi$. This distinction is important. Consider, for example, the sentences $\neg(\phi \vee \neg\phi)$ and $\neg(\psi \vee \neg\psi)$ where ϕ and ψ are two distinct atomic sentences. $\neg(\phi \vee \neg\phi)$ and $\neg(\psi \vee \neg\psi)$ are weakly equivalent but not strongly equivalent. If the occurrence of $\neg(\phi \vee \neg\phi)$ in $\neg\neg(\phi \vee \neg\phi)$ is replaced by an occurrence of $\neg(\psi \vee \neg\psi)$, then the resulting sentence $\neg\neg(\psi \vee \neg\psi)$ is *not* weakly equivalent to the original $\neg\neg(\phi \vee \neg\phi)$. Hence, the Principle of Replacement fails for weak equivalents. Yet it does hold for strong equivalents: if two sentences ϕ and ψ are strongly equivalent, then replacement of an occurrence of ϕ in a sentence χ by an occurrence of ψ will always yield a sentence χ' which is strongly equivalent to the original χ.

EXAMPLES.

- $\neg\neg\phi$ is strongly equivalent to ϕ
- $\phi \vee \psi$ is strongly equivalent to $\neg(\neg\phi \vee \neg\psi)$
- $\phi \wedge \psi$ is strongly equivalent to $\neg(\neg\phi \vee \neg\psi)$
- *may* ϕ is strongly equivalent to $\neg(\phi \rightarrow \neg\phi)$
- *must* ϕ is strongly equivalent to $\neg\phi \rightarrow \phi$[20].

So we see that $\phi \vee \psi$, $\phi \wedge \psi$, *may* ϕ, and *must* ϕ can be considered as mere abbreviations of $\neg(\neg\phi\wedge\neg\psi)$, $\neg(\neg\phi\vee\neg\psi)$, $\neg(\phi\rightarrow\neg\phi)$ and $\neg\phi\rightarrow\phi$, respectively. In other words, in principle it is possible to give a more economical presentation of the present system by taking $\wedge$, $\neg$, $\rightarrow$ (or alternatively $\vee$, $\neg$, $\rightarrow$) as primitive operators and defining the other operators in terms of them. (It is not possible to find three other operators among the ones given with which one can do the same.)

3.3. *may* and *must*

From the above observations it is clear that the logical properties of *may* and *must* are completely determined by the properties of $\neg$ and $\rightarrow$. Still, it seems worthwile to examine to what extent *must* and *may* behave like standard model operators.

- *may* ϕ is strongly equivalent to $\neg$*must* $\neg\phi$
- If $\Vdash \phi$, then $\Vdash$ *must* ϕ
- If ψ is F-stable, then $\Vdash$ *must*$(\phi\rightarrow\psi)\rightarrow($*must* $\phi\rightarrow$ *must* $\psi)$
- $\nVdash$ *must* $\phi\rightarrow\phi$
- $\Vdash$ *must*(*must* $\phi\rightarrow\phi$)
- $\Vdash$ *must* $\phi\rightarrow$ *may* ϕ
- $\Vdash$ *must* $\phi\rightarrow$ *must must* ϕ.

Thus, at first sight, it would seem that *must* and *may* behave like the obligation and permission operators of some system of *deontic* logic. But we also find:

- If ϕ is F-stable, then $\Vdash$ *must may* $\phi\rightarrow$ *must* ϕ, which would be a rather strong result for a system which is marked as deontic.
- $\Vdash$ $\phi\rightarrow$ *may* ϕ, which gives the logic of *may* an *alethic* flavour.
- If ϕ is T-stable then $\Vdash$ $\phi\rightarrow$ *must* ϕ. Cf. Section 2.5.

3.4. Implication

Let us now take a closer look at the logical properties of the operator $\rightarrow$. In many respects $\rightarrow$ behaves like *intuitionistic* implication.

PROPOSITION 6. Suppose that $\wedge$, $\vee$ and $\rightarrow$ are the only operators occurring in the sentence of the argument Δ/ϕ. Then, $\Delta\Vdash\phi$ iff Δ/ϕ is intuitionistically valid.

I shall not prove this proposition here.[21]

The above result does not hold if we permit other connectives to occur in the sentences of an argument. We encountered some counterexamples earlier: every sentence of the form $\neg\neg\phi \rightarrow \phi$ is valid in the sense of Definition 7, but a sentence of that form is in general not intuitionistically valid. On the other hand, every argument of the form $(\phi\wedge\psi) \rightarrow \chi /$ $/ \phi \rightarrow (\psi\rightarrow\chi)$ is intuitionistically valid, whereas according to the present theory the validity of an argument of that form depends on the T-stability of ϕ: If ϕ is T-stable, then $(\phi\wedge\psi) \rightarrow \chi \Vdash \phi \rightarrow (\psi\rightarrow\chi)$, but if ϕ is not T-stable, then it is very well possible that $(\phi\wedge\psi) \rightarrow \chi \nVdash \phi \rightarrow (\psi\rightarrow\chi)$.

In one important respect the behaviour of $\rightarrow$ matches with the behaviour of the *strict implications* occurring in the Lewis Systems[22]: $\neg(\phi\rightarrow\psi) \Vdash may(\phi\wedge\neg\psi)$ and $may(\phi\wedge\neg\psi) \Vdash \neg(\phi\rightarrow\psi)$. This is exactly what one would find if $\rightarrow$ were the implication and *may* the possibility operator of another extension of S 0.5.

However, $\neg(\phi\rightarrow\psi)$ and $may(\phi\wedge\neg\psi)$ are only weakly equivalent and not strongly equivalent. Although we do find that $\neg\neg(\phi\rightarrow\psi) \Vdash \neg may(\phi\wedge\neg\psi)$ or, equivalently, that $(\phi\rightarrow\psi) \Vdash must(\neg\phi\vee\psi)$, it is not the case that $must(\neg\phi\vee\psi) \Vdash (\phi\rightarrow\psi)$; at best we have that $must(\neg\phi\vee\psi) \Vdash \phi \rightarrow must\ \psi$, and even this only for T-stable sentences ϕ.

Our final observations with respect to $\rightarrow$ concern the Principle of *Modus Tollens*.[23] This principle, which holds both in intuitionistic logic and in the systems of strict implication and also in such a weak system as the system R of *Relevance logic*, fails here. It is not generally so that $\phi \rightarrow \psi, \neg\psi \Vdash \neg\phi$. The closest approximation available is this: if ψ is F-stable, then $\phi \rightarrow \psi, \neg\psi \Vdash must\ \neg\phi$.

If ψ is not F-stable, even this weakened version of *Modus Tollens* does not hold. Consider, for example the premises $\phi \rightarrow (\psi\rightarrow\chi)$ and $\neg(\psi\rightarrow\chi)$, where ϕ, ψ and χ are three distinct atomic sentences. Neither $\neg\phi$ nor $must\ \neg\phi$ follow from these premises; we only have that $\phi \rightarrow (\psi\rightarrow\chi), \neg(\psi\rightarrow\chi) \Vdash may\ \neg\phi$.

An example showing that the Principle of Modus Tollens fails in natural language is due here.

Three persons are involved, Allen, Brown and Carr. Perhaps the reader met the three of them before in connection with Lewis Carroll's barbershop paradox[25]. Well, they still run a barbershop, but nowadays they do so according to the following rules: (i) At all times at least one of them must be in the shop. (ii) None of them may ever leave the shop without one of the others accompanying him.

Which of them, do you think, will be in the shop right now? It is clear, of course, that *if Carr is in, then Allen is in if Brown is in*. Furthermore, it may very well be that Allen is out in the company of Carr, while Brown minds the shop, So, *it is not the case that if Brown is in, Allen is in*. Now, by an application of the Principle of *Modus Tollens*, it would follow from the italicized sentences that Carr is out; and then, by a similar argument, one might prove that also Brown and Allen are out...

3.5. Descriptive arguments

Even if one likes the way in which the theory presented here deals with nondescriptive arguments, one may still regret the divergences from classical logic in reasonings with descriptive sentences. However, the departure from classical logic is not as drastic as one might fear at first sight:

PROPOSITION 7. Suppose $\wedge$, $\vee$ and $\neg$ are the only operators occurring in the sentences of Δ / ϕ. If Δ / ϕ if classically valid, then $\Delta \Vdash$ *must* ϕ.

The proof, which is based on Proposition 4 and Proposition 5, is left to the reader.

In other words, if by the standards of classical logic the descriptive sentence ϕ must follow from the descriptive premises Δ, then *at least* 'it must be the case that ϕ' follows from Δ be the standards set here.

NOTES

*) This paper forms part of my doctoral dissertation, which I am writing under the supervision of Johan van Benthem and Hans Kamp. I am greatly indebted to them and to Dick de Jongh for their encouragement and advice. I would also like to express my gratitude to Gerald Gazdar, Jeroen Groenendijk, Theo Janssen, Ewan Klein, Fred Landman, Ieke Moerdijk, Piet Rodenburg, Martin Stokhof, Zeno Swijtink and in particular to Roel de Vrijer for their helpful criticism. Special thanks are due to Ewan Klein and Dick de Jongh for correcting the English.

1. See especially his by now classic 'On What There Is', reprinted as Chapter 1 in QUINE (1961).
2. I want to stress that the improper fact is introduced merely as a technical convenience. In principle, one can dispense with it by taking a

partial combination operation and calling two facts f and g incompatible iff the combination of f and g is not defined.

3. The proof is identical to the proof of the Ultrafilter Theorem for Boolean Algebras. For details, see BELL & SLOMSON 1969, p.15.
4. Let $\mathcal{W}_{<F,\circ,0>}$ be the set of possible worlds generated in this manner by the data lattice $<F,\circ,0>$. Let K be the class of all sets $\mathcal{W}_{<F,\circ,0>}$. Perhaps this class K can be of some use to those who seek to base a possible worlds semantics for counterfactual conditionals on the notion of *minimal change*. (Within each structure $\mathcal{W}_{<F,\circ,0>}$ one can differentiate in an elegant way between propositions which express a fact and propositions which do not.) See POLLOCK 1976, pp.70-93.
5. The proof proceeds along similar lines as the proof of the analogous theorem for Boolean Algebras. (A data set corresponds to a data lattice as a set with the finite intersection property to a Boolean Algebra.) See BELL & SLOMSON 1969, pp.13-14.
6. The position on 'negative facts' taken here is not so different from Mr. Demos' position, which is discussed by Bertrand Russell in 'The Philosophy of Logical Atomism'. See the relevant chapter in RUSSELL 1956.
7. Admittedly, in the absence of a clear cut grammatical criterion to determine which English sentences count as most elementary, this remark is not very illuminating.
8. Taken as a rhetorical question, this question contains the proposal to alter the truth and falsity conditions of atomic sentences ϕ as follows: $\mathcal{D} \Vdash_{\overline{M}} \phi$ iff $I(\phi) \in \mathcal{D}^*$, and $\mathcal{D} \dashv\!\mid_{\overline{M}} \phi$ iff for some $g \in \mathcal{D}^*$, $I(\phi)\circ g = 0$. Here $\mathcal{D}^*$ is defined as $\cap\{E \subseteq F \mid E$ is a proper filter extension of $\mathcal{D}\}$. It is easy to check that $\mathcal{D}^*$ is a filter. So, intuitively, $\mathcal{D}^*$ is the smallest possible situation in which all facts in $\mathcal{D}$ hold.
9. See KRIPKE 1965 and KEISLER 1977. It will be obvious to anyone familiar with the subject that the present paper found some of its inspiration in the notion of forcing.
10. Throughout this paper, I shall assume that the reader is familiar with GRICE 1975.
11. Suppose the sentence 'If Mary went to the party, then John went there, too' is not false on the basis of the data. This does not imply that this sentence is true on the basis of the data, but only that the sentence 'If Mary went to the party, then John *must* have gone there, too' is true on that basis.
12. I am referring here to the requirements set by the authors and co-authors

of the sections on Relevance Logic in ANDERSON & BELNAP 1975. We could, at least partially, meet these requirements by altering our truth and falsity conditions for atomic sentences and implications in the following way:

- If ϕ is atomic, then $\mathcal{D} \Vdash_{M}$ iff (i) $I(\phi) \in \mathcal{D}^{*}$, and (ii) for no proper subset $\mathcal{D}'$ of $\mathcal{D}$, $I(\phi) \in \mathcal{D}'^{*}$. (So, intuitively, an atomic sentence ϕ is *relevantly* true on the basis of $\mathcal{D}$ just in case $\mathcal{D}$ contains a *minimum* of information guaranteeing that the fact described by ϕ obtains in every possible situation in which all the facts in $\mathcal{D}$ obtain.)
- If $\phi = \psi \rightarrow \chi$, then $\mathcal{D} \Vdash_{M} \phi$ iff $\mathcal{D} \cup \mathcal{D}' \Vdash_{M} \chi$ for every $\mathcal{D}'$ such that (i) $\mathcal{D}' \Vdash_{M} \psi$ and (ii) $\mathcal{D} \cup \mathcal{D}'$ is a data set. (Notice that for T-stable antecedents, this condition can replace the condition included in Definition 5, but that it would not work well for T-unstable antecedents.)

Conjecture: Suppose $\rightarrow$ is the only operator occurring in ϕ. Then ϕ is a theorem of Relevance Logic iff for every model M, $\emptyset \Vdash_{M} \phi$.

13. The other cases remain to be studied.
14. See also GAZDAR 1979, pp.59-61.
15. An obvious alternative to the truth and falsity conditions of sentences of the form *must* ϕ and *may* ϕ is the following:

$\mathcal{D} \Vdash_{M}$ *must* ϕ iff for every *maximal* data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \Vdash_{M} \phi$

$\mathcal{D} \dashv\Vert_{M}$ *must* ϕ iff for some *maximal* data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D}' \dashv\Vert_{M} \phi$

$\mathcal{D} \Vdash_{M}$ *may* ϕ iff for some *maximal* data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D} \Vdash_{M} \phi$

$\mathcal{D} \dashv\Vert_{M}$ *may* ϕ iff for every *maximal* data set $\mathcal{D}' \supseteq \mathcal{D}$, $\mathcal{D} \dashv\Vert_{M} \phi$.

Notice that for stable sentences ϕ, the above conditions are equivalent to the ones included in Definition 5.

The main reason that I prefer the clauses of Definition 5 to the ones given above is methodological in nature. In the above clauses reference is made to the *maximal* proper extensions of data sets. However, it can only be proved by using powerful set theoretic methods that data sets have any maximal extensions. To be more specific, Proposition 1(i), though somewhat weaker than the Axiom of Choice, is independent of the axioms of Zermelo Fraenkel Set Theory. Its status as a mathematical truth is not as solidly based as it is for these axioms. Now, if we want the above clauses for *must* and *may* to really work, we must rely on this proposition. The clauses for *must* and *may* given in Definition 5, on the other hand, do not presuppose Proposition 1(i) or any other equally questionable set theoretic proposition. Therefore, from a methodological point of view, the clauses of Definition 5 are to be preferred.

16. See KARTTUNEN 1972, p.12.
17. See VELTMAN (forthcoming).
18. See GROENENDIJK & STOKHOF 1975, pp.83-84.
19. So far, no elaborate pragmatic theory has succeeded in drawing the dividing line between logical and pragmatical-but-not-logical validity precisely as the criterion of cancellability prescribes. It appears that in particular the conclusions of arguments which owe their pragmatical validity exclusively to the maxim of quality defy any attempt to cancellation. (See GAZDAR 1979, p.46.) It is, therefore, perhaps a little premature to suppose that because the inconsistency of the premises cannot be cancelled, it follows that arguments of the form *may* ϕ, $\neg\phi$ / ψ (with F-stable ϕ) are logically rather than just pragmatically valid. Consider, however, the following version of the maxim of quality: *Do not utter a sentence* ϕ *unless* ϕ *is true on the basis of the data at your disposal*. Every argument owing its pragmatical validity exclusively to this version of the maxim of quality is logically valid in the sense of 'logically valid' discussed here, too. So, presumably, data semantics allows for a pragmatics in which 'cancellability' *can* serve as a condition that an argument *must* satisfy in order to be classified as pragmatically but not logically valid.
20. Hence, principle (viii) of Section 3.1 is in fact a special case of principle (vii).
21. See VELTMAN (forthcoming).
22. These systems are extensively discussed in HUGHES & CRESSWELL 1972.
23. Similar observations can be made with respect to the Principle of Contraposition.
24. Modus Tollens does fail in the theory of conditionals put forward in COOPER 1978. However, the evidence and explanation offered by Cooper are quite different from the evidence and explanation offered here.
25. The present example is a slight variant of this paradox, which first appeared in CARROLL 1894. I can hardly imagine that nobody has ever thought of this variant before. In my view, it is much more powerful than the rather innocent barbershop paradox itself. Yet even Cooper, who discusses Carroll's paradox at some length, does not refer to it. Cf. COOPER 1978, pp.204-205.

REFERENCES

ANDERSON, A.R. & N.D. BELNAP, 1975, *Entailment. The Logic of Relevance and Necessity,* Vol. I, Princeton University Press.

BELL, J.L. & A.B. SLOMSON, 1969, *Models and Ultraproducts: An Introduction,* North-Holland Publ. Company, Amsterdam.

CARROLL, L., 1894, 'A Logical Paradox', *Mind* 3, pp.436-438.

COOPER, W.S., 1978, *Foundations of Logico-Linguistics. A Unified Theory of Information, Language and Logic,* Reidel, Dordrecht.

GAZDAR, G., 1979, *Pragmatics. Implications, Presupposition and Logical Form,* Academic Press, New York.

GRICE, H.P., 1975, 'Logic and Conversation', in: *Syntax and Semantics 3: Speech Acts,* P. Cole and J. Morgan (eds), Academic Press, New York, 1975, pp.41-58.

GROENENDIJK, J. & M. STOKHOF, 1975, 'Modality and Conversational Information', in: *Theoretical Linguistics* 2, pp.61-112.

HUGHES, G.E. & M.J. CRESSWELL, 1972, *An Introduction to Modal Logic,* Second edition, Methuen and Co. Ltd., London.

KARTTUNEN, L., 1972, 'Possible and Must', in: *Syntax and Semantics 1,* J.P. Kimball (ed.), Seminar Press, New York, 1972, pp.1-20.

KEISLER, H.J., 1977, 'Fundamentals of Model Theory', in: *Handbook of Mathematical Logic,* J. Barwise (ed.), North-Holland Publ. Company, Amsterdam, 1977, pp.47-105.

KRIPKE, S.A., 1965, 'Semantical Analysis of Intuitionistic Logic I', in: *Formal Systems and Recursive Functions,* J.N. Crossley & M.A.E. Dummett (eds), North-Holland Publ. Company, Amsterdam, pp. 92-130.

LYONS, J., 1977, *Semantics,* Vol. II, Cambridge University Press.

POLLOCK, J.L., 1976, *Subjunctive Reasoning,* Reidel, Dordrecht.

QUINE, W.V.O., 1961, *From a Logical Point of View,* Harper and Row, New York.

RUSSELL, B., 1956, *Logic and Knowledge. Essays 1901-1950,* R.C. Marsh (ed.), George Allen and Unwin Ltd., London.

VELTMAN, F., forthcoming, 'Propositional Data Logic'.

NUMERALS AND QUANTIFIERS IN $\bar{X}$-SYNTAX AND THEIR SEMANTIC INTERPRETATION*

by

Henk J. Verkuyl

0. INTRODUCTION

This paper purports to contribute to the solution of a problem that can be described in two ways. The first is: 'Can we provide Montague-grammar with a syntax satisfying well-established needs of linguists?'. The second is: 'Can we provide Chomsky-grammar with a semantics satisfying well-recognized wishes of logical semanticists?'. Part of the problem is that one can simply deny that there is a problem at all, given the divergent goals of the respective enterprises. Nevertheless, several attempts have been made to bridge the gap (e.g. PARTEE 1975; COOPER & PARSONS 1976). The aim of my paper is to "categorialize" the generative $\bar{X}$-syntax such that it can provide a suitable basis for PTQ-semantics, maintaining its descriptive and explanatory force for linguistic purposes.

1. SOME DESCRIPTIVE MATERIAL

I shall begin with some descriptive material from the internal Noun Phrase structure. It will give an impression of what linguists - given their task to describe natural languages - regard as valuable generalizations. Linguists of all kinds agree upon the need to order the material given in (1).

(1) a. some trees - the trees - nice trees - some nice trees - the nice trees - *the some trees - *some the trees.
    b. these children - three children - these three children - *three these children.

If the basic aim of syntax is "to characterize the various syntactical categories ..." (MONTAGUE 1974, p.233), then we may assume that certain syntactic principles can hold irrespective of the appropriate semantics carried by syntactic structure. To bend one's thoughts directly toward the point in question: linguists consider the notion 'contrastive distribution' as a purely syntactic notion; two members of the same syntactic category never occur in the same syntactic position simultaneously, unless one has to do either with co-ordination or subordination of some kind. That is, in a sentence such as *This girl his sister is ill* we force the second NP into an appositional position.[1]

A second trait of linguists is their interest in the behaviour of NP's in sentences such as (2).

(2) a. There is a child in the house.
b. *There is the child in the house.

Certain principles, though very complicated and not very well understood, block the presence of definite NP's in (2b) given the existential nature of *there* (cf. GUÉRON 1976; CHOMSKY 1977; MILSARK 1977).

There is a third descriptive area that I shall touch upon before going into the syntactic tools under analysis. Linguists are interested in the internal structure of NP's such as (3),

(3) My numerous second three nice little red wooden boxes

discussed in ROOSE (1956), DE GROOT (1949), VENDLER (1968), and CLARK & CLARK (1977), among others. The question is whether NP's do or do not have fixed positions for the elements preceding *boxes* in (3).

The descriptive material given here, is taken from English and Dutch ((3) is in fact a translation from a Dutch example). Both languages show a close correspondence with regard to the material given in (1) - (3). For convenience I shall present English examples as long as there exists a parallelism between Dutch and English. In the next section I shall discuss Jackendoff's proposal concerning the specifier structure of NP's with an eye to (1) - (3).

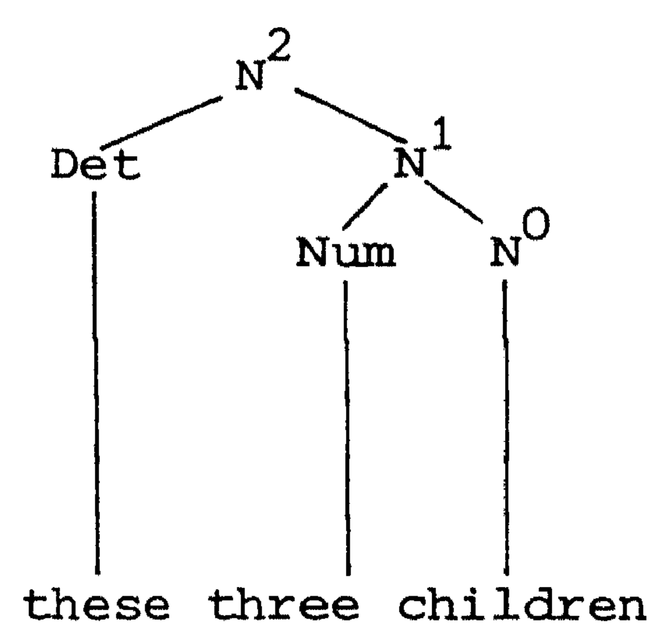
2. THE UNIFORM THREE LEVEL HYPOTHESIS (U3LH)

Chomsky's introduction of $\bar{X}$-syntax into phrase structure grammar solved the problem of how to account for nodes occurring as 'head of a phrase' (LYONS 1968, p.331; CHOMSKY 1970). The rewrite rules of $\bar{X}$-syntax are constrained such that they all fit into the form (4),

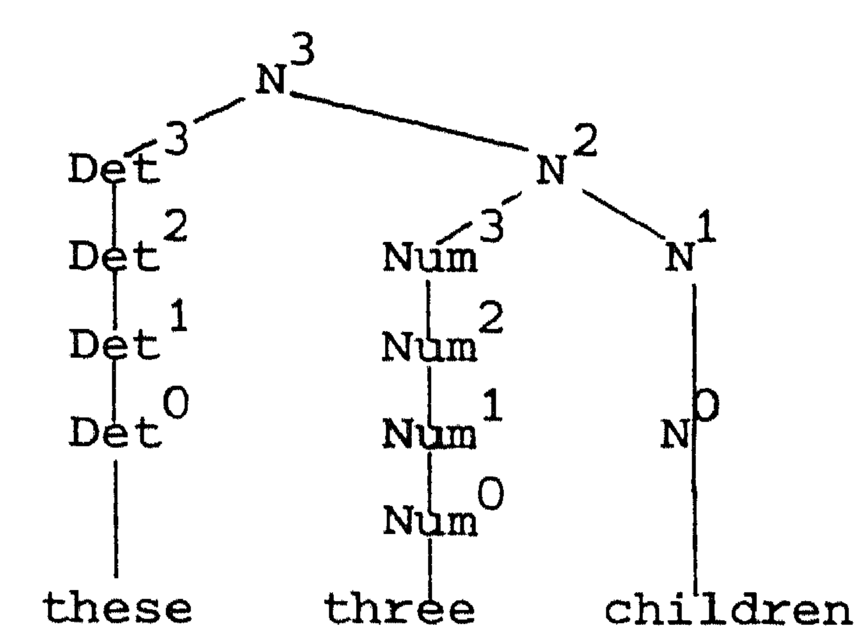
(4) $$X^{i+1} \rightarrow \ldots X^{i} \ldots$$

where X^i is the head of X^{i+1}, and where the lowest node, say X^0, is the head of the whole X-phrase. Values for X are lexical categories such as N(oun), A(djective), V(erb), P(reposition), etc. A lexical category is introduced into phrase structure as X^0, inducing higher values of *i*. The path from X^0 up to and including the highest X-node is called the X-projection line. Scheme (4) allows of the trees (5) and (6),

(5)



(6)



where 'Det" stands for 'determiner' and 'Num' for 'numeral'.

As phrase structure rules work from top to bottom, generative rewrite systems are forced to stipulate how much structure they allow above the X^0-level. JACKENDOFF (1977) is very explicit about this. He says that each lexical category defines a set of supercategories (X^1, X^2, and so on) to be related to each other by rewrite rules of the form (7),

(7) $$X^{n} \rightarrow X^{n-1} \ldots$$

where $n \leq 3$. The canonical form of (7) is (8),

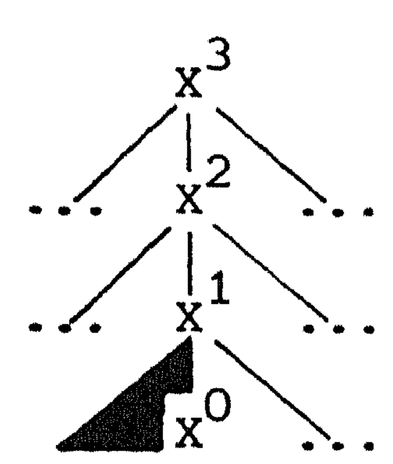
(8) $$X^{n} \rightarrow (C_1) \ldots (C_j) - X^{n-1} - (C_{j+1}) \ldots (C_k)$$

where (i) $1 \leq n \leq 3$; (ii) the values for X are N, V, A, P, Adv(erb), M(odal), Q(uantifier), Art(icle), Deg(ree) and Prt(= particle); and (iii) for all C_i, either C_i is a grammatical formative such as *Past, Possessive,* etc., or $C_i = Y^3$, for some lexical category Y (1977:36). The brackets in (8) indicate that all C_i are optional. Note that (8) excludes (5) as a possible structure for *these three children,* whereas (6) is allowed.

It is not necessary to go very deeply into the motivation for $\bar{X}$-syntax here. It suffices to mention three advantages of this approach. The first is that one can generalize with respect to parallel structural configurations across different phrasal types: *John's refusal of the offer* and *John refused the offer* are the relevant well-known examples. The second is that one can cross-classify among lexical categories, e.g. *refusal* as a noun and *refuse* as a verb have certain lexical-structural properties in common. Finally, one can generalize in terms of rules: in both *the city's destruction by the enemy* and *the city was destroyed by the enemy* passivization takes place. $\bar{X}$-syntax can account for this in terms of corresponding domains (CHOMSKY 1970; JACKENDOFF 1977; HALITSKY 1975; HORNSTEIN 1977).

What do $\bar{X}$-phrase structures contribute to semantic interpretation? As to this question Jackendoff is, in certain respects at least, quite specific. Consider his classification of complements as shown in (9).

(9)



X^2 ... : non-restrictive modification

X^1 ... : restrictive modification

X^0 ... : arguments for X^0

X^1-complements, i.e. sister nodes of X^0, are at the level of interpretation arguments for the predicate X^0, if the value for X is V or N.[2] X^2-complements are restrictive modifiers. For example, in *John saw his three children yesterday* the time adverbial is a V^2-modifier to be taken as a function mapping the V^1-predicate into a V^2-predicate of the same number of arguments, thus "restricting the extension of the sentence" by adding extra truth conditions (1977: 61). On the N-projection line restrictive clauses are dominated by N^2. That is, in *the trees that I like* is *trees that I like* an N^2 having *trees* as its head and the restrictive complement *that I like* as its complement. I restrict myself here to N and V as values for X in (9).

$\bar{X}$-syntax provides again for a generalization over values for X in the case of (9): the internal structure of NP's is claimed to be similar to (or at least parallel to) the internal structure of V^3. The general idea is also clear. The grammatical system, the syntax, provides for schemes that direct semantic interpretation. I shall call this property of Jackendoff's syntax 'rigidity of structure on behalf of semantic interpretation'. That is, a grammar having this property can assign fixed positions to certain categories in terms of syntactic schemata available to speakers of a given language.

ROOSE (1956) claims that the first position in an NP such as (3) is semantically connected with deixis, the second position with relative quantification, the third with ordinality, the fourth with cardinality, etc. CLARK & CLARK (1977) discuss Vendler's analysis on the same matter. From the cognitive point of view, the crucial point is, I believe, whether we have to do with a cognitively determined ordering of reality settling down in syntactic structure which, in turn, determines semantic interpretation, an alternative being that cultural factors play a decisive role. At any rate, Jackendoff's scheme (9) can certainly be related to the discussion about rigidity of phrase structure: it is a mould. The area to the left of the X-projection line in (9) constitutes the specifier structure of X^3. Jackendoff stipulates that the X^1-specifier be empty; in (9) this stipulation is translated into black space. Consequently, Jackendoff has two specifier positions in X-phrases.[3]

I shall now focus on the specifier structure of Noun Phrases by analysing the way Jackendoff treats the material given in (1). He distinguishes three classes to begin with. They are given in (10).

(10) a. DEMONSTRATIVES: demonstrative pronouns, interrogative pronouns, *the*, (possibly) *a*, and (the singular) *some*.

b. QUANTIFIERS : *each, every, any, all, no, many, few, much, little*, and other uses of *some, several*, etc.;

c. NUMERALS : cardinals, *a dozen, a little*, etc.

This tripartition is based on the semantic roles played by these specifiers. The question is, of course, how (10) relates to (9).

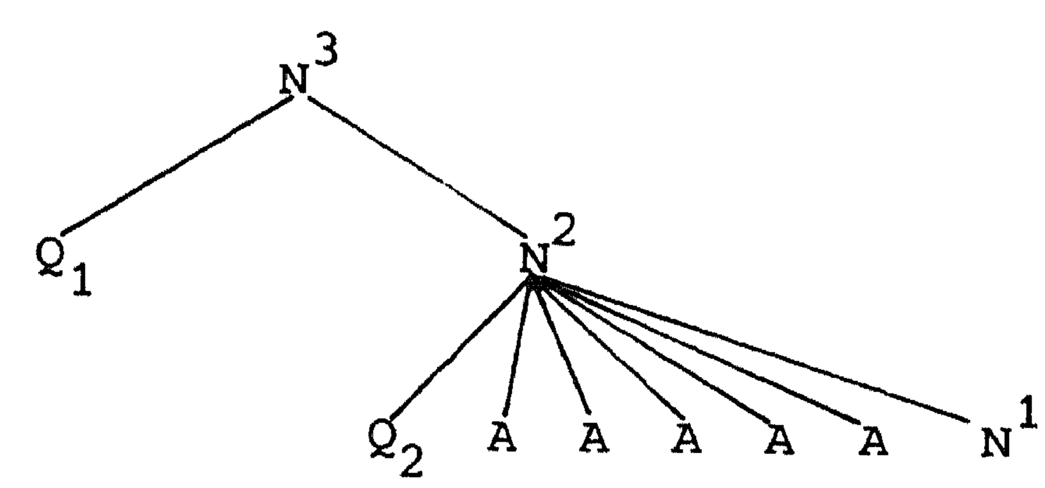
Expressions such as **Fred's all dwarfs, *some the trees, *the no dwarfs* are not well-formed in English, whereas *Fred's several attempts at writing, those few meetings we had*, etc. are well-formed. To solve this problem Jackendoff uses the normal linguistic practice of putting the possessive *Fred's*, the quantifiers *all* and *no*, and the demonstratives *the* and *no* in the same category. Thus their contrastive distribution prohibits the ill-formed examples from being generated. As a result one obtains two syntactic subcategories of the category QUANTIFIER, namely Q_1 (*each, all, no, every*, etc.) and Q_2 (*many, few, several*, etc.). In other words, the tripartition in (10) is resolved into a syntactic bipartition corresponding with the two specifier positions in (9), as shown in (11).

(11) X^3-specifiers: DEMONSTRATIVES, POSSESSIVES, Q_1
X^2-specifiers: Q_2, NUMERALS.

Note that Q_2 and NUM are also mutually exclusive: phrases like **several three trees*, **four few trees*, etc. are correctly ruled out by (11).

The Achilles' heel of this analysis can be demonstrated with the help of diagram (12).

(12)



Ignoring the occurrences of A in (12) for the moment, it can easily be seen that any combination of Q_1 and Q_2 leading to undesirable results cannot be blocked on the basis of contrastive distribution. Indeed, the fact that Jackendoff is committed by his three level hypothesis to assume just two specifier positions in NP's leads to an appeal to *ad hoc* constraints. To block **all several men*, **some few men*, etc., Jackendoff proposes the so-called Specifier Constraint. It reads as follows (1977: 104):

(13) An NP-specifier may contain at most one demonstrative, one quantifier, and one numeral.

I do not like (13) at all. It amounts to an observational statement saying that $Q_1 + Q_2$-combinations are to be blocked. Furthermore, it is redundant in that it forbids **the these children*, which is already excluded by the purely syntactic principle of contrastive distribution. I would say that Jackendoff is lured on to an arbitrary semantic constraint because his U3LH leads him to occupy two fixed positions for his specifier structure, thereby depriving him of the possibility to strictly use a pre-eminently syntactic instrument: contrastive distribution.

Should we conjecture what semantic theory underlies (9) - (13), then I think that Jackendoff connects the X^3-specifier position with deixis of some sort. That is, real deixis in the form of specific reference made by demonstratives and possessives versus possible or claimed reference in the case of Q_1-quantifiers such as *all, every,* and so on. If I say *All trees are well-formed,* I claim that I can say for each individual in the assumed universe *this one is well-formed* and *this one is well-formed*, and so on. The X^2-specifier position could be said to be connected with quantity of some sort, indicated either by measurement or by giving the cardinality of some set.

I am not sure whether this is indeed the semantic background for (11). I simply present this conjecture which seems to relate structures such as (12) to the discussion about phrases like (3), in order to give more flesh to the heel which we are considering at present.

At this point it should be said that the U3LH, though widely assumed in recent theoretical-descriptive work, has come under heavy fire. KEAN (1978) and WILLIAMS (1978) devastated the fundament for Jackendoff's claim that rules of grammar are to be formulated in terms of syntactic features.[4] I ignore this side of the matter here because these features do not play a crucial role in what I have to say against the U3LH as proposed by Jackendoff.

As far as the number of levels and the uniformity are concerned, STURM (1979) has raised some objections which I shall discuss now in some detail. Sturm's criticism is levelled against the two following properties of Jackendoff's U3LH-grammar. In the first place Jackendoff's distinction between the ten lexical categories mentioned in (8ii), each having its three-levelled projection line, generates an enormous amount of superfluous structure. For example, M^2 and M^1 never branch. The same applies to Deg^3 and Deg^1, to mention just a few categories. In the second place certain parts of phrase structure are crammed due to the fact that Jackendoff restricts himself to three levels. I shall illustrate this point now in relation to the V^1-complement and the N^2-specifier structure.

As to the complements of V, Jackendoff allows for at least five sister node positions in V^1, KOSTER (1978) for seven. Though not all these positions will be filled simultaneously, the whole approach leads to some trouble as I have shown in VERKUYL (1979): due to the fact that the Direct Object (DO) and Indirect Object (IO) are sisters of V, Koster is not able to consistently protect his structurally defined Locality Principle in terms of structurally defined auxiliary hypotheses. His only way out would be to promote the IO to a higher structural position, i.e. to a position asymmetrically c-commanding the DO, but this would require that the value for *n* in (8i) be (at least) 4.[5]

STURM (1979) rightly observes that Jackendoff is not consequent in his treatment of V-complementation. The difference between *John hit the* *nail softly of course* and **John hit the nail of course softly* (Jackendoff's judgment) is explained by saying that the "geometry of the sentence predicts that" V^3-complements (in this case, *of course*) must follow V^2-complements (in this case *softly*). However, this sort of restrictions also occurs within V^1 without its leading to a structural difference of V^i-levels. For example, the difference between *I gave my money to my friend* and *?*I gave to my friend the money* should also lead, on exactly the same grounds, to the geometrical prediction that the IO occurs on a higher level than the DO. However, Jackendoff fences his V^1-domain against geometrical structure.

As to the N^2-specifier structure, one can easily see in diagram (12) that this is also packed. Adjectives are generated as daughters of N^2: all the elements between *my* and *boxes* in (3) are sisters of N^1. To save the U3LH Jackendoff has to squeeze the adjectives into the N^2-specifier position. Note that this is a deferring strategy: to interpret N^2-structures such as in (12) requires that a syntax of some sort be given as an interpretive basis. In other words, why does Jackendoff build a syntax for DEM/POS/Q_1 and Q_2/NUM in the base component and not for the adjectives?[6]

We can summarize the second point under consideration by observing that Jackendoff's decision to stack up a lot of constituents as sisters of V and N^1 just amounts to saying that we need an auxiliary syntax for semantic interpretation, because the U3LH does not allow further branching having used up the branchings of scheme (9).

3. THE MINIMAL LEVEL HYPOTHESIS (MLH)

In reaction to the U3LH, STURM (1979) advocates the Minimal Level Hypothesis characterized by a parsimonious tenet: build as much structure as you need. In this respect Sturm strikingly links up with very interesting work from the Dutch structuralist A.W. de Groot, whose book *Structurele syntaxis* (Structural Syntax) written in the forties, can be seen as very much related to the categorial syntactic systems developed in the sixties and seventies (VERKUYL 1980).

De Groot's leading thesis in *Structurele syntaxis* with respect to phrasal structure is that phrasal elements are either co-ordinated or subordinated elements. That is, for every phrase $[_P\ X\ Y]$, (a) P is a co-ordinative construction, or (b) P is a sub-ordinative construction with either

X or Y as the head of P. Suppose that the five A's in (12) are not co-ordinated elements, then the string ...A A A A A N^1... would necessarily have the structure ...$[_{N^6}$ A $[_{N^5}$ A $[_{N^4}$ A $[_{N^3}$ A $[_{N^2}$ A $N^1]]]]]$..., where each N^i occurs as the head of N^{i+1} (see also VAN DER LUBBE 1965).

STURM (1979) is not very explicit about the formal mechanism he wants to use. The basic idea is a syntactic approach from bottom-to-top, so it seems. Sturm wavers between tree formation rules in the sense of McCAWLEY (1968) and rules of the type demonstrated in (14).

(14) $X^i \rightarrow \{X^{i-1}\ C\}$, where (a) no maximum for X^i
(b) there is always a lexical category X^0
(c) C is just one constituent occurring either to the left or to the right of its head
(d) C is either a lexical category or a grammatical formative.

It will be clear that (14) roughly expresses what would lead to a variable binary branching categorial syntax if we reformulate (14) as in (15), given the conditions (a) - (d).

(15) a. If $\alpha \in P_{X^i}$ and $\gamma \in P_C$ (where C abbreviates X^{i+1}/X^i), then

$$F_{lc}(\alpha,\gamma) = [_{X^{i+1}}\ [_C \gamma][_{X^i}\ \alpha]]$$

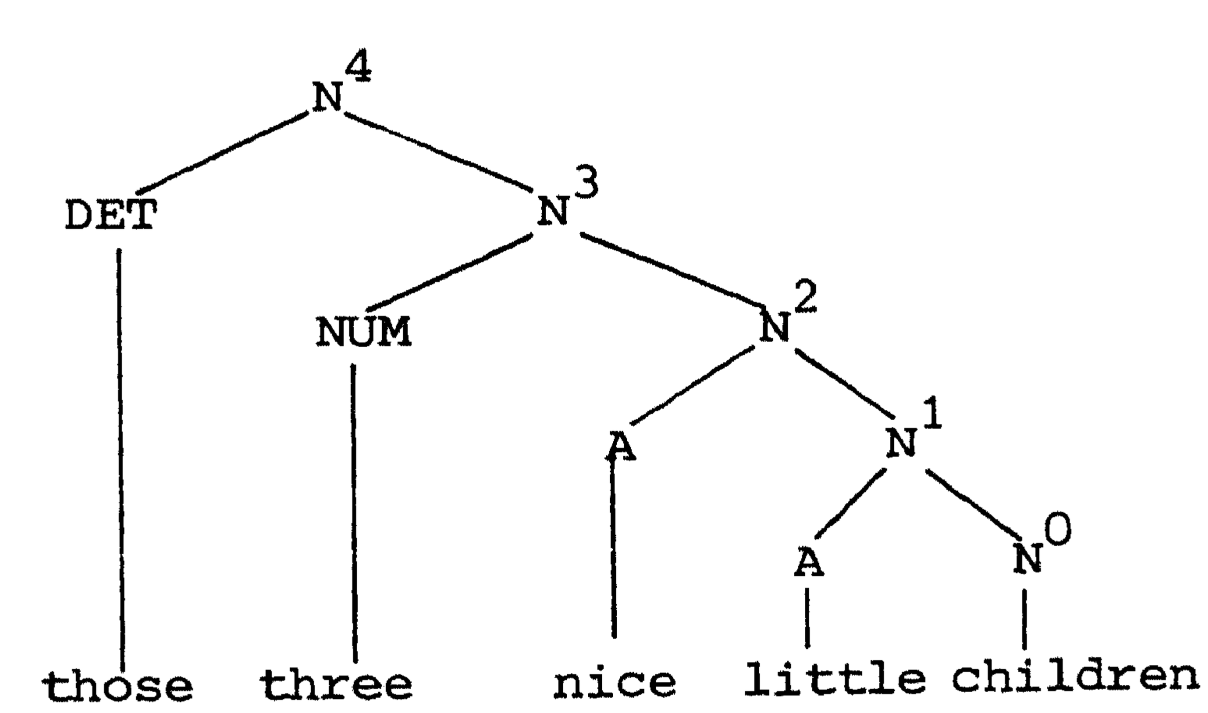
b. If $\alpha \in P_{X^i}$ and $\gamma \in P_C$ (where C abbreviates $X^i\backslash X^{i+1}$), then

$$F_{rc}(\alpha,\gamma) = [_{X^{i+1}}\ [_{X^i}\ \alpha][_C \gamma]]$$

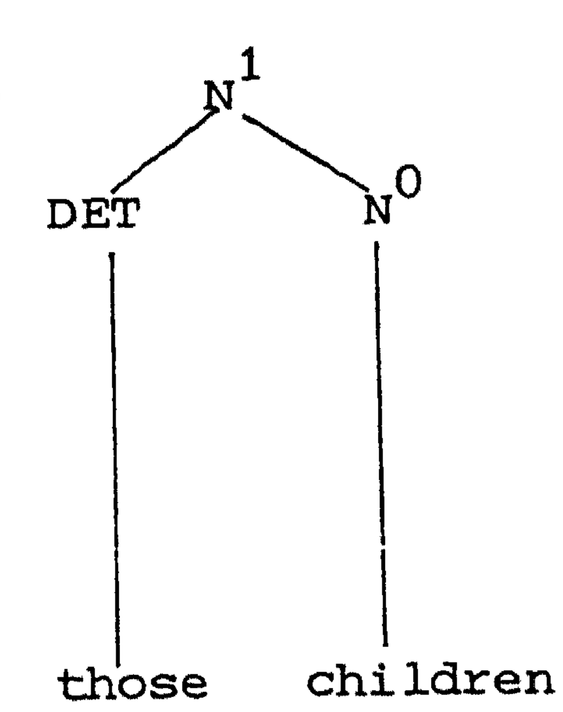
(cf. BARTSCH & VENNEMANN 1972; DAHL 1977; BACH 1979 among others).

$\bar{X}$-structures generated with the help of (15) are shown in (16).

(16a)



(16b)



Translating the representations of structuralists like De Groot and Van der Lubbe into tree diagrams would give us structures such as (16). Transformational linguists cannot be very much disturbed either (Cf. CULICOVER 1977; HALITSKY 1975; HORNSTEIN 1977.) Thus there appears to be some reason to pursue the investigation of the MLH along the lines of De Groot - that is, in a formalized version in terms of (15). Note that the $\bar{X}$-syntax is to be taken as explicitly defining the notion 'head of a phrase' (= X^0). As the head of a phrase is the most deeply embedded element on a projection line, the bottom-to-top approach inherent to (15) seems to be a rather natural mechanism for generating phrases like (16a) and (16b).

Now, there are two approaches to the strengthening of the MLH. The first would argue, for instance, that syntactically spoken a structure like (16a) is to be preferred to (12). Gapping (either taken as a syntactic rule or taken as an interpretive rule) would require that $A+N^i$ be a constituent at each level on the projection line in view of (Dutch) examples such as (17).

(17) Ik houd van grote *snelle Franse auto's* en mijn broer van kleine.
lit: I love big *fast French cars* and my brother small.

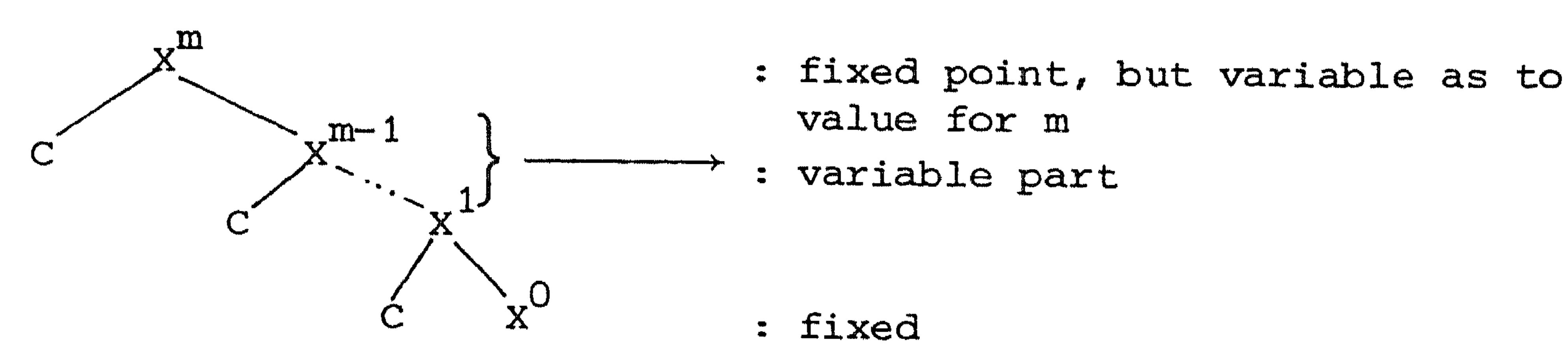
Ik houd van grote snelle *Franse auto's* en hij van grote dure.
lit: I love big fast *French cars* and he big posh.

Figure (16a) satisfies this requirement as opposed to structures such as (12). The corresponding English phenomenon is sometimes analyzed in terms of the so-called *One-substitution*: the pronoun *one* substitutes for the italicized phrases in (17) replacing constituents (cf. CULICOVER 1977: 183-6).

In the remainder of this paper I shall, however, follow a different approach by discussing some features of the MLH with one eye on its usefulness for linguistic analysis and the other eye on its possible contribution to existing binary categorial systems such as in Montague's PTQ and related work.

The first point to be stressed is the flexibility of the system. Since there is no fixed upperbound, we have to assume an X^m-node whose numerical value is variable relative to the structure dominated by this node. In (16a) m=4, in (16b) m=2. The top node X^m can be called 'X maximal'. It must be observed that X^m, though variable as to its projection level, is also fixed in the sense that we always have an X^m. On the other hand, we always know for sure that phrase structure is built up from X^0, the anchorage of the projection. The general idea can most easily be captured by (18).

(18)



STURM (1979) rightly points out that 'X maximal' can be used in the lexical specification without any difficulty at all: it is not necessary to know the numerical value for m; it is sufficient to know that we have to do with the highest node on a projection line. Hence it is possible to characterize the article *the* as in (19a), i.e. as taking the N^{m-1} to form an N^m.

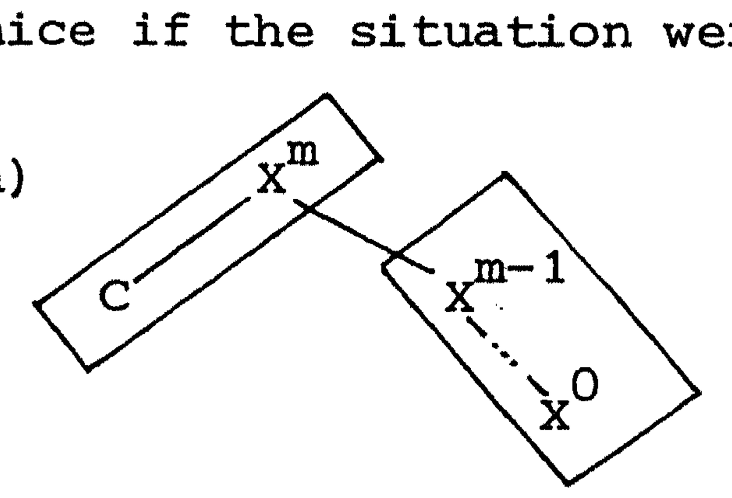
(19a) *the*, +DET, [+ —— N^{m-1}] (19b) ...,[+ —— X^i] ($0 \leq i \leq m-1$).

As to the variable part of (18) we could exploit the structural similarity of adjectives and adverbials. Both categories can be treated as instances of a category having the subcategorization frame (19b). In the case of adjectives the value for X is N, otherwise X is V or A (possibly other categories as well). In this way one can capture a well-known generalization in the formalism.

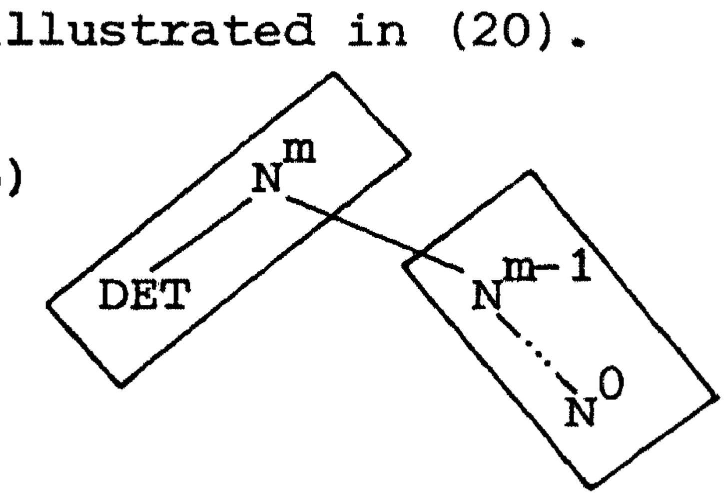
In the present treatment of $\bar{X}$-syntax the notion of projection line is more important than in the Jackendoff version, where structural parallelism is the crucial feature. We can illustrate the difference with the help of (19b): in the MLH adjectives and adverbials both take nodes in the variable part of the projection line, whereas the U3LH cannot account for the parallelism in terms of common behaviour with respect to the same level.[7]

The second point to be raised is the status of the partly fixed top position in (18). Can we use it for a uniquely determined semantic operation, at least in languages such as English and Dutch? Put more generally, it would be nice if the situation were so as illustrated in (20).

(20a)



(20b)



In other words, the constituent C immediately dominated by X^m could be taken as a function operating on X^{m-1} to yield X^m, where X^m and X^{m-1} crucially differ as to their categorial status (and correspondingly to their intensional type), whereas transitions from X^0 up to X^{m-1} keep the categorial status of X^0 constant.[8] Note that in (16b) N^0 would be N^{m-1}. As far as I can see (20a) would also apply to verb phrases and prepositional phrases. For example, Aux could be analyzed as an element changing the V-projection line into a construct of a crucially different nature.

Whatever the generality of (20a) may be, its basic idea seems to apply to (20b). I shall try to show that by following two lines. The first is plotted out in the Chomskyan framework, the second in the Montague framework. My wish is to connect these lines with the help of the MLH.

To begin with, I refer to Chomsky's lecture 'Questions on Form and Interpretation' in which he argues against a one-to-one correspondence between syntax and semantics suggested by Barbara Partee's analysis of restrictive modification. Chomsky argues that the definite article should be taken as a universal quantifier. In a sentence like *The book we ordered arrived* Chomsky considers the definite article as an element determining that all members of a unit class arrived. In *The books we ordered arrived* the article *the* "determines that all members of a class of cardinality greater than or equal to 2 arrived". Thus, he continues, "[± def] corresponds to universal-versus-existential quantification" (CHOMSKY 1977, pp.50-51).

Though I think that Chomsky's argument against Partee is in itself not very convincing, his interpretation of [± def] is very interesting.[9] In MILSARK (1977) we find an extensive analysis of this feature based on that interpretation. I shall discuss it in some detail with the sole purpose of reaching the conclusion that all specifiers treated so far are to be located in the DET-position of (20b). Consider again the sentences in (2).

(2) a. There is a child in the house.
    b. *There is the child in the house.

The opposition between definite and indefinite NP's in sentences like (2) is a much discussed topic. In the sixties the so-called *There-insertion* transformation was proposed to account for the relation between (2a) and the sentence *A child was in the house*. The transformation was constrained so as to exclude (2b) by requiring that the NP to be moved to the position after the copula be [-def]. However, quantifiers like *all, every, each,* etc. cannot occur in sentences such as (2) either, whereas *several, many,* etc. can (Cf. KRAAK & KLOOSTER 1968; MILSARK 1977; GUÉRON 1976, among others):

(21) *There were all children in the house.
*There appeared both elephants in the circus.
*There was every child in the house.
There were several (many, few) children in the house.

Milsark - following or preceding Chomsky, I am not certain which is the case - proposed that [+def] be interpreted as 'universal quantification', whereas [-def] should be taken as 'existential quantification', thus extending the coverage of both features so as to include the Q-quantifiers. Consequently, the ill-formedness of the sentences in (21) is accounted for by the feature [+def]. GUÉRON (1976) noticed that Extraposition from NP also interacts with this feature as shown by the opposition between **Those three books have just come out by Christie* (blocked by [+def]) and *Several books have just come out by Christie* (not blocked on account of the absence of [+def]).

Though it is very clear that a lot of factors are involved complicating the issue considerably, the clear-cut distinction between two classes of NP-specifiers on the basis of the features [+def] and [-def] appears to solve a lot of descriptive problems when applied to the internal noun phrase structure.[10] So let me give the resulting bipartition and see how it takes effect.

(22)

+def	-def
the DEMONSTRATIVES POSSESSIVES, etc. *each, all, every, any*, etc.	*a* *some* *few, several, many*, etc. *two, three, four*, etc.

Let us assume that (22) is an organized list of all members of one and the same syntactic category DET, which takes an N^{m-1} to yield an N^{m}. Then DET has two subcategories, say [+def] and [-def], just as the category NP has as its subcategories proper names, pronouns and full NP's, mutually excluding each other.

On the basis of this assumption a lot of the ill-formed constructions in (1) are automatically blocked. **Some the trees*, **few three children*, **three these children*, **the some trees*, etc. are now excluded on exactly the same ground on which **He the man is walking* is excluded, namely on the basis of contrastive distribution. Due to the restriction that all members of (22) can only occur in one syntactic position, namely the DET-position, these facts follow. As a natural consequence the Specifier Constraint (13) turns out to be superfluous: **all several children*, **any much wine*, **some many trees*, etc. are ruled out on syntactic grounds. Hence the analysis leading to (22) should be preferred to Jackendoff's analysis leading to (11) and (13).

In following this line I have pushed aside several stumbling-blocks on my way to the conjunction where the Montague-line comes in. So my strategy will be to assure that the two lines meet, to be positive about that circumstance, and to show that the advantages outweigh the problems that arise. As a result I shall modify (22) in Section 5 from a different angle.

Condensing the Chomskyan line followed from (20b) up to (22) to its essence one can say that it brings out the 'only one DET-position hypothesis', which says that the top of an N-projection line is characterized by the unique operation at the N^{m-1}-level changing a common noun constituent N^i $(0 \leq i \leq m-1)$ into a noun phrase N^m. Before going more deeply into some of the predictions of this hypothesis, I shall first discuss the Montague approach to determiners in Section 4.

4. DETERMINERS IN THE PTQ-FRAMEWORK

In Montague's PTQ DET would be taken as an abbreviatory notation for the category T/CN, i.e. as a derived category which takes a common noun (CN) to form a term (T). We do not find DET in the lexicon. Montague introduces specifiers such as *the, all, every,* etc. syncategorematically. For example, the article *the* is introduced by the syntactic rule (23).

(23) If $\zeta \in P_{CN}$, then $F_1(\zeta) \in P_T$, where $F_1(\zeta) =$ *the* ζ.

In other words, if ζ is a common noun, then the F_1-operation gives the NP *the* ζ. BENNETT (1975) extended the material presented in PTQ considerably by stating syntactic rules for all the determiners mentioned in (22). By his treatment of the plural he is forced to split up (23) into one rule accounting for *the* occurring with a singular CN and rules accounting for *the* taking a plural CN. Altogether Bennett needs about seventy F-rules to account for less than twenty specifiers.

Corresponding to (23) a translation rule is supposed to operate as shown in (24).

(24) If $\zeta \in P_{CN}$, and ζ translates into ζ', then *the* ζ translates into

$$\lambda P \exists y[\forall x[\zeta'(x) \leftrightarrow x = y] \wedge P\{y\}]$$

where P is the predicate symbol whose place will be occupied by the intransitive verb phrase taken by the term *the* ζ to form a sentence. Again about seventy rules are necessary to account for the specifiers.

The normal reaction of linguists to this sort of treatment is a feeling of repugnance for the use of so many rules, the overall impression being that the system of syncategorematic rules cannot account for the many syntactic correspondences among determiners.[11] However, this feeling should not become a licence for doing away with Montague grammar, since it is easy to modify the organization of the PTQ-framework such that (23) and (24) fit into the linguistic standard mode of organizing a grammar (cf. COOPER & PARSONS 1976, and HAUSSER 1976). Rather than having syncategorematic rules such as (23) and (24), one could apply a rule operating on members of the category DET (i.e. T/CN) and on CN-expressions, as shown in (25). Applied to (16) γ would be N^3 in the case of (16a), and N^0 in the case of (16b).

(25) If $\delta \in P_{T/CN}$ and $\gamma \in P_{CN}$, then $F_{conc}(\delta,\gamma) \in P_T$,

where $F_{conc}(\delta,\gamma) = \delta\gamma$.

Correspondingly, the rule for the translation would be (26).

(26) If $\delta \in P_{T/CN}$ and $\gamma \in P_{CN}$, γ translates into γ',

then $F_{conc}(\delta,\gamma)$ translates into $\delta'({}^\wedge\gamma')$.

As a result the lexical entry for *the* - assuming the correctness of the translation in (24) - would read as (27).

(27) *the*,DET,...,<<s,<<s,e>, t>>, <<s, <<s,e>, t>>,t>>

$$\lambda Q \lambda P \exists y[\forall x[Q\{x\} \leftrightarrow x = y] \wedge P\{y\}]$$

where DET abbreviates (t/IV)/(t//e), i.e. T/CN. Accordingly, *every* would receive an entry such as (28).

(28) *every*,DET,...,<<s,<<s,e>,t>>,<<s,<<s,e>,t>>,t>>

$$\lambda Q\lambda P\forall x[Q\{x\} \rightarrow P\{x\}]$$

given Montague's analysis of *every*.

Summarizing, one can observe that the 'only one DET-position hypothesis' can easily be accounted for in the PTQ-framework by casting rules such as (23) and (24) into a different mould. In the next section I shall discuss some of the predictions of this hypothesis with the help of the framework presupposed by (25) - (28).

5. DETERMINERS, ADJECTIVES AND NUMERALS

In Section 3 I have argued that the MLH-approach of $\bar{X}$-syntax leads to a restrictive and very natural hypothesis about the determiner of an NP: the 'only one DET-position hypothesis' (OODH). In Section 4 I have tried to make clear that the OODH perfectly fits into the PTQ-framework, given a slight conceptual reorganization of the grammar. In the present section I shall confront the OODH with two apparently problematic areas in the specifier structure of NP's. The first one is pre-determiner position, the second is the status of numerals.

It is a fact of English and Dutch that determiners can be preceded by modifying elements. Moreover, some members of (22) occur as specifiers together, one preceding the other. The relevant material is given in (29).

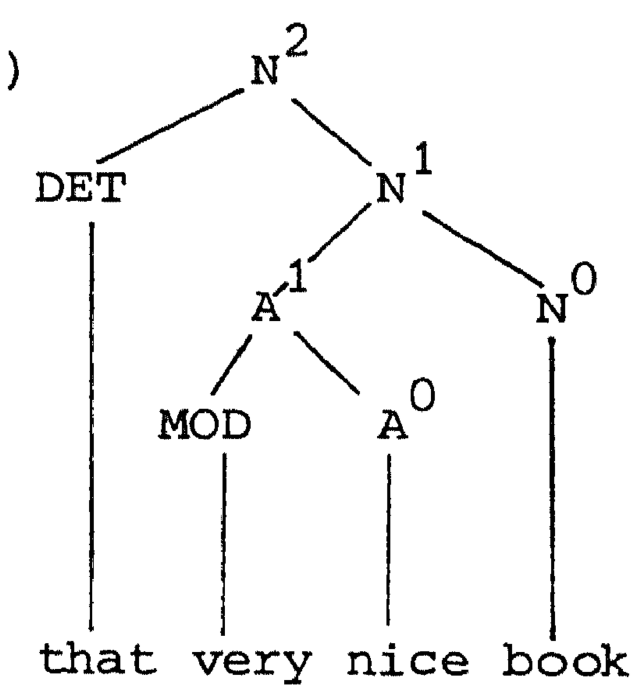
(29)
a. almost every child
b. almost all the children
c. all the/my children
d. *the/my all children
e. nearly all children
f. so very many interesting problems

Extending the range of the descriptive domain, I shall first focus on the internal structure of the italicized constituents in (30) before going further into the problems raised by (29).

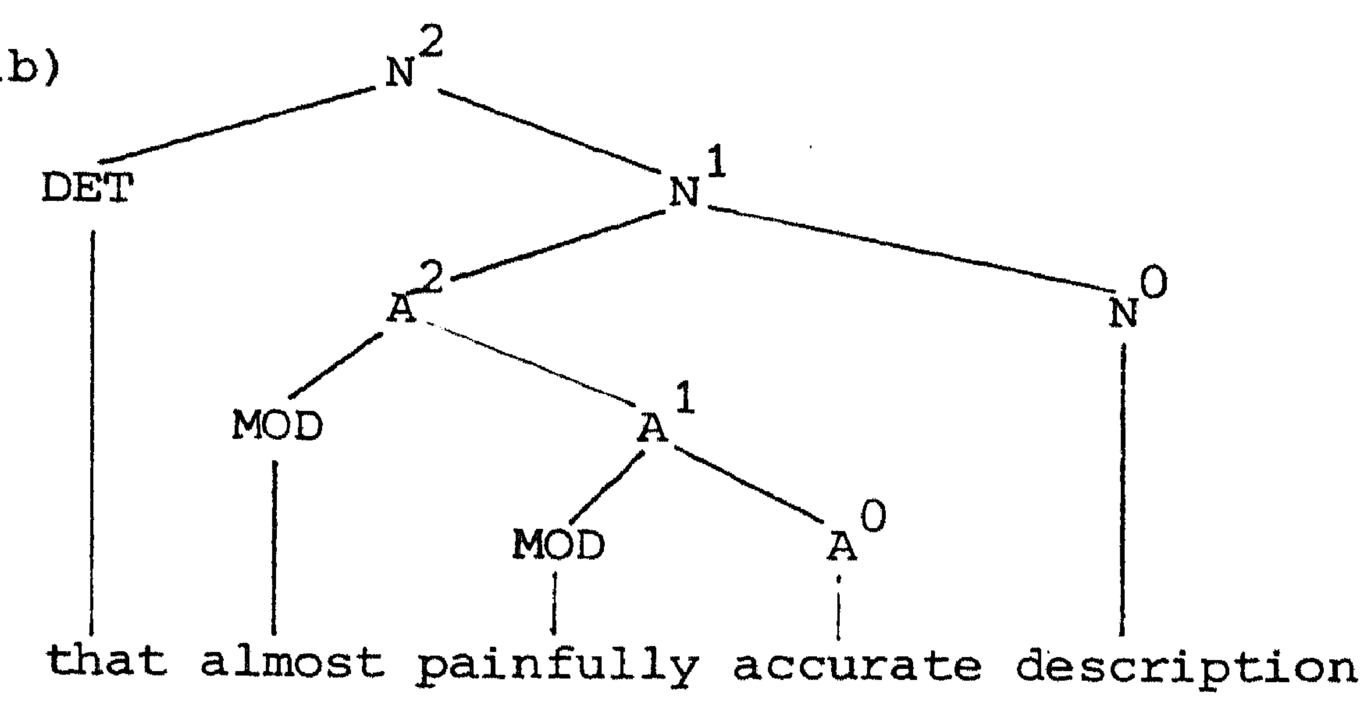
(30)
a. that *very nice* book
b. that *so particularly* nice book
c. that *almost painfully accurate* description

The MLH-hypothesis states that (30a) be analyzed as (31a) and (30c) as (31b).

(31a)



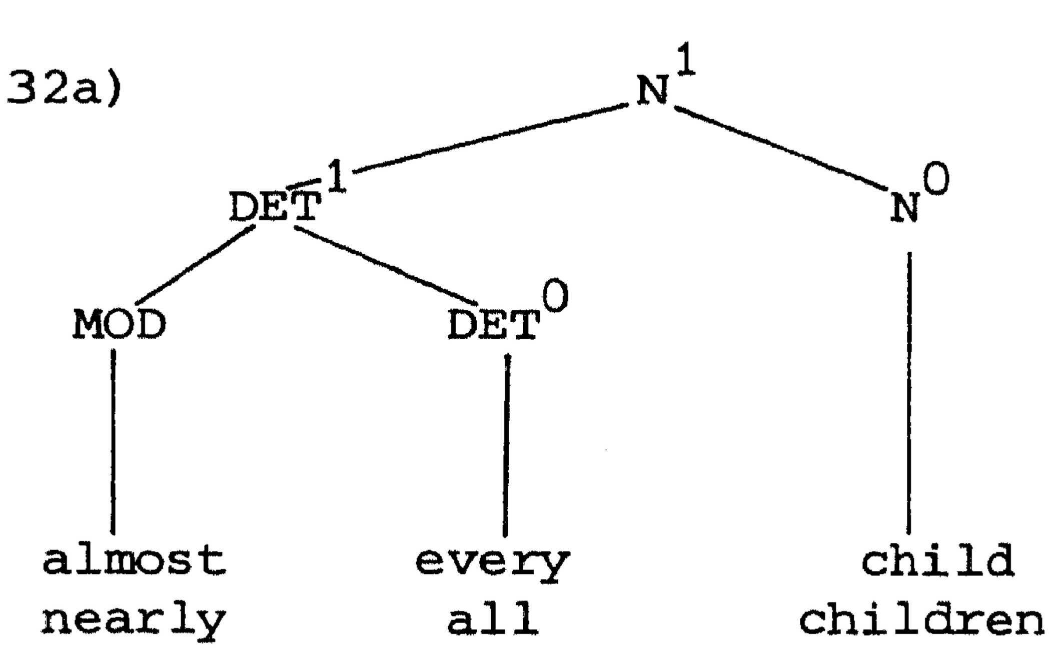
(31b)



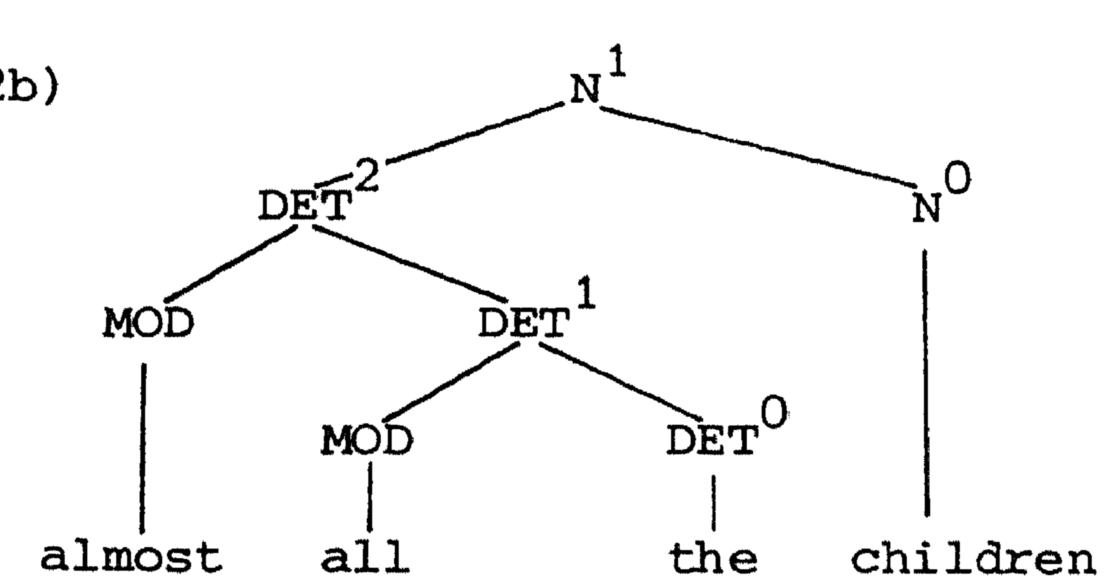
The (adverbial) modifiers are represented here by the label MOD, whose exact $\bar{X}$-status I shall ignore here. The modifiers *almost, nearly, so* and *very* clearly belong to MOD: they can all take adjectives.

Focussing now on (29a) one can say that the OODH forces us into the position of analyzing *almost every* as consisting of a C modifying an X^i, given the MLH-rules (15). The data in (30) strongly suggest that it is *every* rather than *almost* that counts as the head of the construction. In other words, we must accept DET as a possible value for X in the $\bar{X}$-syntax. Following Chomsky, the MLH-approach very reluctantly allows of X-categories as contrasted with the U3LH.[12] This strategic attitude seems to pay off. Consider the diagrams in (32).

(32a)



(32b)



Comparing the DET-structures of (32) with the A-structures of (31) we can observe that there is a structural parallelism. That is, the DET-projection line and the A-projection line seem to share certain properties with respect to their modifiers. So we must ask ourselves whether DET and A have properties in common. An attractive answer is that they have, from the historical point of view. The Old-English predecessors of *every, each,* etc. are often considered adjectives in the traditional literature.

I do not say that determiners *are* adjectives. I merely say that they share certain structural properties. Given the fact that DET is an improductive category, whereas A is productive, one could maintain that DET is a "frozen" adjective, that is, originally an adjective, but having acquired more and more specific properties distinguishing it from real adjectives. DET is frozen in the sense that its property of taking modifiers dates back from the period in which it was an adjective. English and Dutch show the same development in this respect. One might say that both languages have developed such that the N^{m-1} modifier got its specific function of forming a CN into a term (cf. LIGHTFOOT 1979, pp.167-186). An interesting problem arising from this analysis is that constituents such as *almost* and *nearly* in (31b) belong to the category CN/CN, whereas they belong to DET/DET in the case of (32). The question is how semantics accounts for categorial transitions.

Summarizing - and aware of some speculative elements in the above paragraphs - I would say that the OODH entails a DET-projection line, which means that DET is a value for X. Synchronically we capture the structural parallelism with respect to the A-projection line. Observing that the range of modification of DET as well as the number of members of the category DET are very much restricted, we turn back to history. Diachronically seen DET belonged to A. Thus the MLH-approach accounts for the present situation in which DET differs from A in certain respects as well as for the correspondences that remain. The U3LH-approach cannot give such an account on the basis of predictions commanded by (8).

There is an apparent problem with (29b) that we cannot leave out of consideration. In the diagram (32b) *all* occurs as the modifier of *the*. So a distinction is made between a DET *all* and a MOD *all*. Comparing the corresponding Dutch data with the examples in (2), we can easily observe that this distinction comes out in Dutch at the morphological level, as shown in (33).

(33) a. al de kinderen
all the children

b. * alle de kinderen

c. *al kinderen

d. alle kinderen
all children

Example (33b) shows that *alle* cannot modify the definite article as contrasted with *al* in (33a). Dutch makes a distinction between the determiner *alle* occurring in (33d) and the modifier *al*. Note that the modifier does not occur in an indefinite NP as shown in (33c). Again the historical development of the language under analysis can be taken into account. *Alle* results from a fusion of *al* and the definite article. In this sense, *alle kinderen* and *al de kinderen* are to be considered variants. It seems justified to say that they are variants, from the synchronic point of view: there is hardly any semantic difference between the two phrases.

The Dutch data suggest that the English *all* belongs to two different syntactic categories: all_1 occurring in (29b/d/e) must be considered a determiner, whereas all_2 in (29c) is to be taken as a modifier.

The second problematic area with respect to the predictions made by the OODH is shown in (34) and (35).

(34) a. I saw these three children
b. I saw three children
c. *I saw three these children.

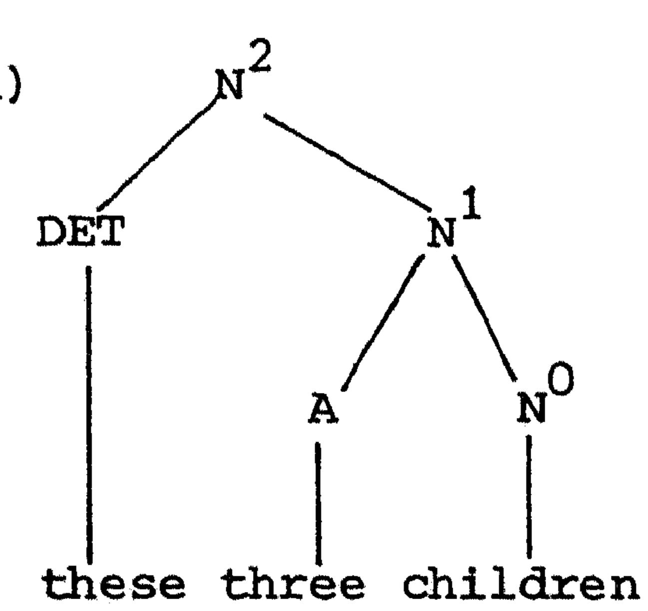
(35) a. I heard about these few attempts to escape
b. I heard about few attempts to escape
c. *I heard about few these attempts.

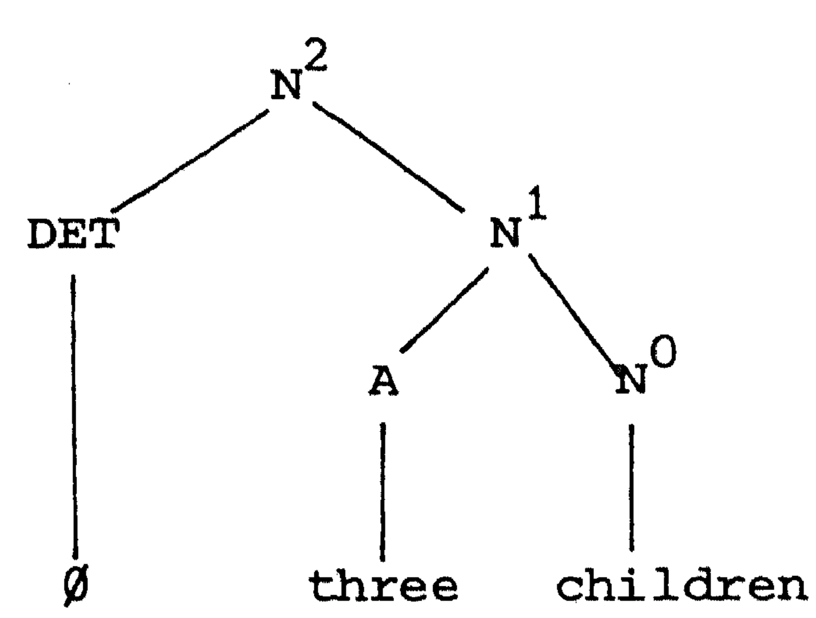
Recall that the U3LH excludes the c-sentences by requiring that *three* and *few* be rewritten on a lower level of phrase structure than *these*. That is, in the lexicon *three* and *few* are syntactically characterized such that they can be inserted only in the N^2-specifier positions, *three* as a Noun and *few* as a Q_2. JACKENDOFF (1977, pp.128-134) defends the position that *three* is a Noun with the argument that numerals cannot be preceded by degree specifiers such as *so, too, how,* etc. Furthermore there are constructions such as *a beautiful two weeks, a dusty four miles,* etc., suggesting that numerals behave like nouns with respect to the specifiers they can occur with. As a result *three children* is to be derived from a structure corresponding to *a six of weeks* generated by the base component. Two local transformational rules are necessary to delete the *a* and the *of*; both are obligatory.

Apart from the weakness of the two arguments cited above (e.g. the determiner *all* can be modified by *nearly*, as contrasted with the derminer *some*; nevertheless, Jackendoff puts *all* and *some* in the same category) and of the other arguments given by Jackendoff, the underlying structure seems highly unnatural in view of constructions such as *approximately twenty books, nearly forty children,* etc. To derive *these almost twenty hits* from *these almost a twenty of hits* appears to me artificial and ill-motivated.

Though the MLH certainly allows an analysis where *three* can be taken as a Noun, I would like to explore the position where numerals are analyzed as taking an N^i to form an N^{i+1}, because I think that such an analysis might contribute to the solution raised in connection with phrases such as (3). That is, by following this line of argumentation I hope to be able to account for the difference between *three nice books* and **nice three books*; more generally, for the principles determining the order of specifiers and adjectival constituents with respect to each other in the pre-nominal position.

Assuming that *these* in (34a) is a DET, *three* must occur at a lower level of phrase structure. Let us provisionally label *three* as an adjective, thus expelling numerals from (22). BARTSCH (1973) defends this position by assuming that numerals belong to the category of plural adjectives. BENNETT (1975, p.132) observes that phrases like *the few gods, the many gods,* and *the twelve gods* "function much like[...] occurrences of adjectives", without being specific about their syntax. As said Bennett introduces these specifiers syncategorematically, which means that *the twelve* is introduces as a whole. The MLH-syntax would represent (34a) as in (38a).[13]

(38a) 

(38b) 

What does a sentence like (34a) *I saw these three children* mean? A set-theoretical basis for an answer to that question seems appropriate. The N^1 can be said to refer to those subsets of the power set *child* that contain three members. (This power set, being the set of all subsets, contains subsets consisting of one member, two members, three members, and so on.) The determiner *these* can be analysed as identifying a particular three-membered set in the set referred to by N^1. Type-logical differences can be applied to guarantee the right combinations, as I shall show below. Going a little bit further into the nature of *these* one can say that this determiner introduces a certain group deictically. I shall symbolize this sort of introduction as $\Theta A[\ldots A\ldots]$, where A refers to a set, and where we can say that (39) holds. ($\exists!A$ meaning 'there is a unique A):

(39) $\Theta A[\alpha(A)] \leftrightarrow \exists! A$ such that A is deictically or contextually or anaphorically given and α is true of A.

The indefinite determiner $\emptyset$ would have an existential quantifier in the position where the definite determiner has the Θ-quantifier: the NP *three children* can be analyzed as introducing the existence of a certain subset in the power set *child*.

The above semantic analysis of the two phrases under consideration closely ties up with work done by BARTSCH (1973) and BENNETT (1975). However, there appear to be three problems with the Bartsch-Bennett position that result from trying to reorganize the current PTQ-framework in terms of the MLH-$\bar{X}$-syntax as in Section 4. Firstly, singular determiners have plural counterparts introduced by different rules. In a Bennett-approach which allows more than seventy syncategorematic rules to describe the specifiers under analysis, this is not felt as an objection. However, in the MLH-approach where we want to have lexical entries it is. We do not want to have an entry for the singular *the* and several entries for the plural *the* if this can be avoided. The same applies to the other specifiers. Secondly, the Bartsch-Bennett position makes it necessary to distinguish between plural and singular predicates and also between plural and singular adjectives. Thirdly, the distributive and collective readings of sentences like (34a) and (34b) cannot be accounted for in a proper way. My intention is to propose a solution to the first two problems thus providing a basis for the solution of the third problem.

Both Bartsch and Bennett introduce PLUR as operators on a singular noun, though in a slightly different way. I believe that this is the root of many troubles, among which the duplication of so many rules. I shall assume that PLUR and SING are both operators of the same level providing for the cardinality of the set to which the N^0 refers. Moreover, I shall assume that PLUR and SING are numerals. In other words, in (38a) PLUR does not appear at all: it is the numeral *three* that provides for the plurality of the noun phrase. Only if numerals like *one, two, three,* etc. are absent, can SING and PLUR occur in the syntactic structure of NP's. I shall show that these assumptions lead to a more satisfactory analysis of the NP-structure than in the literature mentioned here.

A first sight my position would entail that PLUR and SING are to be taken as adjectives, just like numerals. If, for expository purposes, we leave out intensions, this would mean that they belong to the category CN/CN; that is, they would have to be considered as functions operating on properties (of individuals) to yield properties (of individuals). In other words, they would belong to the category $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$. Given our analysis of the plural NP's in (38) this will not do, because PLUR should be taken as a function from properties (of individuals) to properties of properties, that is, as belonging to $\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$. Therefore, I shall propose the following three entries for SING, PLUR and *three* respectively.[14]

(40) a. SING, +NUM, ..., $\langle\langle e,t\rangle, \langle\langle e,t\rangle,t\rangle\rangle$

$$\lambda Q\lambda P[\forall x[P(x) \rightarrow Q(x)] \wedge \#(P) = 1]$$

b. PLUR, +NUM, ..., $\langle\langle e,t\rangle, \langle\langle e,t\rangle,t\rangle\rangle$

$$\lambda Q\lambda P[\forall x[P(x) \rightarrow Q(x)] \wedge \#(P) \geq 2]$$

c. *three*, +NUM, ..., $\langle\langle e,t\rangle, \langle\langle e,t\rangle,t\rangle\rangle$

$$\lambda Q\lambda P[\forall x[P(x) \rightarrow Q(x)] \wedge \#(P) = 3].$$

The symbol # is to be taken as an operator yielding the cardinality of the set referred to by P.

One of the consequences of this proposal is that we have to analyse a phrase like *nice apples* in *I bought nice apples* as [Ø[PLUR[*nice*[*apple'*]]]], and *nice apple* in *I bought a nice apple* as [*a*[SING[*nice*[*apple'*]]]], where *apple'* represents the form of the N^0 being neutral between the plural and the singular form.

I shall now first demonstrate the proposal in some detail with the help of the derivations of (38a) and (38b) before discussing some of its other consequences. In (41) the different types of the elements in (38) are given. NUM replaces the label A.

(41)

N^0 :	$\langle e,t\rangle$	– CN
NUM:	$\langle\langle e,t\rangle, \langle\langle e,t\rangle,t\rangle\rangle$	– (t/CN)/CN
N^1:	$\langle\langle e,t\rangle,t\rangle$	– t/CN
DET:	$\langle\langle\langle e,t\rangle,t\rangle, \langle\langle\langle e,t\rangle,t\rangle,t\rangle\rangle$	– T/(t/CN)
N^2 :	$\langle\langle\langle e,t\rangle,t\rangle,t\rangle$	– T

The complete derivation from bottom to top of the noun phrase *these three* *children* in (34a) on the basis of its syntactic structure (38a) - given the replacement of A by NUM as indicated in (41) - is shown in (42).

(42) N^0 : *child'* $\rightsquigarrow$ child'

N^1 : *three child'* $\rightsquigarrow$ $\lambda Q\lambda P[\forall x[P(x) \rightarrow Q(x)] \wedge \#(P) = 3]$ (child')

$\Leftrightarrow$ $\lambda P[\forall x[P(x) \rightarrow \text{child}'(x)] \wedge \#(P) = 3]$

DET: *these* $\rightsquigarrow$ $\lambda \mathcal{Q}\lambda \mathcal{P}\Theta A[\mathcal{Q}(A) \wedge \mathcal{P}(A)]$

N^2 : *these three children*

$\rightsquigarrow \lambda \mathcal{P}\Theta A[\lambda P[\forall x[P(x) \rightarrow \text{child}'(x)] \wedge \#(P) = 3](A) \wedge \mathcal{P}(A)]$

$\Leftrightarrow \lambda \mathcal{P}\Theta A[[\forall x[A(x) \rightarrow \text{child}'(x)] \wedge \#(A) = 3] \wedge \mathcal{P}(A)]$.

Note that DET correctly introduces predicates of predicates. Assuming that the plural *the* does not differ *semantically* from the singular *the,* I shall assign *this* the same semantic representation as *these* in (42). Consequently, *this child* will be represented as in (43).

(43) $\lambda \mathcal{P}\Theta A[[\forall x[A(x) \rightarrow \text{child}'(x)] \wedge \#(A) = 1] \wedge \mathcal{P}(A)]$.

I think that (43) is a precise formalization of Chomsky's suggestion to analyze definite specifiers such as *the* and *these* in terms of universal quantification over a set containing just one member in the case of singular NP's.

Turning now to (38b) we can give the bottom-to-top derivation of the NP *three children* in (34b):

(44) N^0 : as in (42)

N^1 : as in (42)

DET: $\emptyset \rightsquigarrow \lambda \mathcal{Q}\lambda \mathcal{P}\exists A[\mathcal{Q}(A) \wedge \mathcal{P}(A)]$

N^2 : *three children* $\rightsquigarrow \lambda \mathcal{P}\exists A[\lambda P[\forall x[P(x) \rightarrow \text{child}'(x)] \wedge \#(P)=3](A) \wedge \mathcal{P}(A)]$

$\Leftrightarrow \lambda \mathcal{P}\exists A[[\forall x[A(x) \rightarrow \text{child}'(x)] \wedge \#(A) = 3] \wedge \mathcal{P}(A)]$.

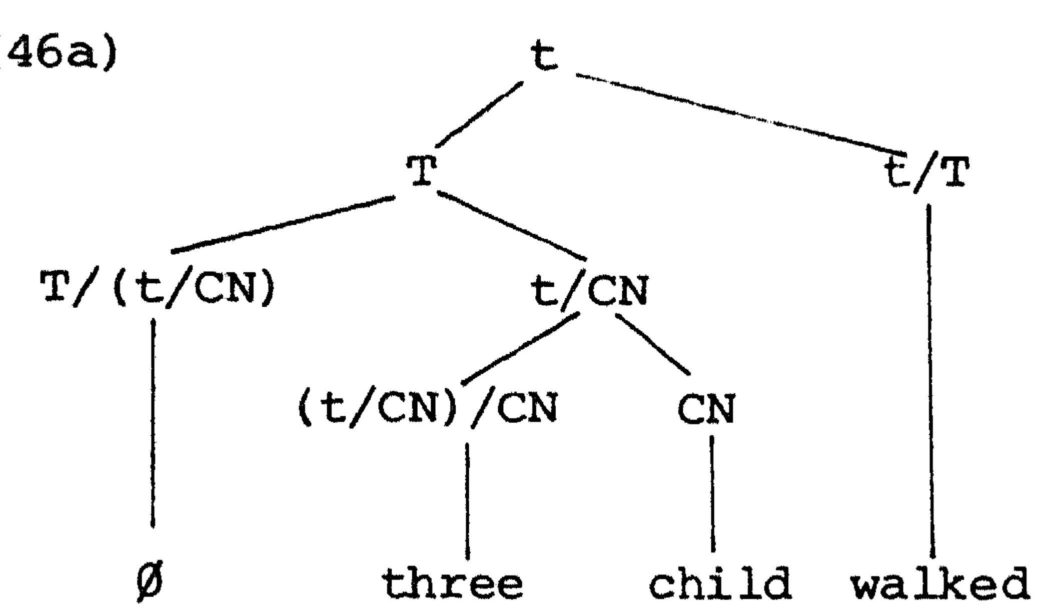
The derivations (42) and (44) show that it is possible to combine Montague-grammar with Chomsky-syntax. Note that (38a) and (38b) are not analysis trees in the PTQ-sense. They are phrase structural configurations. The combinatorial mechanism involved in (42) and (44) is in conformity with Frege's principle.

The same applies to the mechanism involved in the derivation of sentences such as (45):

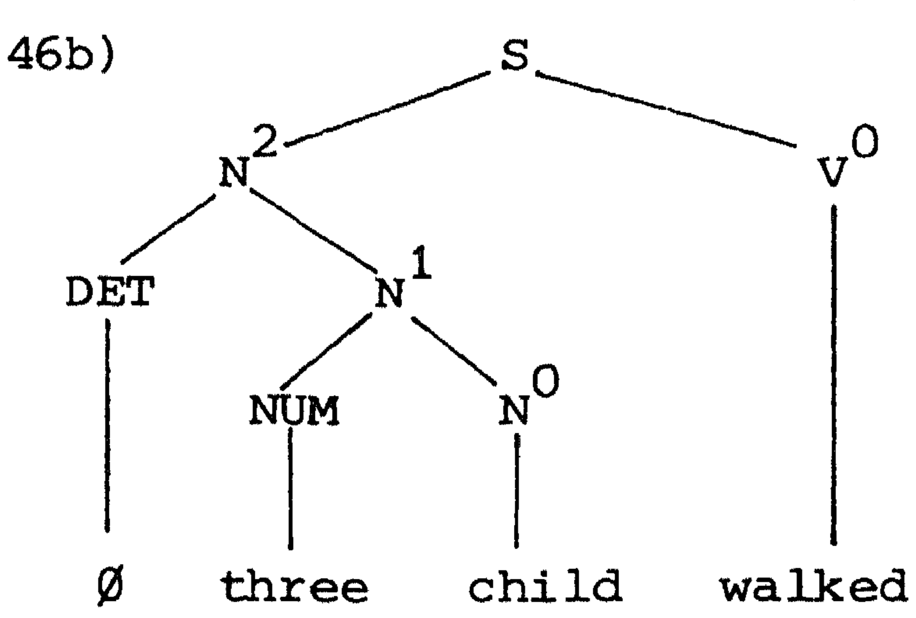
(45) Three children walked.

It will be understood that the standard treatment of *walk* (I ignore tense here) must be adapted to the present analysis of NP's. That is, *walk* must be considered to belong to the type <<e,t>,t> rather than to the type <e,t>. I shall now give the complete derivation of (45) to show that no problems arise. The categorial tree of (45) is given in (46); in (46a) in Montague's notation, in (46b) in Chomsky's notation.

(46a)



(46b)



The semantic derivation corresponding to (46) is given in (47), where I start giving the N^2-representation ending (44).

(47) N^2: $\lambda\mathcal{P}\exists A[[\forall x[A(x) \rightarrow child'(x)] \wedge \#(A)=3] \wedge \mathcal{P}(A)]$

V^0: $\lambda P\forall x[P(x) \rightarrow walk'(x)]$

S : $\lambda\mathcal{P}\exists A[[\forall x[A(x) \rightarrow child'(x)] \wedge \#(A)=3] \wedge \mathcal{P}(A)]\ (\lambda P\forall x[P(x) \rightarrow walk'(x)])$

$\leftrightarrow \exists A[[\forall x[A(x) \rightarrow child'(x)] \wedge \#(A)=3] \wedge \lambda P\forall x[P(x) \rightarrow walk'(x)](A)]$

$\leftrightarrow \exists A[[\forall x[A(x) \rightarrow child'(x)] \wedge \#(A)=3] \wedge \forall x[A(x) \rightarrow walk'(x)]]$.

To complete this sketch of the consequences of my proposal (40) I shall give the entries for *a(n)*, *some* and *every* without much comment.

(48) *a(n)*, DET, ..., <<<e,t>,t>, <<<e,t>,t>,t>>
$\lambda\mathcal{Q}\lambda\mathcal{P}\exists A[\mathcal{Q}(A) \wedge \mathcal{P}(A)]$

some, DET, ..., <<<e,t>,t>, <<<e,t>,t>,t>>
$\lambda\mathcal{Q}\lambda\mathcal{P}\exists A[\mathcal{Q}(A) \wedge \mathcal{P}(A)]$

every, DET, ..., <<<e,t>,t>, <<<e,t>,t>,t>>
$\lambda\mathcal{Q}\lambda\mathcal{P}\forall A[\mathcal{Q}(A) \rightarrow \mathcal{P}(A)]$.

As a consequence of the present analysis the difference between *a(n)* and *some* can be accounted for in a natural way: the entry for *a(n)* and *every* must be extended to include the information that *a(n)* can only c-command the SING-node, whereas *some* can take both SING and PLUR as in *I saw some man at the door* and *I saw some men at the door*, respectively.

I shall now discuss some of the consequences of the above proposal. First of all, my statements with respect to (18) - (20) must be slightly modified. The path from N^0 up to and including N^{m-1} is less homogeneous than presented there: values for N^i can differ in their categorial status. At this point we enter a very interesting area: the $\bar{X}$-syntax requires that all heads along a projection line share a common element, because this constant puts us in a position to express that X^0 is the head of X^m. For all that we see that certain changes in the categorial status can take place, notably the change from <e,t> into <<e,t>,t> on the N-projection line as a result of the NUM-operation. So there is a certain tension between our wish to have a constant element along the N-projection line and the necessity to allow of categorial changes. I think that this tension can be resolved. After all a generalization is made by the simple fact that the <e,t> occurs at the leftmost element in the representations of N^0, N^1 and N^2 in (41).

Secondly, the present approach accounts for the difference between *three nice books* and **nice three books*. Moreover, it accounts for the ungrammaticality of (49).

(49) *These three four books.

The numerical *three* cannot operate on the <<e,t>,t>-phrase *four books*. Observe that (49) has a grammatical counterpart meaning 'these three or four books'. In this case *three* and *four* are co-ordinated. It is interesting to see that there are some circumstances in which numerals can be lowered along the N-projection line as shown in *those beautiful two weeks, a dusty four miles of road*. However, **those beautiful two trees, *a dusty four churches* are not well-formed unless trees and churches are understood as units of measure.[15]

Thirdly, as contrasted with JACKENDOFF (1977) *few* and *many* can be treated as numerals, along the line from (38) to (44), though with a different semantic representation concerning the cardinality of the set A (cf. BENNETT 1975; KLEIN to appear). Again it is interesting to compare the Dutch data with the corresponding facts of English:

(50) a. weinig kinderen — b. *de weinig kinderen die...
few children
c. ?weinige kinderen — d. de weinige kinderen die...
the few children that...

(51) a. veel kinderen — b. *de veel kinderen die...
many children
c. vele kinderen — d. de vele kinderen die...
the many children that...

There is a clear difference between *veel/weinig* and *vele/weinige*. The latter is a declined form occurring most properly after a determiner. However, *vele kinderen* is well-formed though slightly outmoded. In my idiolect (50c) is worse than (51c). Suppose that *de vele kinderen* in (51d) would be analyzed as *these three children* in (38a). In that case *vele* would be a numeral, showing adjectival features by its declined form. Given the data in (50) and (51) it would be unnatural to put *veel* and *vele* in the same category. Consequently one could argue that *veel* is a determiner rather than a numeral, whereas *vele* is a numeral rather than a determiner. If this argument holds, then one could argue in favour of the view that *three* in *three children* is a determiner as well. As said, I have tried out that line of thought in an earlier version of this paper in an effort to account for set-introduction, cardinality and quantification at the level of DET only. As the present analysis of (38) runs quite smoothly as far as the derivations (42) and (44) are concerned, I have cut off the former line of thought for the time being.

As an immediate consequence of the assumption that *few* and *many* must be treated on a par with *three* a somewhat comical effect arise: the structures of *these three children* and *these few attempts* resemble structure (12): *these* asymmetrically c-commands *three* and *few*. Note, however, that (38a) is not rigid as shown in (16a). By removing numerals (among which *few, many,* etc.) from (22), we do not bring ourselves back to a position where we need the Specifier Constraint (13) to account for the data. The fact that numerals are to be asymmetrically c-commanded by DET, follows from the difference between NUM and DET in (41).

Fourthly, the label A in $\bar{X}$-syntax is to be restricted to "pure adjectives" whatever that may be. KAMP (1972) and BENNETT (1972) among others make it clear that adjectives do not form a homogeneous category. At any rate, $\bar{X}$-syntax can provide for the categories showing adjectival behaviour.

Finally, the determiners *these, the,* and *all* have exactly the same semantic representation in the present analysis, except for their degree of Θ-ship. In *all children* the specific subset A is identified either by context or by deixis. The universal quantifier covers all members of A. The same applies to *these* and *the.*

I conclude here my elucidation of (38) - (48). I am aware of the many intricacies of the specifier structure of NP. However, I believe to have shown that the OODH has consequences that can be dealt with adequately in the framework presented here.[16]

After the above discussion about the two problems of the Bartsch-Bennett approach to plurality that I have tried to solve, the third problem will be briefly discussed. Bennett is aware of the "magnitude of this problem" in the discussion of his second fragment (BENNETT 1975, p.133). To account for the collective reading of (34b) Bennett introduces *three* as in (52).

(52) $$\lambda\bar{P}\exists\bar{x}[\underbrace{group'(\bar{x})}_{A} \wedge \underbrace{\forall y[\bar{x}(y)\rightarrow\zeta'(y)]}_{B} \wedge \underbrace{\overset{3}{\exists} y[\zeta'(y) \wedge \bar{x}(y)]}_{C} \wedge \bar{P}\{\bar{x}\}]$$

In the part indicated by A a group is existentially introduced, in B quantification takes place over members of this group and in C the cardinality is given. The $\bar{P}$ and $\bar{x}$ indicate that one has to do with a letter referring to predicates of predicates and groups, respectively.

What does *I saw (these) three children* mean? In both cases one can have four "seeing events": (i) I saw three children together; (ii) I first saw two children together and then a third child; (iii) I saw first one child and then the other two together; and (iv) I saw them one by one.[17] It seems to me that it is the verbal predicate that creates the (felicitous) vagueness of interpretation. The sentence *I see (these) three children* shows that the present tense element has the effect of bringing the collective reading to the fore. This example, which I owe to Martin Stokhof, seems to support the thesis that one should not account for the collective or distributive reading of sentences in terms of representations for noun phrases resulting from (48). In my opinion the collective or distributive readings of sentences are a matter of temporal quantification, i.e. quantification over events, situations or occasions, rather than a matter of noun phrases.

6. CONCLUSION

I have tried to argue that the conceptual gap between Montague grammar and Chomsky-grammar is not so wide as is often assumed. My critical discussion of a particular version of $\bar{X}$-theory led to a variable $\bar{X}$-syntax having a bottom-to-top rule system, the Minimal Level Hypothesis. The 'only one determiner position hypothesis' in the MLH-framework led to certain insights into the syntax and semantics of NP's that seem to corroborate the view that Montague-grammar fits in quite nicely with Chomsky-syntax.

FOOTNOTES

* This paper results from work with my students. I thank Franciska de Jong, Leonoor Oversteegen, Frans Pennings, Karel van Rosmalen as well as Jacqueline Frijn/Fred Weerman for their substantial contribution to the present version. I am also indebted to my colleagues Johan Kerstens, Frederiek van der Leek and Arie Sturm for many suggestions on content and style of an earlier version. Finally, I want to thank Dennis de Champeaux, Jeroen Groenendijk and Martin Stokhof for their substantial portion in getting the things in Section 5 into the formal jacket I wanted to have them in.

1. In the presentation of this paper at the colloquium I used the term 'complementary distribution' where I have 'contrastive distribution' in the text of this article. My use of the former term linked up with JACKENDOFF's (1977, p.104). However, this use of a term widely applied in a different sense might cause some misunderstanding. Therefore, I have changed the terminology. The essence of contrastive distribution is of an analytical nature: given a string $c_1\ c_2 \ldots c_n$ whose structure is to be determined, c_i and c_{i+1} in the structure $[_C\ c_i\ c_{i+1}]$ never belong to the same category unless one has to do with co-ordination, or with subordination where c_i or c_{i+1} is the head of the phrase C.
2. More precisely, for the cluster X^0 of sub-predicates if we take into account Jackendoff's thematic system as presented in JACKENDOFF (1976). See also VERKUYL (1976).
3. It can easily be understood why Jackendoff does not allow for an X^1-specifier. X^0 is taken as a function mapping ordered n-tuples of terms into propositions. Given the stipulation that all members of such an

n-tuple (minus the subject-NP) are located in the complement of X^1 as sisters of X^0, there is semantically no room for any X^1-specifier.

4. That is, the claim that transformational rules can cross-classify on the basis of features proposed by Jackendoff.
5. Koster is not able to define the notion 'prominent' in terms of c-command because this structural relation does not distinguish between sister nodes. Hence Koster appeals to a certain functional hierarchy to define his notion 'more prominent than'. He is forced to do so because he accepts the U3LH. The notion 'c-command' is taken here in the sense of REINHART (1978): a node A c-commands a node B if and only if A does not dominate B and the first branching node dominating A also dominates B.
6. This question is in fact one of the leading motives of this paper. In Chomsky's present theory one of the most important questions concerns the relationship between surface structure (called S-structure) and logical form. In the present paper I follow REINHART (1978) by trying to put as much syntax into surface structure as possible (as a matter of strategy), thus restricting the syntax of logical form.
7. Note also that the MLH requires that transformational generalizations be made on the basis of X^m (or X^0) in the Structural Descriptions rather than on the basis of the same numerical value, as was pointed out in KERSTENS & STURM (1978, p.42).
8. In Section 5 this contention will be slightly modified.
9. Chomsky's analysis hinges on the direct association of the feature [+plur] with the plural morpheme in phrase structure. As I shall argue the feature [+plur] should be located higher up in the NP. Morphological rules can introduce the plural morpheme on the basis of the presence of the syntactic feature [+plur] in the noun phrase.
10. A lot of problems remain, however. For instance, there is an interaction between predicates and quantified expressions. According to Milsark the sentence *Coffee tastes well, dolphins are intelligent* can be analysed as having a universally quantified NP. Though I shall not elaborate that point I believe that the analysis given in Section 5 can in principle account for this sort of problems.
11. More generally, such an approach does not aim at defining the notion 'possible rule' and 'possible grammar' in the sense that restrictions are put on the function and form of rules for the sake of explanatory adequacy.
12. Chomsky seems to restrict the values for X to the productive categories

N, V and A. the present argument shades this picture somewhat in the sense that it also allows of heads of a phrase structural configuration on the basis of structural parallelism with N-, A-, or V-projections.

13. At the colloquium I explored the line of analysing *three children* as $[_{N^1}[_{DET}\text{three}][_{N^0}\text{child -en}]]$. In the present version the representation in (38b) seems appropriate. One needs the DET Ø anyhow to account for indefinite phrases such as *I saw nice children*. See footnote 16.
14. PLUR is represented here as a categorial label. However, it can be taken as a feature as well. It belongs to the category of specified grammatical formatives.
15. Recall that Jackendoff used precisely these phrases to argue that numerals must be taken as nouns. The combination of *weeks* and *two* in *a beautiful two weeks* is an example of what one might call unit-counting. My impression is that in this case the numeral is dominated by N^0, where N^0 is *x - week,* where *x* is an open place for numerals.
16. I mention some problems here. Firstly, I do not know how the present analysis can be related to Hausser's distinction between Presupposing and Assertive Quantifiers (HAUSSER 1976). Secondly, the effect of removing *few, many,* etc. from (22) would be that **all few children* cannot be accounted for any longer. The most obvious way out is to follow the line I took at the colloquium by distinguishing into *few/three* as DET and *few/three* as adjectives (in the modified sense of (41) (cf. KLEIN to appear). In that case the English *all* in **all few children* would modify the DET *few* which is not allowed just as it is not allowed to modify *the* by *nearly*. In Dutch we do not find *al veel kinderen* or *al vele kinderen* as noun phrases, with *al* as a modifier. As said in the text I think that set-introduction, cardinality and quantification are essentially a matter of DET. The general idea is that DET should contain slots where material can be inserted from below by the combinatorial mechanism shown in (42) and (44). I leave this matter open for future research. Note that the problems mentioned here do not affect the main thesis of this paper at all. My purpose was not to solve all descriptive problems that arise with respect to the specifier structure of NP; I wanted to show that a certain approach can lead to fruitful insights. In fact, the basic idea is that by semantic analysis certain things can be said about the syntax of NP's that otherwise would remain at the level of a good guess.

    Note finally that the present analysis seems to clarify the notion of 'unspecified quantity of X' as used in VERKUYL (1976).

17. I do not exclude the possibility of having to do here with an epiphenomenon from the linguistic point of view. On the other hand, recent analyses of quantification in sentences all locate the properties 'distributional' and 'collective' in the NP itself. I doubt very much if that is correct.

REFERENCES

BACH, E., 1979, 'Control in Montague grammar', *Linguistic Inquiry* 10, 515-531.

BARTSCH, R., 1973, 'The semantics and syntax of number and numbers', in: J.P. Kimball (ed.), *Syntax and Semantics*, Vol. 2, New York, 1973, pp.51-93.

BARTSCH, R. & T. VENNEMANN, 1972, *Semantic structures*, Frankfurt.

BENNETT, M.R., 1975, *Some extensions of a Montague fragment of English*, IULC.

CHOMSKY, N., 1970, 'Remarks on nominalization', in: R.A. Jacobs & P.S. Rosenbaum (eds), *Readings in English transformational grammar*, Waltham (Mass.), 1970, pp.184-221.

CHOMSKY, N., 1977, *Essays on form and interpretation*, New York.

CLARK, H.H. & E.V. CLARK, 1977, *Psychology and language*, New York, etc.

COOPER, R. & T. PARSONS, 1976, 'Montague grammar, generative semantics and interpretive semantics', in: B.H. Partee (ed.), *Montague grammar*, New York, etc., 1976, pp.312-362.

CULICOVER, P.W., 1976, *Syntax*, New York.

DAHL, Ö., 1977, 'Operational grammars', in: Ö. Dahl (ed.), *Logic, pragmatics, and grammar*, Lund, 1977, pp.71-116.

GROOT, A.W. DE, 1949, *Structurele syntaxis*, Den Haag.

GUÉRON, J., 1976, 'Interpretation of PP complements: a semantic filter for PP extraposition', in: H.C. van Riemsdijk (ed.), *Green ideas blown up. Papers from the Amsterdam colloquium on trace theory*, University of Amsterdam, 1976, pp.44-68.

HALITSKY, D., 1975, 'Left branch S's and NP's in English: a bar notation analysis', *Linguistic Analysis* 1, 279-296.

HAUSSER, R.R., 1976, 'Scope ambiguity and scope restriction in Montague grammar', in: J. Groenendijk & M. Stokhof (eds), *Proceedings of the Amsterdam colloquium on Montague grammar and related topics,* Amsterdam Papers in formal grammar, vol. 1, Amsterdam, 1976, pp.95-131.

HORNSTEIN, N., 1977, 'S and X' convention', *Linguistic Analysis* 3, 137-176.

JACKENDOFF, R.S., 1976, 'Toward an explanatory semantic representation', *Linguistic Inquiry* 7, 89-150.

JACKENDOFF, R.S., 1977, *$\bar{X}$-Syntax: A study of phrase structure,* Cambridge (Mass.).

KAMP, J.A.W., 1975, 'Two theories about adjectives', in: E. Keenan (ed.), *Formal semantics of natural languages,* Cambridge, etc., 1975, pp.123-155.

KEAN, M.L., 1978, *Some features of features; Remarks on motivating features of syntactic categories,* GLOW-paper.

KLEIN, E., to appear, *Determiners and the category* Q.

KOSTER, J., 1978, *Locality principles in syntax,* Dordrecht.

KRAAK, A. & W.G. KLOOSTER, 1968, *Syntaxis,* Culemborg/Keulen.

LIGHTFOOT, D.W., 1979, *Principles of diachronic symtax,* Cambridge.

LUBBE, H.F.A. VAN DER, 1965, *Woordvolgorde in het Nederlands; een synchrone beschrijving,* Assen, 2nd impression.

LYONS, J., 1968, *Introduction to theoretical linguistics,* Cambridge.

McCAWLEY, J.D., 1968, 'Concerning the base component of a transformational grammar', *Foundations of Language* 4, 243-269.

MILSARK, G.L., 1977, 'Toward an explanation of certain peculiarities of the existential construction in English', *Linguistic Analysis* 3, 1-29.

MONTAGUE, R.H., 1974, *Formal philosophy; selected papers of Richard Montague,* New Haven.

PARTEE, B., 1975, 'Montague grammar and transformational grammar', *Linguistic Inquiry* 6, 203-300.

REINHART, T., 1978, 'Syntactic domains for semantic rules', in: F. Guenthner & S.J. Schmidt (eds), *Formal semantics and pragmatics for natural languages,* Dordrecht, 1978, pp.107-130.

ROOSE, H., 1956, 'Kategorieën van voorgeplaatste bepalingen bij substantieven', *Levende Talen* 186, 474-483.

STURM, A., 1979, *Over het maximale aantal bars in de X-bar-theorie,* Paper Instituut De Vooys 79-NE-02.

STURM, A. & J.G. KERSTENS, 1978, 'Over verklaren in de taalkunde', in: P.C.A. van Putte & H.J. Verkuyl (eds), *Nieuwe tegenstellingen op Nederlands taalgebied,* Utrecht, pp.130-172.

VENDLER, Z., 1968, *Adjectives and nominalization,* The Hague.

VERHAGEN, A., 1979, 'Fokusbepalingen en grammaticale theorie', *Spektator* 8, 372-402.

VERKUYL, H.J., 1976, 'Interpretive rules and the description of the aspects', *Foundations of Language* 14, 471-503.

VERKUYL, H.J., 1979, 'Prominentie, functionele hiërarchie en de uniforme drie-niveauhypothese, *Spektator* 8, 403-414.

VERKUYL, H.J., 1980, *Samenhang,* Den Haag.

WILLIAMS, E., 1978, X-*bar Features,* GLOW-paper.