AMSTERDAM PAPERS IN FORMAL GRAMMAR

VOL. II

CENTRALE INTERFACULTEIT

UNIVERSITEIT VAN AMSTERDAM

The AMSTERDAM PAPERS IN FORMAL GRAMMAR are edited by Renate BARTSCH, Jeroen GROENENDIJK and Martin STOKHOF and published by the CENTRALE INTERFACULTEIT of the University of AMSTERDAM

VOL. I : *PROCEEDINGS OF THE AMSTERDAM COLLOQUIUM ON MONTAGUE GRAMMAR AND RELATED TOPICS, JANUARY 1976*

Jeroen GROENENDIJK & Martin STOKHOF (eds.) f 15,- ($ 7,-)

Contents:

Renate Bartsch, Syntax and Semantics of Relative Clauses, *Jerry Edmonson,* Semantics, Games and Anaphoric Chaining, *Michael Grabski,* Zur Semantik von 'Meinen' und 'Bedeuten', *Jeroen Groenendijk & Martin Stokhof,* Some Aspects of the Semantics and Pragmatics of Performative Sentences, *Roland Hausser,* Scope Ambiguity and Scope Restrictions in Montague Grammar, *Jaap Hoepelman,* Mass-Nouns and Aspects, or: Why We Can't Eat Gingercake in an Hour, *Theo Janssen,* A Computer Program for Montague Grammar: Theoretical Aspects and Proofs for the Reduction Rules, *Peter Lutzeier,* The Counterfactual Reading of 'Before', *Frans Plank,* Misunderstanding Understood Subjects: The Minimal Distance Principle in Montague Grammar, *Helmut Schnelle,* Circ umstances and Circumstantial Expressions, *Frank Veltman,* Prejudices, Presuppositions and the Theory of Conditionals, *Ryszard Zuber,* Operators and Arguments in Natural Language

VOL. II: *PROCEEDINGS OF THE SECOND AMSTERDAM COLLOQUIUM ON MONTAGUE GRAMMAR AND RELATED TOPICS, JANUARY 1978*

Jeroen GROENENDIJK & MARTIN STOKHOF (eds.) f 20,- ($ 9,-)

Copies of the volumes can be obtained by sending a cheque or money order to the Centrale Interfaculteit, Universiteit van Amsterdam, Roetersstraat 15, Amsterdam, The Netherlands, or by transferring the required amount on Postgiro (Postal transfer) account 13500 of the Gemeente Giro Amsterdam, in favour of account B 13021 of the Bureau Centrale Interfaculteit, Universiteit van Amsterdam, stating 'Amsterdam Papers'.

AMSTERDAM PAPERS IN FORMAL GRAMMAR

VOL. II

PROCEEDINGS OF THE SECOND AMSTERDAM COLLOQUIUM

ON MONTAGUE GRAMMAR AND RELATED TOPICS

JANUARY 1978

edited by

Jeroen Groenendijk

and

Martin Stokhof

CENTRALE INTERFACULTEIT

UNIVERSITEIT VAN AMSTERDAM

1978

PREFACE

This second volume of the series "Amsterdam Papers in Formal Grammar" contains most of the papers read at the "Second Amsterdam Colloquium on Montague Grammar and Related Topics", held on january 9-13, 1978. The papers missing are those of Jens Allwood, Colin Biggs, Ewan Klein. Frans Plank and Arnim vom Stechow.

To achieve quick and unexpensive publication, the papers are reproduced exactly as they were recieved. Therefore, only the authors are responsible for form and content of their contribution.

Compared with volume I, the papers included in this volume show even more divergence from Montague Grammar, as is to be expected in the course of the treatment of new problems in the area of formal grammar and the development of interests expanding the original field.
Thus, there is a move towards context dependent interpretation and formal pragmatics, towards lexical semantics, towards the application to the semantics of programming languages, and towards the theory about the form of grammar and the means used in constructing grammars, all of which shows a growing awareness of the problems that arose with the development of the field within the last years, and the merits it may have in the long run.

Renate Bartsch

TABLE OF CONTENTS

ADDRESSES OF THE AUTHORS

Thomas BALLMER, Sprachwissenschaftliches Institut, Ruhr-Universität Bochum, 463 Bochum, West Germany

Renate BARTSCH, Centrale Interfaculteit, Universiteit van Amsterdam, Roetersstraat 15, Amsterdam, The Netherlands

Pieter UIT DEN BOOGAART, Technische Hogeschool Eindhoven, Afdeling Wijsbegeerte en Maatschappijwetenschappen, Den Dolech 2, Eindhoven, The Netherlands

Harry BUNT, Technische Hogeschool Eindhoven, Instituut voor Perceptie Onderzoek, Den Dolech 2, Eindhoven, The Netherlands

Jerold EDMONSON, Fachbereich Sprachwissenschaft, Technische Universität Berlin, 1 West-Berlin 10, West Germany

Peter VAN EMDE BOAS, Instituut voor Toepassingen van de Wiskunde, Universiteit van Amsterdam, Roetersstraat 15, Amsterdam, The Netherlands

Peter EISENBERG, Institut für Deutsche Sprache, Technische Universität, Am Welfengarten 1, D-3 Hannover, West Germany

Michael GRABSKI, Fachbereich 16, Freie Universität Berlin, Habelschwertler Allee 45, 1 Berlin 33, West Germany

Jeroen GROENENDIJK, Centrale Interfaculteit, Universiteit van Amsterdam, Roetersstraat 15, Amsterdam, The Netherlands

Roland HAUSSER, Seminar für Deutsche Philologie II, Universität München, 8 München 40, West Germany

Frank HENY, Instituut voor Algemene Taalwetenschap, Rijksuniversiteit Groningen, Grote Rozenstraat 31, Groningen, The Netherlands

Theo JANSSEN, Mathematisch Centrum, Tweede Boerhaavestraat 49, Amsterdam The Netherlands

Peter LUTZEIER, Fachbereich Germanistik, Freie Universität Berlin, Habelschwerler Allee 45, 1 Berlin 33, West Germany

Martin STOKHOF, Centrale Interfaculteit, Universiteit van Amsterdam, Roetersstraat 15, Amsterdam, The Netherlands

Analysis and Synthesis of Linguistic Surface

Thomas T Ballmer

1. Introduction

In this short paper I want to treat the question whether it is reasonable to distinguish conceptually and formally between synthesis of natural language expressions (NL-expressions) on the one hand and analysis of NL-expressions on the other.

Strong motivations to differentiate the two directions synthesis and analysis for a grammar could be inferred from considerations such as the following: It can be taken as one of the major goals of setting up a theory of language - which in standard cases involves a notion of a formal grammar - to account for the linguistic competence of individual language users, or of typical language users or of language user communities. This competence is to be described or reconstructed in the grammar (system of grammars) referred to in a theory of language.

This competence, as I take it, is a competence of performance and of interpretation, or more explicitly, a competence to perform linguistic actions and to draw conclusions out of such linguistic actions performed by oneself and others. Now, a motivation to differentiate between the two directions synthesis and analysis rests upon the fact that the general framework is, according to our view, one in which there is at least one speaker producing linguistic expressions from cognitive entities such as deep structure, concepts, known or believed propositions, logical forms, and at least one hearer interpreting the linguistic expressions produced by the speaker(s) by means of assigning them cognitive entities (incidentally the set of speakers and hearers may coincide): referring to this framework, a grammar reconstructing the kind of competence I have in mind, i.e. the competence of performance and interpretation, breaks up naturally into two parts: the speaker's competence to perform linguistic actions (synthesis) and the hearer's competence to interpret them (analysis).

Unfortunately this motivation for a theoretical distinction of synthesis and analysis of linguistic expressions though it is a good motivation cannot be taken as a valid argument in favor of the synthesis/analysis distinction.

More thorough considerations force us to go another way. Therefore we shall present two different arguments. After a first more abstract approach which presents a good reason for a theory independent position, we will develop, secondly, a much more specific approach to our problem. This second approach is linguistically more palpable. It concerns the synthesis and analysis of sentences containing pronouns (or more generally proforms). As will be demonstrated by means of this approach there are essential differences between functions synthesizing NL-expressions (from logical forms) and functions analysing NL-expressions (into logical forms). The two processes going to be described - namely the principled synthesis of pronouns from logical constants, and the principled analysis of pronouns into logical constants - we shall call pronominalization and depronominalization, respectively.

Functions of synthesis and analysis, i.e. in our case pronominalization and depronominalization, need not be inverses of each other in the mathematical sense of the word, though they may come close to that. Especially, it will turn out that it is no trivial matter to get from synthesis functions to analysis functions or vice versa.

2. The Two Arguments for a Synthesis/Analysis Distinction

For most theories of language and grammars there is no distinction between synthesis of NL-expressions and analysis of NL-expressions.[1] Chomsky's doctrine that "there is no general notion 'direction of mapping' or 'order of steps of generation' to which one can appeal" (Chomsky 1969, p. 188) has found wide acceptance. The main task for a grammar is considered to be merely the adequate delineation of NL-expressions (and their structures), of pairs of NL-expressions (and their structures) and their logical forms, of pairs of NL-expressions (and their structures) and deep-expressions (and their structures) and the like.[2]

a) the first argument

My first argument is based on the fact that the (natural) inverse of a function having certain properties need not have the same properties. The inverse of an everywhere differentiable function, say, need not be differentiable everywhere. But more to the point, the (natural) inverse of a totally recursive function need not be totally recursive itself.[3] Thus there can be computationally relevant differences of a function and its (natural) inverse. This is a strong argument in favor of distinguishing a function and its inverse from a computational point of view.[4]

b) the second argument

This has been very abstract, so far. - I think we all agree that there is a need for a linguistically more directly interpretable result. Therefore we will turn our attention to the second argument for a synthesis/analysis distinction.

In order to carry through this argument a little preparation is necessary. I will work in the context of logical grammar, expecially in the context of what I call language reconstruction systems (LRSs).[5/6]

The few next steps we are going to present are the following. We present in a very abbreviated form the syntactical head of LRSs. In (1) an analysis of the very simple sentence ALF $RUNS_0$ and in (2) the synthesis of this sentence is given. In (3) the analysis and synthesis (if read from the bottom to the top) of an extremely simple text with pronouns is presented. For more specific information about the syntactical head of LRSs the reader is referred to Ballmer (1978). A catch-word-like list of properties of the syntactical head of LRSs is added.

The Syntactical Head of LRSs

(1) ALF RUNS o

$$\lambda P P(A) \quad \lambda u Runs(u) \quad \kappa\Pi_{VI}\kappa\Pi_{NP}(\Pi_{NP}(\Pi_{VI}))$$

$$\lambda P P(A) \qquad \kappa\Pi_{NP}(\Pi_{NP}(\lambda u Runs(u)))$$

$$\lambda P P(A)\ (\lambda u\ Runs(u)).$$

$$(\lambda u\ Runs(u))(A).$$

$$Runs(A)$$

0. Phonetic-Phonologic Sequence of morphemes (or words)
1. Disambiguation by Indices
2. Word by Word translation into expressions of λ-κ calculus; use of the lexicon!
3. Applying the rule of the grammar

(read in word by word from left to right until you hit a punctuation sign - i.e. κ-expression; reduce κ operators, then reduce λ-operators; go on until no κ- or λ-operator occurs, if possible)

4. The linguistic expression we started with is grammatical just in case λ-κ reduction is possible: *The result of the λ-κ reduction is the meaning (representation) of the linguistic expression in question.*

(2)

$$\mathrm{Runs}(A)$$

$$(\lambda u\, \mathrm{Runs}(u))(A)$$

$$\lambda P P(A)(\lambda u\, \mathrm{Runs}(u))$$

$$\lambda P P(A)\ \kappa\Pi_{NP}(\Pi_{NP}(\lambda u\, (\mathrm{Runs}(u))$$

$$\lambda P P(A)\ \ \lambda u\, \mathrm{Runs}(u)\ \ \kappa\Pi_{VI}\ \ \kappa\Pi_{NP}(\Pi_{NP}(\Pi_{VI}))$$

ALF RUNS o

0. Meaning (Representation), standing for a thought-content a speaker wants to express.
1. Subsequent λ- and κ-extraversions
2. Translation, expression by expression to surface (disambiguated surface); use of the lexicon!
3. Ambiguation by Indices
4. *According to Punctuation Lexicon there are expressible exactly those surfaces which are grammatical in the language fragment under discussion.*

(3) ALF DRINKS o HE IS SAD o

$$\mathrm{ALF}_8\ \ \mathrm{DRINKS}_3\ \ o_4\ \ \mathrm{HE}_2\ \ \mathrm{IS}_7\ \ \mathrm{SAD}_1\ \ o_7$$

$$\mathrm{Al}_8\ \ \lambda u\, \mathrm{Drinks}_3(u)\ \ \kappa\Pi_{VI}\kappa\Pi_N(\Pi_{VI}(\Pi_N))\ \ldots$$

$$\mathrm{Al}_8\ \ \kappa\Pi_N((\lambda u\, \mathrm{Drinks}_3(u))(\Pi_N))\ \ldots$$

$$(\lambda u\, \mathrm{Drinks}_3(u))\mathrm{Al}_8\ \ldots$$

$$\mathrm{Drinks}_3(\mathrm{Al}_8)\ \ldots$$

$$\mathrm{Drinks}_3(\mathrm{Al}_8)\ \mathrm{he}_2\ \ \ \lambda P\lambda u\, P(u)\ \ \ \lambda u\, \mathrm{Sad}_1(u)$$

$$\kappa\Pi_{PA}\kappa\Pi_{Cp}\kappa\Pi_N\kappa\Pi_S(\Pi_S \wedge (\Pi_{Cp}(\Pi_{PA}))(\Pi_N))$$

$$\mathrm{Drinks}_3(\mathrm{Al}_8)\wedge((\lambda P\lambda u\, P(u))(\lambda u\, \mathrm{Sad}_1(u)))(\mathrm{he}_2)$$

$$\mathrm{Drinks}_3(\mathrm{Al}_8)\wedge \mathrm{Sad}_1(\mathrm{he}_2)$$

Some Properties of the Syntactical Head of LRS:

1. The Syntactical Head of LRSs is CF

   (This does not contradict the classical result of Chomsky mathematically, but linguistically, *because I speak of a different, and apparently more adequate set of linguistic entities:* namely texts, including suprasegmental information!)

2. Words, Rules, Transformations in the traditional senses are not structurally different! Symbolically:
   Words ≡ "Rules" ≡ "Transformations"
   namely: all these entities are λ-κ expressions.

3. There is a distinction between primary words (morphemes) and secondary words (morphemes) at a finer level!
   Secondary words (morphemes) contain suprasegmental information.

phonetic-phonologic:	*Intonation, Rhythm, ...*
morpho-syntactic:	*"Rules", "Transformations"*
semantico-pragmatic:	*Illocutionary Force*
procedural:	*Stop and Think*

4. CF-"transformations"
5. LRSs are textgrammars
6. LRSs are procedural
7. Function/Argument Distinction plays no role
8. Intuitive translations of surface expressions into λ-expressions!
9. There is only one rule (λ-κ reduction/expansion)
10. No need for Categorial Grammar
    (CG only for convenience)
11. No λPP(A)-translations of names necessary (cf. Montague Grammar)
12. Lexicon plays a major role:
    Every language is characterized by exactly one lexicon.

The study of language is hence thrown back to the study of LEXICA.

13. LRSs are tight. Not too powerful, not too powerless: disconnected constituents, scope ambiguity, but still CF.
14. Computational and (Quasi-)Psychological [procedural] adequacy:
    1. left to right analysis/synthesis
    2. only left-partial understanding necessary
15. (Suggested) Biological Hypotheses: The one and only rule in LRSs to analyse NL-expression is a realization of an (ideally) innate procedure: the behavior cycle of action-nonaction
    (e.g. reading - thinking, being awake - sleeping, living - being dead).
16. To my knowledge LRSs are the only descriptively adequate uniformly organized grammar types, i.e. no use is made of alien metalinguistic means.

I shall exemplify the processes of semantic analysis and synthesis with two processes which are, as I mentioned in the introduction, called *depronominalization* and *pronominalization*. These are processes which are connected intimately with the textual character of names und pronouns. In order to explain the two notions of depronominalization and pronominalization a hierarchy of different logics is to be discussed. The lowest level is called hard logic, upper levels are called soft logics. Natural language is taken to be the softest logic of this hierarchy.[7]

Let me present an extremely simple example which illustrates what I am about to explain with hard and soft logics, their relation to each other mediated by semantic synthesis and analysis. The following is an expression of hard logic:

Stylistic Modifications/Demodifications

(4) ⌜Enter (Jack) ∧ See (Jack, Jill) ∧ Embrace (Jack, Jill)⌝ *depth*

It contains only logical and non-logical constants of the sort usual in predicate logic. A near-surface expression with a similar meaning is the following:

(5) ⌜Enter (Jack) ∧ See (he, Jill) ∧ Embrace (he, she)⌝ *pronominalization/ depronominalization*

And even nearer to the surface, i.e. even softer are the expressions:

(6) ⌜Enter (Jack) ∧ See (he, Jill) ∧ Embrace (he, her)⌝ *case*

(7) ⌜Enters (Jack) ∧ Sees (he, Jill) ∧ Embraces (he, her)⌝ *inflection*

(8) ⌜JACK ENTERS$_0$ HE SEES JILL$_0$ HE EMBRACES HER$_0$⌝ *surface*

The claim is that it is possible to find a principled way of translating levels containing expressions such as (4), or (5), or (6) in order to connect hard logic and linguistic surface. The direction of translating "up" or "down" matters very much, as will be seen in the presentation of our procedures.

We will now give a very short summary of context change logics:

Context Change Logic (Basic Semantics/Pragmatics)

$$h(\ulcorner \downarrow\varphi \urcorner, i) \equiv h(\ulcorner \not\downarrow\varphi \urcorner, \underbrace{\xi(\ulcorner \downarrow\varphi \urcorner, i)}_{j})$$

$\xi(\ulcorner \downarrow\varphi \urcorner, i) \equiv$ nearest Context j relative to i in which φ is true *(one standard way to do it)*

$\downarrow$: force operator
$\ddagger$: ad acta force operator *(may be the empty sign)*
h : interpretation function
ξ : force function
$\equiv$: irreversible metalinguistic identity

important case:

$$h(\ulcorner \alpha \circ \beta \urcorner, i) \equiv h(\ulcorner \alpha \urcorner, \underbrace{\eta(\ulcorner \alpha \urcorner, i)}_{k}) \circ h(\ulcorner \beta \urcorner, \underbrace{\xi(\ulcorner \beta \urcorner, i)}_{j})$$

$\ulcorner \alpha \urcorner, \ulcorner \beta \urcorner$: formulas
$\ulcorner o \urcorner$: two place logical function symbols (e.g. $\wedge, \Rightarrow, \ldots$)
η : another force function

Force operators are different from Modal Future operators they change what is actual, present

special applications

Depronominalization
Pronominalization

- Generally: reference (def. / indef. articles, proforms, ...), textgrammar, speech acts: assertives, directives, expressives, commissives, declaratives ...

special task for each case of applications

- Determine context (context type)
- Specify force function

More about context change logics can be found in Ballmer (1977, 1978).

Let us now have a look at two special kinds of context change logics by way of examples accounting for depronominalization and pronominalization.

Depronominalization

(9) $h(\ulcorner \pi \urcorner, I) \equiv \pi$ — $\ulcorner \pi \urcorner$: *propositional constant*

$h(\ulcorner \alpha \wedge \beta \urcorner, I) \equiv h(\ulcorner \alpha \urcorner, I) \wedge h(\ulcorner \beta \urcorner, \xi_{DI}(\ulcorner \alpha \urcorner, I))$

$h(\ulcorner \Pi\gamma \urcorner, I) \equiv \Pi(h(\ulcorner \gamma \urcorner, I))$ — $\ulcorner \gamma \urcorner$: *standard/nonstandard constant*

$h(\ulcorner \delta \urcorner, I) \equiv \delta$ — $\ulcorner \delta \urcorner$: *standard constant*

$h(\ulcorner he_2 \urcorner, I) \equiv \varepsilon x(x \in I)$ — $\ulcorner he_2 \urcorner$: *nonstandard constant*

(10) $\xi_{DI}(\ulcorner \Pi(\gamma) \urcorner, I) \equiv \xi_{DI}(\ulcorner \gamma \urcorner, I)$

$\xi_{DI}(\ulcorner \delta \urcorner, I) \equiv I \cup \{\delta\}$

$\xi_{DI}(\ulcorner he_2 \urcorner, I) \equiv I$

(11) $h(\ulcorner Drinks_3(Al_8) \wedge Sad_1(he_2)\urcorner, \emptyset) \equiv$

$h(\ulcorner Drinks_3(Al_8)\urcorner, \emptyset) \wedge h(\ulcorner Sad_1(he_2)\urcorner, \xi_{DI}(\ulcorner Drinks_3(Al_8)\urcorner, \emptyset)) \equiv$

$$\left[\begin{array}{l} \xi_{DI}(\ulcorner Drinks_3(Al_8)\urcorner, \emptyset) \equiv \\ \xi_{DI}(\ulcorner Al_8\urcorner, \emptyset) \equiv \\ \{AL_8\} \end{array}\right]$$

$h(\ulcorner Drinks_3(Al_8)\urcorner, \emptyset) \wedge h(\ulcorner Sad_1(he_2)\urcorner, \{Al_8\}) \equiv$

$Drinks_3(h(\ulcorner Al_8\urcorner, \emptyset)) \wedge Sad_1(h(\ulcorner he_2\urcorner, \{Al_8\})) \equiv$

$Drinks_2(Al_8) \wedge Sad_1(Al_8)$

Pronominalization

(12) $e(\ulcorner \pi \urcorner, E) \equiv \pi$

$e(\ulcorner \Pi(\gamma)\urcorner, E) \equiv \ulcorner \Pi(\urcorner \frown e(\ulcorner \gamma\urcorner, E) \frown \ulcorner)\urcorner$

$e(\ulcorner \alpha \wedge \beta\urcorner, E) \equiv e(\ulcorner \alpha\urcorner, \eta(\alpha, E)) \frown \ulcorner \wedge \urcorner \frown e(\ulcorner \beta\urcorner, \xi(\ulcorner \alpha\urcorner, E))$

$e(\ulcorner \delta\urcorner, E) \equiv \delta \quad \text{if } \delta \in E$

$e(\ulcorner \delta\urcorner, E) \equiv he \quad \text{if } he \in E$

$\vdots$

(13) $\eta(\ulcorner \delta\urcorner, E) \equiv E \cup \{\ulcorner \delta\urcorner\}$

$\eta(\ulcorner \Pi(\gamma)\urcorner, E) \equiv \eta(\ulcorner \gamma\urcorner, E)$

$\xi(\ulcorner \delta\urcorner, E) \equiv E \cup \{\ulcorner he\urcorner\}$

$\xi(\ulcorner \Pi(\delta)\urcorner, E) \equiv \xi(\ulcorner \delta\urcorner, E)$

(14) $e(\ulcorner Walk(John) \wedge Look\text{-}Around(John)\urcorner, \emptyset) \equiv$

$e(\ulcorner Walk(John)\urcorner, \eta(\ulcorner Walk(John)\urcorner, \emptyset)) \frown \ulcorner \wedge \urcorner \frown e(\ulcorner Look\text{-}Around(John)\urcorner, \xi(\ulcorner Walk\text{-}Around(John)\urcorner, \emptyset)) \equiv$

$e(\ulcorner Walk(John)\urcorner, \eta(\ulcorner John\urcorner, \emptyset)) \frown \wedge \frown e(\ulcorner Look\text{-}Around(John)\urcorner, \xi(\ulcorner John\urcorner, \emptyset)) \equiv$

$e(\ulcorner Walk(John)\urcorner, \{\ulcorner John\urcorner\}) \frown \wedge \frown e(\ulcorner Look\text{-}Around(John)\urcorner, \{\ulcorner he\urcorner\}) \equiv$

$\ulcorner Walk(\urcorner \frown e(\ulcorner John\urcorner, \{\ulcorner John\urcorner\}) \frown \ulcorner) \wedge Look\text{-}Around(\urcorner e(\ulcorner John\urcorner, \{\ulcorner he\urcorner\})) \equiv$

$\ulcorner Walk(John) \wedge Look\text{-}Around(he)\urcorner$

Some Examples

We now present a rather fragmentary list of examples which can be treated along the lines suggested by the depronominalization and the pronominalization technique. These examples comprise mixed cases of pronouns and quantified phrases (a), reflexives (b), possessives (c,d), TOO-proforms (e,f) and sloppy identity (f). It should be clear how these cases are treated by context change logics. An explicit solution to some harder examples is given subsequently.

(a) TREES STAND THERE. THEY ARE SHADOWY.
TREES STAND THERE. THERE ARE PEOPLE NEARBY. { *THEY ARE SHADOWY. / THE TREES ARE SHADOWY. }

(b) Shave (John,John) ↔ Shaves (John,himself) ↔ JOHN SHAVES HIMSELF.

(c) V_1 x(Dance(x) ∧ Father (x,George)) ↔ Dances (father(George)) ↔ GEORGE'S FATHER DANCES.

(d) V_1 x(Dance(x) ∧ Father (x,he)) ↔ Dances (father (he)) ↔ HIS FATHER DANCES.

(e) Kiss (Bob,Ann) ∧ Kiss (Bill,Ann) ↔ Kisses (Bob,Ann) ⩓ Too (Bill) ↔ BOB KISSES ANNA AND BILL TOO.

(f) - Kiss (Bob,Ann) ∧ Wife (Ann,Bob) ∧ Kiss (Bill,Ann)
  ↔ Kiss (Bob, his (Wife)) ⩓ Too (Bill)
- Kiss (Bob,Ann) ∧ Wife (Ann,Bob) ∧ Kiss (Bill,Joan) ∧ Wife (Joan,Bill)
  ↔ Kiss (Bob,his (Wife)) ⩓ Too (Bill)

Further Examples[8]

The following example is a problem of pronominal reference going back to Geach (1962). It is a problem of dangling variables. This is just one of the many cases where context change logics prove to give an enlightening and elegant solution. Beside the idea of context change only first order predicate logic is used. The correct reading (all-quantification also, for the donkeys!) is gotten by systematic elimination of the non-standard constructs. The constant ⌜b⌝ introduced by the quantifier ⌜One⌝ is non-standard. We call it adjustable constant. Its elimination rule is not

$$\vdash \ulcorner \Pi(b) \urcorner \;//\; \vdash \ulcorner \Lambda\, x\; \Pi(x) \urcorner$$

but rather

$$\vdash \ulcorner \Pi(b) \urcorner \;//\; \vdash \ulcorner V\, x\; \Pi(x) \urcorner$$ [9]

(g) EVERY MAN WHO OWNS A DONKEY BEATS IT_o

$\lambda P \lambda Q$ Every x (Px,Qx) λu Man(u) $\kappa\Pi^1_{CN}\gamma\Pi_{V2}\gamma\Pi_{DET}\gamma\Pi^2_{CN}\, \lambda u\, (\Pi^1_{CN}(u) \triangleright ((\Pi_{DET}(\Pi^2_{CN}))(\lambda w \Pi_{V2}(u,w))))$

$\lambda v \lambda w$ Own(v,w) $\lambda P \lambda Q$ One y(Py,Qy) λu Donk(u) $\lambda v \lambda w$ Beat(v,w) it

$\kappa\Pi_N \kappa\Pi_{V2} \kappa\Pi_{CN} \kappa\Pi_{DET}(((\Pi_{DET}(\Pi_{CN}))(\lambda u\, \Pi_{V2}(u,\Pi_N)))$

$$\text{Every } x\, (Man(x), \text{One } y\, (Donk(y),\ Own(x,y)),\ Beat(x,it))$$

$$\left[\begin{array}{l}
h(\ulcorner \text{Every } x\, (Px,Qx)\urcorner, A) \equiv \bigwedge x\, (h(\ulcorner Px\urcorner, A) \rightarrow h(\ulcorner Qx\urcorner, A)) \\
h(\ulcorner \text{One } y\, (Py,Qy)\urcorner, A) \equiv \bigvee y (h(\ulcorner Py \wedge y = b\urcorner, A) \wedge h(\ulcorner Qy\urcorner, A \cup \{b\})) \\
h(\ulcorner \alpha \triangleright \beta\urcorner, A) \equiv h(\ulcorner \alpha\urcorner, A) \wedge h(\ulcorner \beta\urcorner, A) \\
h(\ulcorner it\urcorner, A) \equiv \varepsilon x\, (x \in A)
\end{array}\right]$$

$$\bigwedge x((Man(x) \wedge \bigvee y(Donk(y) \wedge y = b \wedge Own(x,y)) \rightarrow Beat(x,b))$$

$$\bigwedge x \bigwedge y((Man(x) \wedge Donk(y) \wedge y = b \wedge Own(x,y)) \rightarrow Beat(x,b))$$

$$\bigwedge x \bigwedge y((Man(x) \wedge Donk(y) \wedge Own(x,y)) \rightarrow Beat(x,y))$$ [10]

The following problem is Seuren's (1976). It is a more complex version of an example given by Geach (1962). It is again of the dangling variable type. Its first solution still involves an adjustable constant. This can be eliminated, however, if an appropriate intensional logic is used.

(h) PLATO BELIEVES THAT SOCRATES OWNS A DOG AND HOPES THAT SOCRATES LOVES IT_o

$$\begin{array}{l}
Plato\ \lambda u \lambda \emptyset Bel(u,\emptyset)\ \kappa\Pi_{V1,1}\gamma\Pi_N\gamma\Pi_{V2}\gamma\Pi_{DET}\gamma\Pi_{CN}\lambda u(\Pi_{V1,1}(u,((\Pi_{DET}(\Pi_{CN}))(\lambda w \Pi_{V2}(\Pi_N,w))) \\
\quad Socr\ \lambda v \lambda w\ Own(v,w)\ \lambda P \lambda Q\ \text{One } y(Py,Qy)\ \lambda u\ Dog(u)\ \kappa\Pi_{V1}\kappa\Pi_N\lambda P(\Pi_{V1}(\Pi_N) \wedge P(\Pi_N)) \\
\quad \lambda u \lambda \emptyset\ Hope(u,\emptyset)\ \kappa\Pi_{V1,1}\gamma\Pi_N^1\gamma\Pi_{V2}\gamma\Pi_N^2\ \lambda u\ (\Pi_{V1,1}(u,(\Pi_{V2}(\Pi_N^1,\Pi_N^2)))) \\
\quad Socr\ \lambda v \lambda w\ Love(v,w)\ it\ \kappa\Pi_{V1}\kappa\Pi_{NP}(\Pi_{NP}(\Pi_{V1}))
\end{array}$$

$$Bel(Plato,\ \text{One } y(Dog(y),Own(Socr,y))) \wedge Hope(Plato,Love(Socr,it))$$

$$\left[\begin{array}{l}
h(\ulcorner \alpha \wedge \beta\urcorner, A) \equiv h(\ulcorner \alpha\urcorner, A) \wedge h(\ulcorner \beta\urcorner, \xi(\ulcorner\alpha\urcorner, A)) \\
\xi(\ulcorner ___ \ \text{One } y(Py,Qy)\urcorner, A) \equiv \xi(\ulcorner \text{One } y(Py,Qy)\urcorner, A) \equiv A \cup \{b\} \\
\text{especially: } \xi(\ulcorner \Pi_{V1,1}(\delta, \text{One } y(Py,Qy))\urcorner, A) \equiv \xi(\ulcorner \text{One } y(Py,Qy)\urcorner, A)
\end{array}\right]$$

$$Bel(Plato, \bigvee y(Dog(y) \wedge y = b \wedge Own(Socr,y)) \wedge Hope(Plato,Love(Socr,b))$$ [11]

Thus we may standardize further as promised:

$$\bigvee y(Bel(Plato,Dog(y) \wedge Own(Socr,y)) \wedge Hope(Plato,Love(Socr,y)))$$

This does not, of course, imply:

$$\bigvee y\ E(y) \wedge (Bel(Plato,Dog(y) \wedge Own(Socr,y)) \wedge Hope(Plato,Love(Socr,y)))$$ [12]

The first of the next two examples are known by the name Bach-Peters paradox. It is easily treated by the formalism developed so far. The second of the two examples exhibits the same "turtling all the way down" but is linguistically much less controversial. To my knowledge it has not been discussed in the literature.

(i) THE BOY WHO LOVES HER KISSES THE GIRL WHO HATES HIM_o

$$\lambda P \lambda Q\ \overset{M}{The}\ x(Px,Qx)\ \lambda u\ Boy(u)\ \kappa\Pi_{CN}\gamma\Pi_{V2}\gamma\Pi_{N}\lambda u(\Pi_{CN}(u)\rhd\Pi_{V2}(u,\Pi_{N}))$$

$$\lambda v\lambda w\ Love(v,w)\ her\ \lambda v\lambda w\ Kiss(v,w)\ \lambda P\lambda Q\ \overset{F}{The}\ y(Py,Qy)\ \lambda u\ Girl(u)$$

$$\kappa\Pi_{CN}\gamma\Pi_{V2}\gamma\Pi_{N}\lambda u(\Pi_{CN}(u)\rhd\Pi_{V2}(u,\Pi_{N}))\ \lambda v\lambda w\ Hate(v,w)\ him$$

$$\kappa\Pi^2_{CN}\kappa\Pi^2_{DET}\kappa\Pi_{V2}\kappa\Pi^1_{CN}\kappa\Pi^1_{DET}((\Pi^1_{DET}(\Pi^1_{CN}))((\lambda v\Pi^2_{DET}(\Pi^2_{CN}))\lambda w\Pi_{V2}(v,w)))$$

⋮

$$\overset{M}{The}\ x\ (Boy(x)\rhd Love(x,her),\overset{F}{The}\ y\ (Girl(y)\rhd Hate(y,him),Kiss(x,y)))$$

with

$$\left[\begin{array}{l}
h(\overset{M}{The}\ x\ (\Pi x,\Sigma x),A^M,A^F)\ \equiv \\
\qquad \bigvee_1 x\ (h(\Pi x\wedge x=a,\xi(\Sigma x,A^M\cup\{a\},A^F))\ ,\ h(\Sigma x,\xi(\Pi x,A^M\cup\{a\},A^F))) \\
h(\overset{F}{The}\ y\ (\Pi y,\Sigma y),A^M,A^F)\ \equiv \\
\qquad \bigvee_1 y\ (h(\Pi y\wedge y=b,\xi(\Sigma x,A^M,A^F\cup\{b\}))\ ,\ h(\Sigma y,\xi(\Pi y,A^M,A^F\cup\{b\}))) \\
h(\alpha\rhd\beta,A^M,A^F)\ \equiv\ h(\alpha,A^M,A^F)\wedge h(\beta,A^M,A^F) \\
h(him,A^M,A^F)\ \equiv\ \varepsilon v(v\in A^M) \\
h(her,A^M,A^F)\ \equiv\ \varepsilon w(w\in A^F) \\
\xi(\overset{M}{The}\ x\ (\Pi x,\Sigma x),A^M,A^F)\ \equiv\ A^M\cup\{a\} \\
\xi(\overset{F}{The}\ y\ (\Pi y,\Sigma y),A^M,A^F)\ \equiv\ A^F\cup\{b\}
\end{array}\right]$$

Translational interpretation gives with $A^M = A^F = \emptyset$:

$$\bigvee_1 x\ (Boy(x)\wedge x=a\wedge Love(x,b),\bigvee_1 y\ (Girl(y)\wedge y=b\wedge Hate(y,a),Kiss(x,y)))$$

The two place quantifiers $\bigvee_1$("There is exactly one") may be eliminated according to $\bigvee_1\delta(\Pi(\delta),\Sigma(\delta))\ \equiv\ \bigvee\delta(\bigwedge\gamma(\Pi(\gamma)\leftrightarrow\delta=\gamma)\wedge\Sigma(\delta))$

This is equivalent to:

$$\bigvee x(\bigwedge v((Boy(v)\wedge v=a\wedge Love(v,b))\leftrightarrow x=v)\wedge$$

$$\bigvee y(\bigwedge w((Girl(w)\wedge w=b\wedge Hate(w,a))\leftrightarrow y=w)\wedge Kiss(x,y)))$$

This in turn is equivalent to:

$$\bigvee x \bigwedge v \bigvee y \bigwedge w(((\mathrm{Boy}(v) \wedge v = a \wedge \mathrm{Love}(v,b)) \leftrightarrow x = v) \wedge$$
$$((\mathrm{Girl}(w) \wedge w = b \wedge \mathrm{Hate}(w,a)) \leftrightarrow y = w) \wedge \mathrm{Kiss}(x,y))$$

$$\bigvee x \bigwedge v \bigvee y \bigwedge w(((\mathrm{Boy}(v) \wedge \mathrm{Love}(v,y)) \leftrightarrow x = v) \wedge$$
$$((\mathrm{Girl}(w) \wedge \mathrm{Hate}(w,x)) \leftrightarrow y = w) \wedge \mathrm{Kiss}(x,y))$$

And finally the linguistically less controversial case of a syntactically non-eliminable pronoun:

(k) THE WOMAN WHO KNOWS THAT SHE DREAMS SLEEPS$_o$

$$\lambda P \lambda Q\ \overset{F}{\mathrm{The}}\, y\, (Py,Qy)\ \lambda u\ \mathrm{Wom}(u)\ \kappa\Pi_{CN}\gamma\Pi_{V1,1}\lambda u \lambda \emptyset(\Pi_{CN}(u) \triangleright \Pi_{V1,1}(u,\emptyset))$$
$$\lambda u \lambda \emptyset\ \mathrm{Know}(u,\emptyset)\ \kappa\Pi_{V1,1}\gamma\Pi_{N}\gamma\Pi_{V1}\lambda u(\Pi_{V1,1}(u,\Pi_{V1}(\Pi_N)))$$
$$\mathrm{she}\ \lambda u\ \mathrm{Dream}(u)\ \lambda u\ \mathrm{Sleep}(u)\ \kappa\Pi_{V1}\kappa\Pi_{CN}\kappa\Pi_{DET}((\Pi_{DET}(\Pi_{CN}))(\Pi_{V1}))$$

$$\overset{F}{\mathrm{The}}\, y\, (\mathrm{Wom}(y) \triangleright \mathrm{Know}(y,\mathrm{Dream}(\mathrm{she})),\mathrm{Sleep}(y))$$

$$V_1 y(\mathrm{Wom}(y) \wedge y = b \wedge \mathrm{Know}(y,\mathrm{Dream}(b)),\mathrm{Sleep}(y))$$

$$V_1 y(\mathrm{Wom}(y) \wedge \mathrm{Know}(y,\mathrm{Dream}(y)),\mathrm{Sleep}(y))$$

$$\bigvee y(\bigwedge w((\mathrm{Wom}(w) \wedge \mathrm{Know}(w,\mathrm{Dream}(w))) \leftrightarrow y = w) \wedge \mathrm{Sleep}(w))$$

3. Conclusion

I have presented two arguments in favor of a substantial distinction of synthesis and analysis of natural language expressions. After a formal argument involving considerations of calculability which calls into question Chomsky's dictum that "there is no general notion of 'directions of mapping'", a linguistically more relevant argument has been represented. Pronominalization and depronominalization, i.e. synthesis and analysis of NL-expressions containing pronouns, has been studied and shown to be sufficiently different from each other: The functions reconstructing pronominalization and depronominalization need not be inverses from each other, in the mathematical sense.

From these arguments we can conclude that a grammar should neither be (exclusively) generative nor (exclusively) interpretive, it rather should be both. Language Reconstruction Systems (LRSs), of which I was able to present a small semantical fraction, are able to cope simultaneously with both of these aspects in a formally explicit way, without referring to alien metalinguistic means.

LRSs are therefore superior to every grammar which does not incorporate the productive (generative) and the interpretive aspect simultaneously in a formally explicit way. Thus many known linguistic and logical grammars may be criticized on that level of argumentation.

Moreover we were able to present an argument in favor of the reconstructive view of grammar. As we made amply explicit the first, abstract demonstration of a synthesis/analysis distinction was conducted solely in the descriptive framework. This demonstration shows that it is possible to transcend the descriptive framework from within. The solution of the dilemma goes in the direction of a reconstructive conception, as the second approach successfully shows. Considering what enabled us to formalize synthesis and analysis in closed form, the reconstruction which is at stake here turns out to be the reconstruction of the changing of the context by linguistic means (or instruments) in linguistic actions and non-actions to overcome possible failures of action and communication.

FOOTNOTES

[1]This is true for the different versions of Transformational Grammars, including, Relational Grammar and Double Hierarchy Grammar, but this is true also for the different versions of more wellknown Logical Grammars such as Montague Grammar (EFL, UG, PTQ), Cresswell's and Lewis' Categorially Based Grammars. For some grammars, a distinction is referred to but not formally taken into account. I think this is true for Generative Semantics.

[2]Working in the framework of reconstructing processes of the linguistic understanding of sound and the linguistic expression of thoughts, I tend to opt, perhaps too hastily, for a careful distinction of synthesis and analysis. According to the reconstruction view such a grammar should of course do more than, say, enumerate pairs of phonologic matrices and logical forms, because a reconstruction aims for more than a description. A reconstruction is a mathematical and maybe even psychological modelling of processes which are connected with the production we are trying to describe adequately.
From saying that my approach is reconstructive - which includes providing for an adequate description of the linguistic facts, of course - rather than exclusively descriptive, we could not, however, deduce that the traditional descriptive view is mistaken. The traditional descriptive view is simply different. My line of argument runs rather as follows. Already, on the assumptions of the traditional descriptive view we are lead to accept that synthesis and analysis are to be taken as radically different, and therefore a reconstructive methodology has to be adopted. Thus let us first stay in the descriptive paradigm in order to work out its own limitations.

[3]Or more explicitly, if a function f is partially recursive, its domain D(f) and its range R(f) are recursively enumerable, but not recursive. If the domain D(f) happens even to be recursive, f is called totally recursive. For totally recursive functions the range need not be recursive, also in cases where f is injective, i.e. invertible. In such a case the inverse defined on the recursively enumerable but not recursive range is only partially recursive but not totally recursive.

[4]In the scientific paradigm Chomsky has set up for descriptive linguistics, computational properties play an essential role. There are a number of ways this can be demonstrated.

a) Chomsky's computational hierarchy of formal languages.
b) His methodological restrictions on the basis of computational properties: finite alphabet, finitely many rules, finite applications of rules.
c) His interest in decidability questions.

Let us therefore apply what has just been said about the (possible) differences of computational properties of a function and its inverse to a specifically linguistic problem. Let us consider therefore the function mapping the (stylistically disambiguated) logical forms to meaning disambiguated syntactical forms. It may be totally recursive, whereas at the same time its inverse may be partially recursive. To accept this is to accept on very general grounds that properties of computation may vary in a relevant way depending on whether we are concerned with synthesis or with analysis.

[5]There is not enough time to say but a very few things about these grammatical systems. It should become clear, as has been argued already, that the theoretical result we aim at is largely independent of its demonstration in this special framework. I hope that - as a side result - the easiness and, finally, the possibility of demonstrating our point will speak for the conception and the form of language reconstruction systems themselves.

[6]Language reconstruction systems (LRSs) are grammatical systems, which are to my knowledge the only ones using the same mechanism to settle questions of the grammaticality on levels which are normally treated altogether differently: namely on the base level and on the transformational level. There are those approaches which exclude deliberately the levels of transformation (in some sense Montague [1970], Cresswell [1973]) and those who claim that the classical base structure is sufficient (Thomason [1974]). Most others (Chomsky [1957, 1965, 1969, 1974], Lakoff [1968], Bartsch-Vennemann [1973], Partee [1975]) use two altogether different mechanisms for the base level and the transformational level. There is, however, an alternative way to the LRS approach with respect to the question which we are discussing right now (cf. Lieb 1976).
LRSs make use of two kinds of morphemes, primary morphemes and secondary morphemes. The primary morphemes are morphemes in the traditional sense. The secondary morphemes are morphemes which contain, generally speaking, suprasequental information of various kinds: phonetic-phonologic, morpho-syntactic, semantic-pragmatic, suprasequental information. Syntactically, secondary morphemes - occurring at the surface! - contain compository information - i.e. what traditionally is in base rules and transformations - of a very well defined structure and complexity. LRSs make use of only one proper rule, as far as the question of grammaticality is concerned. It says more or less that a linguistic expression (sentence, phrase) is wellformed in case the secondary morphemes are such that they allow for a complete reduction of all - primary and secondary - morphemes and assign an appropriate meaning representation.

[7]Semantic Analysis and Synthesis are reconstructed and shown to be translation processes which connect logical expressions of the softer levels with logical expressions of the harder levels of the hierarchy and vice versa. The languages of the hard logics only contain classical logical constants such as n-place logical connectives ⌜¬⌝, ⌜v⌝, ⌜∧⌝, quantifiers ⌜V⌝, ⌜Λ⌝ and the like. Soft languages also contain non-classical constants such as extensional and intensional operators ⌜ˇ⌝, ⌜ˆ⌝, modal operators ⌜◇⌝, ⌜□⌝. The non-classical logical constants we are interested in here in this paper are pronominal constants of different kinds and a certain type of name constant.

[8]Due to some remarks during the conference I feel that it is helpful for the reader to demonstrate that it is a straightforward matter to calculate some "harder cases" of pronominal reference. None of the examples took more than half an hour to calculate in this framework of context change logics. As far as I can see these are the first syntactically and semantically fully formalized solutions to these problems.

[9] An adjustable constant is so to speak an externally existentially quantified constant. The other non-standard constant ⌜it⌝ poses no special problems, it simply refers to the adjustable constant ⌜b⌝ which has been introduced by an embedded quantifier phrase.

[10] The ⌜γ⌝ operator works exactly like a ⌜λ⌝ operator as far as conversion and extraversion are concerned. The difference is that it is part of punctuation and as such the timing of its conversion (extraversion) is like that of the ⌜κ⌝ operator.

[11] This can be standardized further (to the point where additionally the adjustable constant ⌜b⌝ is eliminated) if an appropriate intensional logic is used, namely an intensional logic for which it does not matter whether the quantifiers stand in or out of the scope of intensional predicates. A translational interpretation of such a logic would be as follows:

$h(\ulcorner\pi\urcorner, i) \equiv \pi(i)$

$h(\ulcorner\Pi x\urcorner, i) \equiv \Pi(i)x(i)$

$h(\ulcorner\Box\alpha\urcorner, i) \equiv \forall j(iR_{\Box}j \rightarrow h(\ulcorner\alpha\urcorner, j))$

$h(\ulcorner Bel(\delta,\alpha)\urcorner, i) \equiv \forall j(iR^{\delta}_{BEL}\, j \rightarrow h(\ulcorner\alpha\urcorner, j))$

$h(\ulcorner Vx\alpha\urcorner, i) \equiv \exists x(x=a \wedge h(\ulcorner\alpha\urcorner, i))$ ⌜a⌝: adjustable constant for quantifier ⌜∃x⌝

$h(\ulcorner E(x)\urcorner, i) \equiv \exists\kappa(\kappa=x(i))$ ⌜E⌝: existence predicate (≡ there is an incarnation)

Existence quantification only says that the concept exists in the framework (model) so far built up. It is not the case that $Vx\alpha$ implies $Vx(E(x)\wedge\alpha)$. Thus no incarnation need be existent. The following is valid as one easily checks:

$Bel(\delta, Vx\alpha) \equiv Vx\, Bel(\delta,\alpha)$

[12] In order to have nice duality properties for the quantifiers V and Λ slightly different definitions are useful. These definitions have some other advantages as well (they are generalizable to generic cases and to coreference problems related to the coreference problems discussed here). They are slightly more cumbersome to penetrate, discussing our examples here, however.

$h(Vx\alpha, i) \equiv \exists x(x \in A \wedge h(\alpha, i))$

$h(\Lambda x\alpha, i) \equiv \forall x(x \in A \rightarrow h(\alpha, i))$

(A is an adjustable predicate constant). These clauses allow for reasonable interdefinitions of the two quantifiers V and Λ.

References

Ballmer, Th. (1977) — The Instrumental Character of Natural Language. Habilitation Thesis, Ruhr University Bochum. 1073 pp.

(1978) — Logical Grammar. North Holland, Amsterdam (forthcoming).

Bartsch, R.; Vennemann, Th. (1973) — Semantic Structures. Athenäum. Frankfurt.

Chomsky, N. (1969) "Deep Structure, Surface Structure, and Semantic Interpretation", in: Steinberg, D.; Jacobovits, L. (Eds.) Semantics. Cambridge Univ.Press, Cambridge, 1971.

Geach, P. (1962) Reference and Generality. Cornell Univ. Press, Ithaca.

(1972) Logic Matters. Blackwell, Oxford.

Lieb, H.H. (1976) "Principles of Semantics", mimeo, Berlin.

Seuren, P. (1976) Zwischen Sprache und Denken. Athenaion, Wiesbaden.

INFINITIVES AND THE CONTROL PROBLEM IN CATEGORIAL GRAMMAR

Renate Bartsch

The question of how infinitives should be handled in grammar, and specifically in categorial grammar, has been tackled in grammatical studies for quite a while. Frans Planck (1975) has reported on this discussion, in which several times it has been argued that infinitives should be treated as verbs, and, on the other hand side, that infinitives should be treated as noun-phrases. The main argument for the first treatment is that they have complements like verbs do; the main argument for the second treatment is that they are used in subject and object position as complements of verbs, like noun phrases are. Planck (1975) has proposed a syntactic category for infinitives that takes into account both these facts: he treats infinitives as terms (like noun phrases) that can have an internal structure like verbs, namely take complements. An n-place verb has a corresponding infinitive with (n-1)-places. Infinitives get the category $[V^{n-1} / V^{n}]^{m}$; this notation means that the infinitive has m places to be filled by terms and that, if it is applied to an n-place verb, it reduces the number of term places to (n-1), i.e. it fills one term place of the verb. All term places added together, the new complex verb should have (n+m)-1 term places. This way the infinitive has the characteristics of a term, namely to reduce the number of term places by one, when being applied to a verb; and terms can be applied to the infinitive like they can be applied to verbs, where they reduce the number of term places by one.

The virtue of Planck's analysis is, among others, that infinitives are of a single category scheme, namely that of terms, no matter whether they are used in subject position or not. The analysis of Partee 1973, according to which infinitives are of the same category as finite intransitive verbs (IV = t/e), requires quite a different category for infinitives in subject position, like in <u>To walk is easy</u>. Further in Partee 1973, like in Montague 1974, verbs with infinitive complements, like <u>try to</u>, <u>wish to</u>, are of category IV//IV, and thereby not semantically different from adverbials, what I find rather awkward. Further, it is fairly counter-intuitive to parse <u>try to walk</u> into <u>try to</u> and <u>walk</u>, but otherwise to parse <u>to</u> together with the infinitive, as I suppose, Partee would do in <u>to walk is easy</u>; and how would Partee parse <u>I waited for John to come</u>, would she put a verb <u>wait to</u> into the lexicon? Thus it seems rather ad hoc to analyze <u>try to come</u> into <u>try to</u> and <u>come</u>.

Thomason 1974 analyzes infinitives as t//e and treats <u>to</u> as an operator that makes out of an expression of type t/e an infinitive. Thus, semantically infinitives are basic predicates (which he calls AB). Verbs with n-places are treated as expressions that are applied to terms with the result that one place of the verb gets filled by the term and thus we receive a verb with one place less. Thus intransitive verbs are of category t/T, where T is t/(t/e). The weak point in Thomason's and in Partee's analysis is how they handle the control problem.

The control problem consists in how to account for the fact that verbs like promise and permit behave rather differently with regard to which term (noun phrase) governs the hidden subject of the infinitive complement. Examples:

(1) John promises Mary to come.

(2) John permits Mary to come.

In (1) it is John who will come and in (2) it is Mary. For this problem, Thomason and also Partee, propose a syntactic solution. Thomason classifies promise as (IV/INF)/T, this means, as a verb that first combines with a term, e.g. Mary to result into promise Mary, and then takes an infinitive, e.g. to come to receive the result promise Mary to come, which is an intransitive verb. On the other hand side, permit is of category TV/INF. This means, that it first takes an infinitive, e.g. to come with the result permit to come. This is a transitive verb that can take a term, e.g. Mary to yield permit Mary to come. Thus, promise Mary is a phrase while permit Mary is not; promise to come is not a phrase in (1), while in (2) permit to come is. On the other hand side, promise to come is a phrase in John promises to come, where promise is classified as a IV/INF, like try, want, etc. In the terminology of Partee 1975 we have the same: promise is of category (IV//IV)/T, and thus takes Mary to form a phrase; permit to is of category TV/IV, and thus takes come to form a phrase.

The treatment of Thomason and Partee tries to give a solution to the control problem in sentences like (1) and (2) by assigning different syntactic structures to these sentences. The only argument in favour of a distinction between promise and permit in this way comes from the different behaviour of both verbs in the passive construction:

(1') * Mary is promised by John to come.

(2') Mary is permitted by John to come.

As Planck points out, there are, in German, passive constructions with promise, for example (3') and (4').

(3') Mary ist von John versprochen worden kommen zu dürfen.
'Mary is promised by John to be allowed to come.'

(4') Mary ist von John versprochen (worden) eingeladen zu werden.
'Mary is promised by John to be invited.'

which corresponds with (3) and (4) that preferably have object control in German, but are also possible with subject control.

(3) John versprach Mary kommen zu dürfen.
'John promised Mary to be allowed to come.'

(4) John versprach Mary eingeladen zu werden.
'John promised Mary to be invited.'

In constructions like these, promise behaves like permit with respect to the control problem. The same holds for promising something unfavourable, namely for threatening (Vennemann in personal communication).

(5) John droht dem Mädchen an zu kommen.
'John is threatening the girl to come.'

(5') * Dem Mädchen wird von John angedroht zu kommen.
'The girl is being threatened by John to come.'

(6) John droht dem Mädchen an kommen zu müssen.
'John is threatening the girl to have to come.'

(6') Dem Mädchen is von John angedroht kommen zu müssen.
'The girl is being threatened by John to have to come.'

Examples (3), (4) and (5) suggest another syntactic analysis for promise and threaten that also would put them into the same class as permit, namely TV/INF in Thomason 1974 and TV/IV in Partee 1975, respectively. Thus Thomason should have at least three different category assignments for promise and threaten. Planck 1976 points out that this is an undesirable consequence. He suggests "that is is not the matrix verb alone (certainly not its syntactic category), but its interplay with the structure of the complement that determines the solution to the control problem", i.e. the solution to understanding what the complement's hidden subject is. Planck shows that the control depends on the internal properties of the complement, namely the semantical information it includes, the voice (active or passive) and modal modifications included in the complement. Thus, with respect to one single verb, e.g. beg or promise, and even with permit, the control can be different. Planck notices that there are verbs that are inherently neutral with respect to control, and others that are rather fixed with regard to that, but even there we find exceptions. An example is permit:

> Lady Churchill erlaubte der Holländischen Regierung finanziell (von ihr) unterstützt zu werden.
> 'Lady Churchill permitted the Dutch government to be financially supported (by it).'

Such an ambiguity has been recognized with respect to English ask: The boy asked the teacher to leave the room. Here control is ambiguous and decided on the basis of knowledge about the situation then sentence describes.

Planck writes: "I suggest that in German, there does not even exist a class of strict promise-type matrix verbs; such items as versprechen, anbieten 'offer', versichern 'assure', geloben 'vow' all seem to be inherently neutral. Mostly it is the context of use, rather than the structure of the INF-complement, that occasionally succeeds in disambiguating the identity relations of understood complement subjects." (1976: 213).
Before I want to give a non-syntactical, but a lexical solution to the control problem, I want to point out that the argument from passivization does not hold for German, and, historically, I guess it did not hold for English in a time where accusative and dative were morphologically marked by different case endings. Even, now, when we use the morphologically marked dative in English, which is built by using to prepositionally, Thomason's argument does not hold. This gives us the hint to look for another reason for the unacceptability of (1') rather than try to find it in categorial syntax. In German we have the following passive constructions:

(7a) Den Mädchen wurde vom Lehrer erlaubt zu gehen.
* 'The girls (dative) was permitted by the teacher to go.'

(8a) Den Mädchen wurde vom Lehrer versprochen einen Ausflug zu machen.
* 'The girls (dative) was promised by the teacher to go on a school trip.'

These constructions, in German, contain den Mädchen as a dative phrase, and the number of the verb is not marked for plural because the sentence subject is not marked for plural; the sentence subject being to go and to make a school trip, respectively. In English we can imitate this construction by using the morphologically marked dative to the girls, and to reckon with English word order which requires 'subject first'.

(7E) To go was permitted to the girls by the teacher.
(8E) To go on a school trip was promised to the girls by the teacher.

The so-called English Dative-passive does not correspond to the German sentence (7a):

(7E') The girls were permitted by the teacher to go.

Historically, this seems to be a reanalysis of (7a): when dative marking was lost in English the girls, a former dative in the literal translation of (7a), was reanalyzed as a nominative, and thereby as the subject; and on the basis of this, verb congruence demanded a plural marking on the verb (see 7E'). This reanalysis carried out with respect to (8a) would yield:

(8E') ? The girls were promised by the teacher to go on a school trip.

Sentence (8E') apparently seems stranger to English speakers than (7E'), when interpreted such that the girls will go on a school trip. Passive constructions in German require that we treat the infinitive of promise and permit as a term that is the direct object in the active sentence and is the subject in the passive sentence. If we do that, then infinitives are terms; and, indeed, they can take all the positions in the sentence a direct object can take.

Further passive constructions:

(7b) Zu gehen wurde den Mädchen vom Lehrer erlaubt.
(8b) Einen Ausflug zu machen wurde den Mädchen vom Lehrer versprochen.
(7c) Vom Lehrer wurde den Mädchen erlaubt zu gehen.
(8c) Vom Lehrer wurden den den Mädchen versprochen einen Ausflug zu machen.

Further active constructions:

(9g) Der Lehrer {erlaubte / versprach} dem Jungen zu gehen.
'The teacher {permitted / promised} the boy to go.'

(9b) Dem Jungen {erlaubte / versprach} der Lehrer zu gehen.

(9c) Zu gehen {erlaubte / versprach} der Lehrer dem Jungen.

Comparing these constructions with constructions involving a direct object noun phrase (accusative phrase), we see that the infinitive phrase behaves like a direct object.

<u>passive</u>

(10a) Den Mädchen wurde vom Lehrer eine Englischstunde gegeben.
lit. * 'The girls was given an English lesson by the teacher.'
'The girls were given an English lesson by the teacher.'

(10b) Eine Englischstunde wurde den Mädchen vom Lehrer gegeben.

(10c) Vom Lehrer wurde den Mädchen eine Englischstunde gegeben.

<u>active</u>

(11a) Der Lehrer gab dem Jungen einen Stift.
'The teacher gave the boy a pencil.'

(11b) Dem Jungen gab der Lehrer einen Stift.

(11c) Einen Stift gab der Lehrer dem Jungen.

The same, of course, holds for subordinate clause word order, for which I give examples in the active voice only:

(12a) (Weil) der Lehrer dem Jungen $\left\{\begin{matrix}\text{einen Stift}\\ \text{zu geben}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{gab}\\ \text{erlaubte/versprach}\end{matrix}\right\}$.

(12b) (Weil) dem Jungen der Lehrer $\left\{\begin{matrix}\text{einen Stift}\\ \text{zu geben}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{gab}\\ \text{erlaubte/versprach}\end{matrix}\right\}$.

(12c) (Weil) $\left\{\begin{matrix}\text{einen Stift}\\ \text{zu geben}\end{matrix}\right\}$ der Lehrer dem Jungen $\left\{\begin{matrix}\text{gab}\\ \text{erlaubte/versprach}\end{matrix}\right\}$.

Topicalization in subordinate clauses is hardly possible; therefore (12b) and (12c) are rather strange, although (11b) and (11c) are perfectly allright, because in main clauses topicalization is normal. The infinitive phrase can be topicalized like every noun phrase that is a main constituent of the sentence, i.e. a term that is applied to the verb in one of its term-places.

To illustrate the points I want to make in this paper I only need one syntactic rule, namely the rule that connects verbs with the terms that fill their term-places. Infinitives and <u>that</u>-subject and -object-clauses will be treated as subclasses of terms, one subclass for infinitives and another for <u>that</u>-clauses.

The categorial rule for the application of terms to verbs is:

<u>S1</u>: If β' is an n-place verb (V^n) with the set of n term-places K, $i \in K$, and α' is a term (T), then $(\alpha', i)(\beta')$ is a (n-1)-place verb (V^{n-1}) with the set of term-places K-{i}. For this we write: $(\alpha'_{T<i>}(\beta'_{V^n})_{V^{n-1}})$.

Since the terms can be of different subcategory and we will need variables of different logical types to characterize the term places of the verbs with respect to these subcategories, we use a translation scheme rather than a translation rule corresponding with S1.

<u>Translation scheme of S1:</u>

T1: If α'' is the translation of term α' and $\lambda\nu_j \ldots \nu_m\, \beta_{\nu_j \ldots \nu_m}''$ with n variables ν_κ with the place markings $\kappa := j \ldots m$ is the translation of the n-place verb β', then the translation of $(\alpha', i)(\beta)$ is: $\lambda\nu_j \ldots \nu_i{}' \; {}'\nu_i \ldots \nu_m(\alpha''(\lambda\nu_i\, \beta_{\nu_j \ldots \nu_m}''))$, with $\nu_i{}'$ as the variable preceding ν_i and ${}'\nu_i$ as the variable following ν_i.

This we can write shorter by using κ for the set of variables following λ: $\lambda\kappa{-}\{\nu_i\}\; (\alpha''(\lambda\nu_i\, \beta_\kappa''))$.

The place marking can be skipped if we use different variables x, y, z, etc. in different places. Note, that every n-place predicate can be written in this form, namely λ followed by the n variables and by a sentence open in these variables.

Rules S1 and T1 have been constructed in Bartsch 1975 and 1976a to be able to deal with "free" word order and cases in German. The System of Montague 1974 and Thomason 1974 could be adapted for the treatment of cases and "free" word order, but on the cost of a very complicated syntax and lexicon, where for every word order the order of application of the verb to the terms would be different and thereby the verb would get assigned as many categories as there are word orders. This had been worked out by Ballmer (1975). The other strategy to solve the problem would be the transformational approach, where one word order is taken to be basic and the others derived. This to me seems an unnecessary blowing up of the transformational device which rather should be as small as possible. The rule S1 proposed here, with its translation scheme handles all word orders and case assignments by one rule. It is accompanied by a rule of serialization that rewrites the hierarchical categorial structures in linear order. Serialization rules are language specific, but there are universal properties of serialization (cf. Vennemann 1973 and 1977).

<u>Rule of serialization:</u> (for German)

If a categorial structure consists out of terms and a verb the hierarchy of application $(\alpha'(\beta'))$ will be realized as $\alpha\beta$. (This is the general case; now the marked cases follow.)

If the structure is to be realized as a main clause then:

(1) If the term that is the highest in the hierarchy is to be topicalized, then the verb is inserted directly behind it.

(2) If there is no topicalization, the verb is preposed and es is put in front.

(3) If the term α' is sentential ("that"-clause) then $\alpha'(\beta')$ is serialized as $\beta\alpha$.

(4) If α' is an infinitival term, then $\alpha'(\beta')$ may be serialized as $\beta\alpha$.

Example:

(13) (Weil) der Junge dem Vater den Apfel gibt.
(unmarked)
lit.*'Because the boy Father the apple gives.'
Der Junge gibt dem Vater den Apfel.
'The boy gives Father the apple.'
Es gibt ein Junge dem Vater einen Apfel.
'There is a boy giving an apple to Father.'

Furthermore, to realize the categorial structure we do not only need serialization rules but also morphological rules for the realization of the term place markings as case markings on noun phrases.

Morphological rule:
If not marked differently, noun phrase term places of verbs will be realized as follows: 1 -- nominative; 2 -- accusative; 3 -- dative.

Since geb- 'give' behaves regular with respect to this rule we have

1(geb) = nominative
2(geb) = accusative
3(geb) = dative

(This need not be spelled out in the lexicon.)
The verb gedenk 'think of' behaves unregular. For this verb we need a marking in the lexicon:

2(gedenk) = genitive

The lexicon specifies with what kind of complements the term places of a verb are filled (compare the "Valenzwörterbuch der deutschen Sprache" by Helbig and Schenkel 1969). For every verb (and also adjective and noun) we have a corresponding expression of intensional logic.

<u>List of variables of different types used in the paper</u>

Type	Variable
e	$\underline{z}$, $\underline{y}$, $\underline{z}$
<e,t>	$\underline{P}$, $\underline{Q}$
<s,<e,t>>	$\mathfrak{P}$, $\mathfrak{Q}$
<s,t>, intensions of sentences	$\mathfrak{p}$
<<s,t>,t>, predicates of intensions of sentences	$\mathfrak{T}$
<<s,<e,t>>,t>, predicates of intensions of predicates of type <e,t>	$\mathscr{P}$

<u>List of term translations treated in the paper</u>

noun phrases, extensional	$\lambda\underline{P}\ldots$
infinitives, if α" is the translation of the one-place verb	$\lambda\mathscr{P}\,\mathscr{P}(\hat{}\alpha")$
subject-object sentences	$\lambda\mathfrak{T}\,\mathfrak{T}(\mathfrak{p})$

<u>Translation of the matrix verb</u>

$\underline{\text{versprech}}'_1$, V^3: $\lambda\underline{x}\mathfrak{P}\underline{y}$ $\underline{\text{versprech}}"(\underline{x},\hat{}([{}^{\vee}\mathfrak{P}](\underline{x})),\underline{y})$ (1-place controlling)
$\underline{\text{erlaub}}'$, V^3 : $\lambda\underline{x}\mathfrak{P}\underline{y}$ $\underline{\text{erlaub}}"(\underline{x},\hat{}([{}^{\vee}\mathfrak{P}](\underline{y})),\underline{y})$ (3-place controlling)
$\underline{\text{versprech}}'_2$, V^3: $\lambda\underline{x}\mathfrak{P}\underline{y}$ $\underline{\text{versprech}}"(\underline{x},\hat{}([{}^{\vee}\mathfrak{P}](\underline{y})),\underline{y})$ (3-place controlling)

examples: Hans verspricht dem Freund zu kommen
'John promises his friend to come'
Hans erlaubt dem Freund zu kommen
'John permits his friend to come'
Hans verspricht dem Freund das Auto fahren zu dürfen
'John promises his friend to be allowed to drive the car'

There is a <u>lexical schema</u> for the translation of infinitive terms:

If α" is the translation of a one-place verb, then the translation of the infinitive formed from this verb stem is:
$\lambda\mathscr{P}\,\mathscr{P}(\hat{}\alpha")$; where $\mathscr{P}$ is of type <<s,<e,t>>,t>, i.e. a variable for predicates of intensions of predicates of type <e,t>.

The morphological schema for the formation of infinitives is:
$\text{INF}(\alpha_{V}1) = (\underline{\text{zu}})\alpha\text{-}\underline{\text{en}}_{T<inf>}$, where α is the verb stem and INF the syntactic operator that denotes the operation of infinitive formation; <inf> is the subcategory marker.

example: $\text{INF}(\underline{\text{komm-}}_{V}1) = (\underline{\text{zu}})\ \underline{\text{kommen}}_{T}$.

We thus need a rule to construct infinitives from one-place verbs, on the morphological, categorial, and logical language level:

Morph, INF
If α is a V^1 then $INF(\alpha) = (\underline{zu})\gamma\text{-}\underline{en}$, where γ is the stem of the lexical verb of α.

Categorial syntax rule:
S, INF
If α' is a V^1, then $INF(\alpha')$ is an infinitive term.
We write $(INF(\alpha'_{V1})_{T<inf>})$ or shorter $\alpha'_{T<inf>}$.

Translation rule:
T, INF
If α'' is the translation of the one-place verb α', then the translation of $INF(\alpha')$ is $\lambda\mathcal{P}\,\mathcal{P}({}^\wedge\alpha'')$.

example:
$\underline{komm}'_{V1}$: $\lambda\underline{x}\ \underline{komm}''(\underline{x})$; then
$\underline{komm}'_{T<inf>}$: $\lambda\mathcal{P}\,\mathcal{P}({}^\wedge\lambda\underline{x}\ \underline{komm}''(\underline{x}))$

Example:
(14) Der Junge verspricht dem Vater zu kommen
(Weil) der Junge dem Vater zu kommen verspricht
'The boy promises Father to come'
$(\underline{der\ Junge}'_{T<1>}(\underline{dem\ Vater}'_{T<3>}(\underline{zu\ kommen}'_{T<inf>}(\underline{verspricht}'_{V3})_{V2})_{V1})_{V0})$
with $\underline{a}$ referring to the boy and $\underline{b}$ referring to Father, this translates, after λ-conversion, into: $\underline{versprech}''(\underline{a}, {}^\wedge(\underline{komm}''(\underline{a})), \underline{b})$.

Accordingly,
(15) Der Junge erlaubt dem Vater zu kommen
translates into: $\underline{erlaub}''(\underline{a}, {}^\wedge(\underline{komm}''(\underline{b})), \underline{b})$.

Both sentences, (14) and (15), the one with verspricht 'promise' and the other with erlaubt 'permit', have the same categorial structure. The difference is due to a purely lexical property of the respective verbs. In the lexicon certain properties that are syntactical on the level of logical syntax, are incorporated in the translation schema for certain classes of verbs. Thus we have a schema for first-place controlling verbs, the class of $promise_1$, that is different from that of the third-place controlling verbs, the class of permit. Some verbs are a member of both classes and thus have two translations, with respect to infinitive constructions. This is true of English ask, i.e. German bitten.

$\underline{bitt'_1}$, V^3: $\lambda \underline{x}\underline{y}\mathcal{P}$ $\underline{bitt}''(\underline{x}, \underline{y}, \hat{}([{}^\vee\mathcal{P}](\underline{x})))$ (1-place controlling)

example: Der Junge bittet den Lehrer zu kommen
'The boy asks the teacher to come'

$\underline{bitt'_2}$, V^3: $\lambda \underline{x}\underline{y}\mathcal{P}$ $\underline{bitt}''(\underline{x}, \underline{y}, \hat{}([{}^\vee\ \mathcal{P}]\ (\underline{x})))$ (2-place controlling)

example: Der Junge bittet den Lehrer kommen zu dürfen
'The boy asks the teacher to be allowed to come'
'The boy asks the teacher to come'

We also have corresponding constructions with object clauses (and subject clauses in the passive, accordingly).

$\underline{versprech'}$, V^3: $\lambda \underline{x}\,\gamma\,\underline{y}$ $\underline{versprech}''(\underline{x}, \gamma, \underline{y})$

and for the other verbs accordingly. γ is a variable of type <s,t>.

Example:
(16) Hans verspricht dem Freund, dass er kommt.

The translation of a sentential term is two types higher than γ, since it has to be applied to a verb, following S1. If α" is the translation of a sentence then the translation of the sentential term $\underline{dass}$ α", 'that α', is: $\lambda\mathcal{G}\,\mathcal{G}(\hat{}\alpha)$, where $\mathcal{G}$ is of type <<s,t>,t>.

example: $\underline{\text{verspricht, dass er kommt}}$
'promises that he will come'
$(\underline{\text{dass er kommt}}''_{T<s>}(\underline{\text{versprich}}''_{V3})_{V2})$: $\lambda\underline{xy}(\lambda\mathcal{G}\,\mathcal{G}(\hat{}\underline{\text{er kommt}}''))(\lambda\gamma\ \underline{\text{versprich}}''(\underline{x}, \gamma, \underline{y}))$
$= \lambda\underline{xy}\ \underline{\text{versprich}}''(\underline{x}, \hat{}\underline{\text{er kommt}}'', \underline{y})$.

Before I give examples of infinitive constructions with passives, and with pronouns, I want to present the list of variables and translations used in this paper.

The passive construction

Since, with regard to versprech- and erlaub- the infinitive phrase holds the same position with respect to the verb as does the accusative noun phrase, the infinitive phrase becomes the subject phrase of the passive construction, i.e. takes the position of the nominative with respect to passive verbs.
There is a morphological rule for passive formation:

Morph. Poss.

If α is the stem of an n-place verb that can form passive then $(\alpha\{{}^{PERF}_{PART}\}\ (\underline{werd}'_{HV})_{V}n)$ is the process-passive verb and $(\alpha\{{}^{PERF}_{PART}\}\ (\underline{sei}'_{HV})_{V}n)$ is the state-passive verb. $\alpha\{{}^{PERF}_{PART}\}$ is realized morphologically as $\underline{ge}$-α-$\underline{t}$, or, if marked as exception, according to the paradigma of α.

Translation rule

(T, PASS)
If α' translates as $\lambda\ldots\alpha''(\mu, \nu, \ldots)$ then $(\alpha'\{{}^{PERF}_{PART}\}(\underline{werd}_{HV})_{V}n)$ translates into $\lambda\ldots\breve{\alpha}''(\nu, \mu, \ldots)$, where $\breve{\alpha}''$ is the translation of α''.

There is a schema of meaning postulates that relates α'' with its converse $\breve{\alpha}''$.
Necessarily: $\forall\nu\mu(\breve{\alpha}''(\nu, \mu, \ldots) \leftrightarrow \alpha''(\mu, \nu, \ldots))$.

According to these schemas the passives of versprech- and erlaub- translate as:

versprochen werd', V^3: $\lambda \underline{x}\underline{y}\,\breve{\mathcal{P}}\ \underline{versproch}''(\hat{}(\,[{}^{\vee}\mathcal{P}\,](\underline{x})), \underline{x}, \underline{y})$
(2-place controlling)

erlaubt werd', V^3 : $\lambda \underline{x}\underline{y}\,\mathcal{P}\ \underline{erlaubt}''(\hat{}(\,[{}^{\vee}\mathcal{P}\,](\underline{y})), \underline{x}, \underline{y})$
(3-place controlling)

(Note, that the order of the variables after λ does not matter; they are unordered, but marked for the place by the use of different letters.)

gebeten werd', V^3 : $\lambda \underline{x}\underline{y}\,\mathcal{P}\ \underline{gebeten}''(\underline{y}, \underline{x}, \hat{}(\,[{}^{\vee}\mathcal{P}\,](\underline{y})))$
(1-place controlling)

versprochen werd', V^3: $\lambda \underline{x}\underline{y}\,\breve{\mathcal{P}}\ \underline{versprochen}''(\hat{}(\,[{}^{\vee}\mathcal{P}\,](\underline{y})), \underline{x}, \underline{y})$
(3-place controlling)

gebeten werden$^{1}_{2}$, V^3: $\lambda \mathcal{P}$ xy gebeten''(y, x, ^([$^\vee \mathcal{P}$] (x)))
(2-place controlling)

Examples: Compare (7a) and (8a) for 'promise' and 'permit' in German.

(31) Die Suppe aufzuessen wurde der Mutter (dat.) vom Vater {erlaubt / versprochen}

(31') Der Mutter (dat.) wurde vom Vater {erlaubt / versprochen} die Suppe aufzuessen

(32) Eine Stunde später zu kommen wurde das Mädchen vom Lehrer gebeten

(32') Das Mädchen wurde vom Lehrer gebeten eine Stunde später zu kommen
'The girl was asked by the teacher to come an hour later'

Control problems with pronouns in infinitive constructions

With the proper translation procedure for pronouns we will be able to handle the following examples:

(33) John verspricht {Peter / Mary} sich zu waschen
'John promises {Peter / Mary} to wash himself'

(34) John erlaubt Peter sich zu waschen
'John permits Peter to wash himself'

(35) (a) John verspricht Peter ihn zu waschen
'John promises Peter to wash him'
(b) John's Vater verspricht Peter ihn zu waschen
'John's father promises Peter to wash him'
(c) John verspricht Mary ihn zu waschen
'John promises Mary to wash him'

(37) John erlaubt sich zu kommen
'John permits himself to come'

(38) John verspricht Peter ihn waschen zu dürfen
'John promises Peter to be allowed to wash him'

(39) {Dem Vater / John} wird von Peter erlaubt ihn zu waschen
John is allowed by Peter to wash him'

(40) John wird von Peter erlaubt sich zu waschen
'John is allowed by Peter to wash himself'

Translation procedure for sich', 'himself':

(T, Reflex): sich'_T, 'himself': $\lambda \underline{P}P(\nu)$,

with $\nu := \nu_1$, and ν_1 is the variable occurring in place 1 of the verb to which himself is applied according to syntactical rule S1. Thus, we have a context sensitive translation rule for the reflexive pronoun.

Example:
$\text{sich}'_{T<2>}(\text{wasch}'_{V2})_{V1}$: $\lambda \underline{x}((\lambda \underline{P}P(\underline{x}))(\lambda \underline{y}\ \text{wasch}''(\underline{x}, \underline{y}))) = \lambda \underline{x}\ \text{wasch}''(\underline{x}, \underline{x})$;
according to T1 and (T, Reflex).

(T, PRO): Translation procedure for personal pronouns in obligue cases:
A personal pronoun translates into $\lambda \underline{P}P(\nu)$, with ν being any constant or variable with the following restrictions:
(1) ν may not stand for the variable or constant occurring in place 1 of the verb to which the personal pronoun is applied according to syntactical rule S1.
(2) ν may not stand for a variable or constant occupying the n-th place of an n-th place controlling verb.
Note: the control properties show in the logical structure of the verb, i.e. in their translations.

There are certain restrictions on possible co-referential noun phrases that are not special to pronouns in oblique cases. I tried to formulate a search procedure for noun phrases co-referential with pronouns in texts in Bartsch 1976b and 1978, that includes a hierarchy of qualification for being a noun phrase that can serve as co-referential with regard to a pronoun.

Restriction (2) for pronouns in infinitive complements takes care of the fact that in sentence (35a) and (35c) the pronoun cannot refer to John, in (35b) not to John's father, and in (36a) not to Peter, in (36c) not to Peter's father, and in (38) not to Peter.

Restriction (1) secures that pronouns in obligue cases are not understood in the way reflexive pronouns are.

With the translation procedures for pronouns and the restrictions (1) and (2) for pronouns in infinitive and complements, and in oblique cases we can handle the above examples, as well as their passive versions and more complicated examples where several such constructions are embedded into eachother, e.g. John promises Peter to allow him to wash himself; or John verspricht Peter sich zu erlauben ihn zu waschen, 'John promises Peter to allow himself to wash him'. It is not very usual that somebody allows something to himself but it is certainly not grammatically odd; maybe it is behaviourally odd; but we are not concerned with that.

Treatment of the examples:

(33')

$(\underline{John'}_{T<1>}(\{\frac{Peter}{Mary}\}_{T<3>}((INF(\underline{sich'}_{T<2>}(\underline{wasch'}_{V}2)_{V}1)_{T<inf>})(\underline{versprech''}_{V}3)_{V}2)_{V}1)_{V}0)$

After application of the translation rules and λ-conversion we receive

$\underline{versprech}''(\underline{j}, \ulcorner\underline{wasch}''(\underline{j}, \underline{j})\urcorner, \{\frac{p}{m}\})$

(34')

$(\underline{John'}_{T<1>}(\underline{Peter'}_{T<3>}((INF(\underline{sich'}_{T<2>}(\underline{wasch'}_{V}2)_{V}1)_{T<inf>})(\underline{erlaub''}_{V}3)_{V}2)_{V}1)_{V}0)$

$\underline{erlaub}''(\underline{j}, \ulcorner\underline{wasch}''(\underline{p}, \underline{p})\urcorner, \underline{p})$

(35'a)

Here, we receive two readings because ihn 'him' is ambiguously interpretable between $\nu := \underline{p}$ or some other $\nu: \neq \underline{p}$, $\nu: \neq \underline{j}$. Let: $\nu := \underline{p}$.

Then $(\underline{ihn'}_{T<2>}(\underline{wasch}_{V}2)_{V}1)$ is translated into:

$\lambda\underline{x}((\lambda\underline{P}\underline{P}(\underline{p}))(\lambda\underline{y}\ \underline{wasch}''(\underline{x}, \underline{y}))) = \lambda\underline{x}\ \underline{wasch}''(\underline{x}, \underline{p})$.

Otherwise the categorial structure and the translation procedure is like in (33').
We receive:

$\underline{versprech}''(\underline{j}, \ulcorner\underline{wasch}''(\underline{j}, \underline{p})\urcorner, \underline{p})$.

Let $\nu := r$. Then: $(\underline{ihn'}_{T<2>}(\underline{wasch}_{V}2)_{V}1)$ translates into $\lambda\underline{x}\ \underline{wasch}''(\underline{x}, \underline{r})$; and we receive $\underline{versprech}''(\underline{j}, \ulcorner\underline{wasch}''(\underline{j}, \underline{r})\urcorner, \underline{p})$ as the second reading.

(35'b)

Here, the pronoun can refer to Peter, or to John, or to somebody else, let's say r, who is not John's father. Our search- and translation procedure for pronouns gives us these three possibilities. (1): $\nu:=\underline{p}$, or (2) $\nu := \underline{j}$, or (3) $\nu := \underline{r}$, $\underline{r}$ does not refer to John's father. If $\underline{f}$ is the constant referring to John's father, we receive:

(1) versprech"(f, ^wasch"(f, p)), p)
(3) versprech"(f, ^wasch"(f, j)), p)
(4) versprech"(f, ^wasch"(f, r)), p)

(35'c)
Here, the gender marking of Mary excludes reference of ihn 'him' to Mary. Thus the only possibility is that ihn refers to a third person, let's say to the one referred to by r in the context of speech. Then we receive from the categorial structure and the translation procedure:
versprech"(j, ^wasch"(j, r)), m).

(36'a)
The categorial structures are not syntactically different from those above. I will present then in a shortened version:
(a) $\text{John}'_{T<1>}(\text{Peter}'_{T<3>}(\text{ihn zu waschen}'_{T<inf>}(\text{erlaub}''_{V3})_{V2})_{V1})_{V0})$.
Here, according to our pronoun search and translation procedure, ihn 'him' can be translated by using ν: = p, or by some other constant or variable with ν: ≠ p, ν: ≠ j. Thus, if ν: = p, we receive
erlaub"(j, ^wasch(p, j)), p).
If ν: = r, gotten from the context by the search procedure, we receive
erlaub"(j, ^wasch"(p, r)), p).

(36'b)
Here ν: = m is excluded by the gender restriction and thus we only receive a reading with some constant found from the context, with ν: ≠ p. If that constant is r, we get:
erlaub"(m, ^wasch"(p, r)), p).

(36'c)
Here ν: = j, or ν: = p, or ν: = r, with r as some constant from the context and r ≠ f, if f refers to Peter's father.
(1) erlaub"(j, ^wasch(f, j)), f)
(2) erlaub"(j, ^wasch(f, p)), f)
(3) erlaub"(j, ^wasch(f, r)), f)

(37')
The categorial structure is
$(\text{John}'_{T<1>}(\text{sich}'_{T<3>}(\text{zu kommen}'_{T<inf>}(\text{erlaub}'_{V3})_{V2})_{V1})_{V0})$
zu kommen erlaubt translates into λxy erlaub"(x, ^komm"(y)), y).

Then sich zu kommen erlaubt translates into: $\lambda \underline{x}\ \underline{erlaub}'(\underline{x}, {}^\wedge\underline{komm}''(\underline{x}), \underline{x})$, according to the translation rule for reflexive pronouns.
For the whole sentence we then receive: $\underline{erlaub}''(\underline{j}, {}^\wedge\underline{komm}''(\underline{j}), \underline{j})$.

(38)
Here, we have two infinitives, zu dürfen and waschen without zu 'to', since darf- is a modal verb. The English translation is structurally different from the German since it uses the passive of permit or allow accompanied by an infinitive with to. Here we use the dative controlling $\underline{versprech}'_2$.
$(\underline{John}'_{T<1>}(\underline{Peter}_{T<3>}(INF(INF(\underline{ihn}'_{T<2>}(\underline{wasch}'_{V2})_{V1})_{T<inf>})(\underline{darf}_{V2})_{V1})_{T<inf>})(\underline{versprech}''_{2_{V3}})_{V2})_{V1})_{V0})$.

The pronoun ihn can be translated by $\nu := \underline{j}$, or some other constant, say $\underline{r}$, $\underline{r} \neq \underline{p}$, received from the context. Let us choose $\nu := \underline{j}$.
$\underline{\text{ihn waschen}}_{T<inf>}: \lambda \mathcal{P}\, \mathcal{P}\, ({}^\wedge(\lambda \underline{z}\ \underline{wasch}''(\underline{z}, \underline{j})))$

$\underline{\text{ihn zu waschen darf}}_{V1}: \lambda \underline{x}\ \lambda \mathcal{P}\, \mathcal{P}\, ({}^\wedge(\lambda \underline{z}\ \underline{wasch}''(\underline{z}, \underline{j})))(\lambda \varphi\ PER({}^\wedge(([{}^\vee \varphi](\underline{x})))) = \lambda \underline{x}\ \underline{PER}({}^\wedge(\underline{wasch}''(\underline{x}, \underline{j})))$

$\underline{\text{ihn waschen zu dürfen}}_{T<inf>}: \lambda \mathcal{P}\, \mathcal{P}({}^\wedge\lambda \underline{x}\ \underline{PER}({}^\wedge(\underline{wasch}''(\underline{x}, \underline{j}))))$

$\underline{\text{ihn waschen zu dürfen verspricht}}_{2_{V2}}: \lambda \underline{xy}\ \underline{versprech}''(\underline{x}, {}^\wedge\underline{PER}({}^\wedge(\underline{wasch}''(\underline{y}, \underline{j}))), \underline{y})$

The whole sentence then translates into:
$\underline{versprech}''(\underline{j}, {}^\wedge\underline{PER}({}^\wedge(\underline{wasch}''(\underline{p}, \underline{j}))), \underline{p})$

Although we have two verbs versprech' 'promise' in German which differ in their control properties (first-place controlling versus third-place controlling), there is only one 3-place predicate
versprech" in intensional logic. If we would translate sentence (38) by using $\underline{versprech}'_1$ we would get a formula that is well-formed but that is semantically odd because it says that John promised Peter that he himself (= John) would be permitted to wash him (= Peter).

The translation of sentence (38) with a reflexive instead of the non-reflexive ihn is straight forward, the only difference being that $\underline{\text{sich wasch}}'_{V1}$ translates as $\lambda \underline{z}\ \underline{wasch}''(\underline{z}, \underline{z})$. Thus, for John verspricht Peter sich waschen zu dürfen 'John promises Peter to be allowed to wash himself' we get
$\underline{versprech}''(\underline{j}, {}^\wedge\underline{PER}({}^\wedge(\underline{wasch}''(\underline{p}, \underline{p}))), \underline{p})$.

(39')

Passive formation in German changes first and second place in the formation of the converse relation. There is no change involved in the third place, the dative marked one. Thus, the passive verb erlaubt werd' is, like the active verb erlaub', third-place controlling. This means that ihn may be translated by any constant or variable except the one standing for John.

$(\underline{\text{John}}'_{T<3>}(\underline{\text{von Peter}}'_{T<2>}(\underline{\text{ihn zu waschen}}'_{T<inf>}(\underline{\text{erlaubt wird}}'_{V3})_{V2})_{V1})_{V0})$.

Since the details of the passive construction have not been worked out we treat the agentive phrase simply as a term in place two of the passive verb. Let us assume $\nu := \underline{p}$. With that $\underline{\text{ihn zu waschen}}_{T<inf>}$ translates into:

$\lambda\mathcal{P}\ \mathcal{P}(\hat{}\lambda\underline{x}\ \underline{\text{wasch}}''(\underline{x}, \underline{p}))$.

The infinitive phrase takes the first place of the passive verb erlaubt werd". The translation amounts to:

$\underline{\text{erlaubt werd}}''(\hat{}(\underline{\text{wasch}}''(\underline{j}, \underline{p}), \underline{p}, \underline{j})$.

Using the meaning postulate mentioned above, this is equivalent to

$\underline{\text{erlaub}}''(\underline{p}, \hat{}(\underline{\text{wasch}}''(\underline{j}, \underline{p}), \underline{j})$.

(40')

Here we have $\underline{\text{sich zu waschen}}_{T<inf>}$ and therefore: $\lambda\mathcal{P}\ \mathcal{P}(\hat{}\lambda\underline{x}\ \underline{\text{wasch}}''(\underline{x}, \underline{x}))$.

Since the passive verb erlaubt werd' is third-place (dative) controlling), the translation of sich zu waschen erlaubt werd' is (see schema for passives)

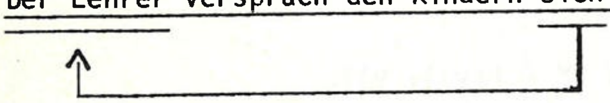
$\lambda\underline{xy}\ \underline{\text{erlaubt werd}}''(\hat{}(\underline{\text{wasch}}''(\underline{y}, \underline{y})), \underline{x}, \underline{y})$.

The whole sentence translates into:

$\underline{\text{erlaubt werd}}''(\hat{}(\underline{\text{wasch}}''(\underline{j}, \underline{j}), \underline{p}, \underline{j})$.

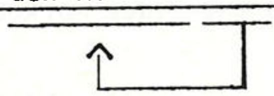
Note, that we get two different translations for the reflexives, according to whether we have a $\underline{\text{versprech}}_1$, which has first-place control or a $\underline{\text{versprech}}_2$, which has third-place control. Thus we have:

Der Lehrer versprach den Kindern sich zu waschen



and

Der Lehrer versprach den Kindern sich waschen zu dürfen



To be able to treat the examples given by Jan Anward (in personal communication) for Swedish, where himself can refer to John in a sentence of the form lit. John permits Mary to wash himself, which is acceptable in Swedish, we have to provide for the possibility to apply an infinitive of a two-place verb, e.g. wash, to permit. In this way we receive permit to wash. Syntactically this was formulated in the notation of Planck 1975.

Syntactical rule

If α is an m-place infinitive and β an n-place verb taking infinitive complements then $\alpha(\beta)$ is an (n+m)-place verb.

Note that from an n-place verb we get an (n-1)-place infinitive term by applying the infinitive operation, for example

$\underline{tanz}'_{V1}$, $(INF(\underline{tanz}'_{V1})_{T<inf>})$ with 0 term places.

$\underline{wash}'_{V2}$, $(INF(\underline{wash}'_{V2})_{T1<inf>})$ with 1 term place.

(T, INF): If $\lambda\nu_1\ldots\nu_m\ \alpha''_{\nu_1\ldots\nu_m}$ is an m-place verb then

$\lambda\nu_2\ldots\nu_m\ \mathcal{P}\,(\ \mathcal{P}\,(\hat{}\lambda\nu_1\ \alpha''_{\nu_1\ldots\nu_m}))$ is the (m-1)-place infinitive.

(T, INF, COMP): If $\lambda\nu_2\ldots\nu_m\ \mathcal{P}\,(\ \mathcal{P}\,(\hat{}\lambda\nu_1\ \alpha''_{\nu_1\ldots\nu_m}))$ is the translation of the (m-1)-place infinitive and $\lambda\mu_1\ldots\mu_n\ \beta''_{\nu_1\ldots\nu_m}$ is the translation of the n-place verb β' taking an infinitive complement in its i-th place, then the translation of $\alpha'(\beta')$ is:

$\lambda\nu_2\ldots\nu_m\ \mu_1\ldots\mu'_i\ '\mu_i\ldots\mu_j(\lambda\mathcal{P}\mathcal{P}(\hat{}\lambda\nu_1\ \alpha''_{\nu_1\ldots\nu_m}))(\lambda\mu_i\ \beta''_{\mu_1\ldots\mu_m})$.

Here: $\mu_i := {}^{\vee}\mathcal{P}$.

Using this rule, we can build up

$\lambda\underline{xyz}(\lambda\mathcal{P}\mathcal{P}(\hat{}\lambda\underline{u}\ \underline{wash}''(\underline{u}, \underline{z})))(\lambda\ {}^{\vee}\mathcal{P}\ \underline{permit}''(\underline{x}, \hat{}(\ [{}^{\vee}\ {}^{\vee}P\ ](\underline{y})), \underline{y}))$.

This, after λ-conversion, amounts to

$\lambda\underline{xyz}\ \underline{permit}''(\underline{x}, \hat{}(\underline{wash}''(\underline{y}, \underline{z})), \underline{y})$.

If we apply himself' to this with respect to the place marked by z we get

$\lambda\underline{xy}\ \underline{permit}''(\underline{x}, \hat{}(\underline{wash}''(\underline{y}, \underline{x})), \underline{y})$,

and then with John as T<1> and Mary as T<2> we receive the translation

$\underline{permit}(\underline{j}, \hat{}(\underline{wash}''(\underline{m}, \underline{j})), \underline{m})$

as a translation for the literal English equivalent for the Swedish construction.

The translation for infinitives gives us a form that can be used to apply terms to. The places marked by $\nu_2 \ldots \nu_m$ can be filled out by applying terms to the infinitive in these places. To do this a simple extension of rules S1 and T1, which are formulated for the application of terms to finite verbs, can be used. This provides us with an alternative to the procedure adopted in this paper, where we constructed one-place verbs first before we applied the infinitive operation. We constructed the infinitive term ihn wash" as $(\underline{\mathrm{INF}}(\underline{\mathrm{ihn}}'_{T<2>}(\underline{\mathrm{wash}}'_{V}2)_{V1})_{T<inf>})$. But we could also do it by applying INF to the lexical verb and then apply the terms: $(\underline{\mathrm{ihn}}'_{T<2>}(\underline{\mathrm{INF}}(\underline{\mathrm{wash}}_{V}2)_{T1<inf>})_{T<inf>})$.

Conclusion

The main difference between Groenendijk and Stokhof's (1977) treatment of infinitives and mine is that they introduce a special kind of variables, context variables, that are interpreted in a certain way in a model, such that in the context of promise the context variable denotes the entity introduced into the context (within the model) by the interpretation of the subject phrase of promise and in the context of permit the denotation of the context variables is the entity introduced into the context by the interpretation of the indirect object phrase. What Groenendijk and Stokhof do within the semantics, i.e. within a model and the context dependent interpretation of expression in the model, I do in the translation procedure. Thus I keep the model theory simpler, but make the translation into intensional logic more complicated. The reason why I do that is, that I need surface word order information for the interpretation of pronouns in general and this information can get lost by translating sentences into intensional logic, except if we change the syntax of intentional logic in such a way that the linear order of expressions is going to play a role and that it mirrors surface order such that, for example, p & q is not anymore equivalent to q & p in the moment q and p contain expressions that are interpreted context dependently. Another difference between mine and Groenendijk & Stokhof's proposal is that I treat infinitives not as the characteristic set of properties of certain propositions but as the characteristic set of properties of properties, like Thomason. But other than Thomason, I do not need meaning postulates to account for the paraphrases between the infinitive constructions and the corresponding clause constructions. Rather, the propositional structure attributed to infinitives by Groenendijk and Stokhof is, in my proposal, not attributed to the infinitive phrases themselves but to the internal structure of the matrix verb.

In the translation procedure the translation of an infinitival phrase is integrated into the propositional structure of one of the term-places of the matrix verb. Thus, I have two categorizations for the matrix verb in the categorial grammar of German (or English), one taking an infinitive and one taking a sentential complement, but only one type for the matrix verb on the level of the syntax of intentional logic, and of course only one lexical content analysis of the verb. I do not derive infinitives from imbedded propositions and I do not derive pronouns and reflexives from artificial syntactic variables as in Bennett's (1976) and Thomason's (1976) analysis.Rather, I construct infinitive constructions directly on the level of categorial syntax; and I derive the infinitives morphologically from the verb-stems; and pronouns, including reflexive pronouns, are basic expressions, except for morphological and morphophonemic case, number, and gender modifications. I could also imagine a modification of my proposal in the way that the infinitive is translated by a sentence open in the subject position, such that, for example, John promises Mary to come would translate into promise"(j, ^(come"(x)), m). If we do that we need to restrict the possible assignment functions when interpreting phrases like this in a model. The arguments of Groenendijk and Stokhof (1977) against this course are not decisive since they assume that restriction to be unduely heavy. How strict the restriction is depends on the kind of context. Thus for the sentence mentioned it is stricter than for Peter promises John that he would come.

With regard to the sentences with that- complements the requirement is fairly loose: The sentence ... is true if there is an assignment g with $g(\underline{x}) \neq F(\underline{j})$ such that $[\ldots]^{g\ldots} = 1$.Thus $g(\underline{x})$ might be Peter or someone else. But even that is too strict since we can interpret the above sentence even with $g(\underline{x}) = F(\underline{j})$ if the promise is a kind of threat or of it means that Peter promises John that he would be invited. Here, it is John who is the referent of he. The same is not possible for the infinitive construction: there $g(\underline{x})$ cannot be somebody different from John or Mary; in the first place it must be John; but we have discussed examples where it can be otherwise, for example, John promises Mary to be invited. Here, $g(x)$ more likely will be Mary and not John. Thus, as Planck pointed out, control should be subject to context properties. These context properties can be formulated as restrictions on the assignment functions used to interpret variables. If we take this course we need not have in the lexicon $\underline{\text{promise}}_1$ and $\underline{\text{promise}}_2$, that differ in their control-properties. Then we can do with one promise in natural language, like we can, anyhow, in the syntax of intensional logic. But the drawback of this position has been already mentioned: we need to consider surface properties to formulate the restrictions on the assignment functions, if we want to use this method for the interpretation of all pronouns. Therefore, besides structures of categorial syntax we need to formalize the surface properties that are relevant for the choice of the proper kinds of assignment functions and store them in a component parallel with and referring to the categorial structure. This can be skipped if we use all this relevant surface information directly in the translation procedure into intensional logic, as done in Bartsch 1976, 1978, and in this paper.

The point of the lexical treatment given here is, that there is a fair amount of logical structure included in the lexical information of lexical items, which we know when we know the meaning of a lexical item. This information need not be part of the syntax of natural language and thus the syntactical structures of natural language can be simpler than they would be if all that lexical knowledge would be explicit in the categorial syntax of natural language sentences.

Nevertheless, we have lexical information of this kind available when interpreting sentences. In the logical language this kind of lexical knowledge is made explicit in the syntactic forms. The alternative would be to make quite complicated context dependent rules of interpretation in a model that express these properties. The third way, namely to force these differences into the syntax of natural language, where they do not appear, in fact, I reject. Thus I reject making an artificial natural language syntax with never realized artificial pronouns like $\underline{he}_n$ in Montague grammar, and further employed by Partee 1975, or with traces as in transformational grammar, e.g. in Chomsky 1976.

REFERENCES

Ballmer, Th. (1975), Sprachrekonstruktionssysteme und einige ihrer Anwendungsmöglichkeiten in Satz- und Textlinguistik. Scriptor Verlag, Kronberg, Ts. 1975.

Bartsch, R. (1975), Topik-Fokus-Struktur und kategoriale Syntax. In: Ehrich and Finke (eds.) 1975: 85-100.

Bartsch, R. (1976a), The role of categorial syntax in grammatical theory. In: A. Kasher (ed.) 1976, Language in focus: foundations, methods and systems. Essays in memory of Y. Bar-Hillel. De Reidel, Dordrecht-Holland 1976: 503-540.

Bartsch, R. (1976b), Syntax and semantics of relative clauses. In: Groenendijk and Stokhof (eds.), 1976: 1-24.

Bartsch, R. (1978), The syntax and semantics of subordinate clause constructions and pronominal coreference. To appear in: F. Heny and H. Schnelle (eds.) Syntax and semantics, vol. 10: Selections from the 3rd Groningen Round Table. Seminar Press, New York, San Francisco.

Chomsky, N. (1976), Reflections on Language. Pantheon Books, Random House, New York.

Ehrich, V. and P. Finke (eds.) 1975, Beiträge zur Grammatik und Pragmatik. Scriptor Verlag, Kronberg, Ts. 1975.

Groenendijk, J. and M. Stokhof (eds.) 1976, Amsterdam papers in formal grammar, vol. I. Papers of the symposium 'Montague grammar and related topics', held in Amsterdam on January 12-16, 1976. Centrale Interfakulteit, Universiteit van Amsterdam.

Groenendijk J. and M. Stokhof (1977), Infinitives and context in Montague grammar. Ms. Centrale Interfaculteit, University of Amsterdam. To appear in the Proceedings of the conference on Montague grammar, philosophy and linguistics, held at the State University of New York ar Albany, April 23-24, 1977.

Helbig, G. and W. Schenkel (1969), Wörterbuch zur Valenz und Distribution deutscher Verben. VEB Bibliographisches Institut Leipzig.

Montague, R. (1974), The proper treatment of quantification in ordinary English. In: Thomason (ed.) 1974: 247-270.

Partee, B. (1973), Some transformational extensions of Montague grammar. In: Journal of philosophical logic, Vol. 2, no. 4: 509-534.

Partee, B. (1975), Montague grammar and transformational grammar. In: Linguistic inquiry 6: 203-300.

Planck, F. (1975), Ein Verb ist ein Verb ist ein Verb? Zur kategorialen Syntax von Infinitiven. In: Ehrich and Finke (eds.) 1975: 111-132.

Planck, F. (1976), Misunderstanding understood subjects: the minimal distance principle in Montague grammar. In: Groenendijk and Stokhof (eds.), 1976: 194-216.

Thomason, R. (1974), Introduction to Thomason (ed.) 1974: 1-70.

Thomason, R. (1974), Formal philosophy. Selected papers of Richard Montague. Yale University Press, New Haven and London.

Vennemann, T. (1973), Explanation in syntax. In: J. Kimball (ed.), 1973, Syntax and semantics, vol. 2. Seminar Press, New York and London: 1-50.

Vennemann, T. (1977), Categorial grammar and the order of meaningful elements. In: A. Juilland (ed.), Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday. Anma Libri: Saratoga, Cal.: 615-634.

FORMAL TREATMENT OF TENSES AND TEMPORAL CLAUSES

Pieter Uit den Boogaart, University of Technology,
Eindhoven, Netherlands

In this paper I will present a number of formal systems, L_0 *through* L_8 *. The semantics of each of them is meant to formalize partially the way temporal clauses behave in natural languages. As only temporal aspects of sentence building are taken into consideration, the syntactic parts of the systems are very simple, if not trivial. The semantic parts, on the other hand, are far from simple, at least in the cases of the latest three systems. Their relative complexity is, as I hope,justified by the advantage of having systems in which al kinds of temporal clauses are treated as meaningful units of the same semantic type. I include tenses in the variety of possible temporal clauses.*

L_8 *might be the "logical" framework for temporal constructions in formal languages that are more similar to natural languages - just like intensional logic is the logical base for Montague's fragments of ordinary English. I hope I will be able in due time to present some formalized fragments of ordinary Dutch language that have a kind of* L_8*-semantics: at this stage,however, I restrict myself exclusively to the more esotheric systems with syntaxes simple but hardly adapted to the demands of ordinary language usage.*

The formal definitions of the systems consist of five parts: the declaration of the alphabet, the declaration of the names of syntactic classes, the list of the syntactic rules, the definition of what models are supposed to be, and,finally, the list of the semantic rules which are cuonterparts of the syntactic rules.

In all of the systems there is a syntactic category called "formulae". In L_0 *the semantic rules assign truth values* 0 *and* 1 *to formulae for falsehood and truth respectively. In other systems formulae are assigned "indexed truth values", i.e. mappings from the set of moments into truth values.*

Throughout the paper, we will use the symbol U *for referring to the set of moments. This set is taken to be isomorphic to the set of real numbers with respect to order and with respect to distance measure. This metaphysical point of view is not beyond philosophical criticism. I believe, however, it is pretty harmless in the present context.*

The power set of a given set S *will be referred to with* 2^S .
Particularly,2^U *denotes the set of subsets of* U .

The system L_0

1 Alphabet

The L_0-alphabet consists of two kinds of variables: term variables $\tau_0, \tau_1, \tau_2, \ldots$ *and proposition variables* $p_0, p_1, p_2, \ldots$.
In the sequel we will use symbols for sets of variables:

$\overline{\overline{P}}$ *will denote the set of proposition variables, and* $\overline{\overline{T}}_0$ *the set of term variables.*

2 Syntactic Categories

L_0 *has three syntactic categories:* L_0*-propositions,* L_0*-formulae and* L_0 *- terms.*
In this system and the following, we will speak simply of propositions, formulae and terms, instead of L_0*-formulae,* L_0*-terms and so on. From the context must be clear what kind of formulae or terms we have in mind. The same is valid for those syntactic categories that will be developed later on.*
We will use $\overline{\overline{F}}$ *as denoting the set of formulae, and* $\overline{\overline{T}}$ *as denoting the set of terms. (Of course, in different contexts, the above symbols may have different denotations).*

3 Syntactic rules

The expressions of L_0 *are generated completely by the following rules:*
0-Synt*(i)* $\overline{\overline{P}}$ *is the set of propositions;*
0-Synt*(ii)* $\overline{\overline{T}}=\overline{\overline{T}}_0$;
0-Synt*(iii) if* $\tau\in\overline{\overline{T}}$, *and* $\alpha\in\overline{\overline{P}}$, *then* $\tau\alpha\in\overline{\overline{F}}$.

4 Models

A L_0*-model is a pair* (σ_1,σ_2) *such that* $\sigma_1:P\to\{0,1\}^U$ *and* $\sigma_2:T\to U$

5 Interpretation functions

Let $\mathfrak{M}$ *be the* L_0*-model* (σ_1,σ_2).
The denotation function $\mathfrak{M}$ *is the triple* $(\zeta_1,\zeta_2,\zeta_3)$ *defined as follows:*
0-Sem *(i)* $\zeta_3=\sigma_1$
0-Sem *(ii)* $\zeta_2=\sigma_2$
0-Sem *(iii) if* $\tau\in\overline{\overline{T}}$, *and* $\alpha\in\overline{\overline{P}}$, *then* $\zeta_1(\tau\alpha)=\zeta_3(\alpha)(\zeta_2(\tau))$.

Remarks

Of course, the rules could be given far more simply than we have done now. The way they are presented is uniform with the presentations of the other systems.
In a more relaxed style we can say that formulae consist of a term variable and a proposition variable, that the proposition variable denotes a function from moments into truth values, that the term variable denotes a moment, and that the denotation of the formulae is yielded by application of the denotation of the proposition to the denotation of the term. So, with respect to a model that assigns to α *a function like* $\lambda t\in U$.*it is raining at the moment* t, *and to* τ *the moment 3-10-'77, 12.00 h.,* $\tau\alpha$ *will be true if and only if it actually rained on that specific moment.*
So far our system L_0 . *It only serves as an introduction into the next systems.*

The System L_1

1 Alphabet

The L_1-alphabet is just the set of proposition variables $\bar{\bar{P}}$.

2 Syntactic Categories

L_1 *has only one syntactic category: the set of* (L_1)*-formulae* $\bar{F}$.

3 Syntactic Rule

The expressions of L_1 *are generated completely by* 1-Synt(*i*) $\bar{\bar{P}}=\bar{F}$.

4 Models

An L_1*-model is a mapping* $\mathfrak{M}$ *such that* $\mathfrak{M}:\bar{\bar{F}}\to\{0,1\}^U$

5 Semantic Rule

Let $\mathfrak{M}$ *be an* L_1*-model. We associate to it the* (L_1)*-denotation function* $\tilde{\mathfrak{M}}$*, defined by*

1-Sem(*i*) $\tilde{\mathfrak{M}}=\mathfrak{M}$

Remarks

Of course, L_1 *is practically worthless for almost all purposes that can be imagined. We have constructed this system in order to show the fundamental behavior of time dependent sentences: they are true or false with respect to a given moment of time - this moment may be considered as the moment of utterance. In the sequel we will use meta-expressions like "α is true at (the moment) t". Statements like these presuppose a given, temporarily fixed, model* $\mathfrak{M}$*. Their sense is obviously the following: "(The first coordinate of)* $\tilde{\mathfrak{M}}$ *assigns to* α *a function that assigns* 1 *to* t *".*

The System L_2

1,2,3 Alphabet, Syntactic Categories, Syntactic Rules

Like in L_0.

4 Models

An L_2*-model is a pair* (σ_1,σ_2) *such that* $\sigma_1:\bar{\bar{P}}\to\{0,1\}^U$ *and* $\sigma_2:\bar{\bar{T}}\to U$.

5 Semantic Rules

Let $\mathfrak{M}$ *be the* L*-model* (σ_1,σ_2)*. We associate to it the* (L_2)*-denotation function* $\tilde{\mathfrak{M}}=(\zeta_1,\zeta_2,\zeta_3)$*, defined as follows.*

2-Sem(*i*) $\zeta_3=\sigma_1$

2-Sem(*ii*) $\zeta_2=\sigma_2$

2-Sem(*iii*) $\zeta_1:\bar{F}\to\{0,1\}^U$ *and*

if $\tau\in\bar{\bar{T}}$ *and* $\alpha\in\bar{\bar{P}}$ *then* $\zeta_1(\tau\alpha)(t)=\zeta_3(\alpha)(\zeta_2(\tau))(t))$.

Remarks

With respect to their syntaxes, L_2 *and* L_0 *are the same languages. Semantically, they differ in that formulae of* L_2 *have truth values depending on the moment of utterance. So the semantic type of* L_2*-formulae is equal to the semantic type of formulae in* L_1 *. This will be the case in each of the remaining systems.*

The dependency on time of the truth value of L_2*-formulae is made possible by the dependency on time of the denotation of terms. Terms in* L_2 *refer to moments, but it depends on the moment of utterance which moment is referred to by a given term* τ *with respect to a given model* $\mathfrak{M}$ *. Hence the meanings of terms in* L_2 *are conceived of as mappings from moments (of utterance) to moments (of reference).*

Now, let ϕ *be a function from moments to moments, equal to* $\sigma_2(\tau)$.
σ_2 *is intended to be the second coordinate of an* L_2*-model, and* τ *to be a term. We give a few examples of possible properties of* ϕ:

1. ϕ *may be monotously increasing and continuous.*
   In that case ϕ *may correspond to clauses in ordinary language like "exactly one year ago", "exactly five hours ago"(or "now", in which case* ϕ *assigns to each moment itself).*
2. ϕ *may be monotously increasing and discrete (i.e. constant in separate intervals).Here we think of functions corresponding to clauses like "yesterday at noon", "the latest time it started raining".*
3. ϕ *may be constant. In that case the term* τ *refers to one moment, say, 12th of October, 1977, 10.00 h. p.m.,independently of the moment of utterance. Of course, if* L_2*-models were restricted to those having constant functions in their second coordinate, the system* L_2 *would be essentially the same as* L_0.

One of the disadvantages of L_2 *is that formulae have an obligatory "temporal clause". We would like to get rid of this rather unnatural feature. This will be done in our system* L_3.

The System L_3

1 Alphabet

Like in L_0

2 Syntactic Categories

L_3 *has two syntactic categories:* (L_3)*-formulae and* (L_3)*-terms.*

3 Syntactic Rules

The expressions of L_3 *are completely generated by the following rules:*

3-Synt*(i)* $\bar{\bar{\bar{P}}} \subset \bar{\bar{\bar{F}}}$

3-Synt*(ii)* $\bar{\bar{\bar{T}}}_0 = \bar{\bar{\bar{T}}}$

3-Synt*(iii)* *if* $\tau \in \bar{\bar{\bar{T}}}$ *and* $\alpha \in \bar{\bar{\bar{F}}}$, *then* $\tau\alpha \in \bar{\bar{\bar{F}}}$

4 Models

An L_3*-model is a pair* (σ_1, σ_2), *satisfying the conditions of an* L_2*-model.*

5 Semantic Rules

Let $\mathfrak{M}$ *be the* L_3 *-model* (σ_1,σ_2). *We associate to it the* (L_3)*-denotation function* (ζ_1,ζ_2), *defined by the following rules:*

$\zeta_2:\bar{\bar{F}}\to\{0,1\}^U$ *and*

3-Sem*(i)* $\zeta_1\subset\sigma_1$

3-Sem*(ii)* $\zeta_2=\sigma_2$

3-Sem*(iii)* *if* $\tau\in\bar{\bar{T}}$ *and* $\alpha\in\bar{\bar{F}}$ *then* $\zeta_1(\tau\alpha)=\zeta_1(\alpha)\circ\zeta_2(\tau)$.

(Here the sign "$\circ$" is used for function composition.)

In order to avoid ambiguities, we shall use the symbol $\tilde{\tilde{\mathfrak{M}}}$ *instead of* $\tilde{\mathfrak{M}}$ *for the denotation function depending on* $\mathfrak{M}$, *if* $\mathfrak{M}$ *is to be considered as an* L_3*-model.*

Remarks

L_3 *is the first system where recursive definitions of syntactical categories and semantic functions are applied. It is an extension of* L_2. *Each formula consists of a (possibly empty) sequence of terms, followed by a proposition variable.*

We show now that L_3 *is an extension of* L_2 *with respect to its semantics too. Indeed, let* $\mathfrak{M}$ *be an* L_2*-model (thus also an* L_3*-model), and let* α *be an* L_2*-formula (and hence an* L_3*-formula too). Furthermore, let* (ζ_1,ζ_2) *and* (ζ_1',ζ_2') *be the* L_2 *-denotation function and the* L_3*-denotation function associated to* $\mathfrak{M}$ *respectively.*

α *must be of the form* τp , τ *being a term and* p *a proposition variable.*
We are going to show that $\zeta_1'(\alpha)=\zeta_1(\alpha)$:

$\zeta_1'(\alpha)=\zeta_1'(\tau p)=\zeta_1'(p)\circ\zeta_2'(\tau)=\zeta_1(p)\circ\zeta_2(\tau)=\lambda t\in U.\zeta_1(p)(\zeta_2(\tau)(t))=$
$\lambda t\in U.\zeta_1(\tau p)=\zeta_1(\alpha)$.

It is not difficult to show that a sequence of terms corresponds semantically to the reverse composition of the "meanings" of the terms. Let $\mathfrak{M}$ *be a model, and* (ζ_1,ζ_2) *the denotation function associated to* $\mathfrak{M}$. *For terms* τ_1 *and* τ_2, *we assume* $\zeta_2(\tau_1)=\phi$ *and* $\zeta_2(\tau_2)=\chi$. *Let* τ_3 *be a third term, such that* $\zeta_2(\tau_3)=\chi\circ\phi$

Now for each formula α, $\zeta_1(\tau_1\tau_2\alpha)=\zeta_1(\tau_2\alpha)\circ\phi=\zeta_1(\alpha)\circ\chi\circ\phi=\zeta_1(\alpha)\circ\zeta_2(\tau_3)=\zeta_1(\alpha\tau_3)$.

It may be worth noticing that in many cases composition of term meanings is commutative. Two terms with meanings like "one year ago" and "two years ago" will have the combined effect of "three years ago" ; the order of the terms is here irrelevant. But, as function composition is basically not commutative, we may expect cases too where meaning is changed by changing order. This happens for example to pairs of terms of which one is assigned a constant value. Whenever ϕ *and* χ *are meanings of terms, while* ϕ *is a constant function,* $\phi\circ\chi$ *will be equal to* ϕ − $\chi\circ\phi$ *will be a constant too, but in general different from* ϕ . *This consider ation mirrors our intuition that sentences like "a year ago it rained at noon on the 29th of June in 1960" are equivalent to the same sentence without the redundant "a year ago". On the other hand, the sentence "at noon on the 29th of June in 1960 it rained a year ago" is rather equivalent to "it rained at noon of the 29th of June in 1959."*

Our next system will be essentially richer than the system we have been studying now. Although syntactically the same as L_3, L_4 *is able to express temporal quantifiers like "always" and "during an hour".*

The system L_4

Some preliminaries

Here and in the sequel we will use some abbrevations.
So we define: $W=\{0,1\}^U$;
$W^+=W^W$.

1,2,3 Alphabet, Syntactic Categories, Syntactic Rules:

Like in L_3.

4 Models

An L_4*-model is a pair* (σ_1,σ_2) *such that*

$\sigma_1:\bar{\bar{P}}\to W$ *and* $\sigma_2:\bar{\bar{T}}\to W^+$.

5 Semantic Rules

Let $\mathfrak{M}$ *be the* L_4*-model* (σ_1,σ_2). *We associate to it the* L_4*-interpretation function* (ζ_1,ζ_2) , *defined by the following rules:*

4-Sem*(i)* $\zeta_1 = \sigma_1$
4-Sem*(ii)* $\zeta_2=\sigma_2$
4-Sem*(iii)* *if* $\tau\in\bar{\bar{T}}$ *and* $\alpha\in\bar{\bar{F}}$ *then* $\zeta_1(\tau\alpha)=\zeta_2(\tau)(\zeta_1(\alpha))$.

Remarks

In modal logic systems, it is convenient to identify the meaning or "sense" of a sentence with the set of possible words in which the sentence is true (or with the characteristic function of such a set).

In the same way it is convenient to identify, within tense logic systems, the meaning of a sentence with the set of moments at which the sentence is true. Actually, in L_1 *through* L_3 *we did identify sentence meanings with characteristic functions of sets of those moments at which a given sentence is true.*

In all of our systems terms must be thought of as temporal clauses. They change the meaning of a sentence into a different sentence meaning. So, if we agree on taking sentence meanings as elements of the set W, *it is quite natural to take temporal clause meanings as elements of the set* W^+.

Now it is easily seen that all possible denotations of terms in L_3 *can be expressed in* L_4 *too. Indeed, let* σ_2 *be the second part of an* L_3*-model. Now, to each element* ϕ *of* U^U *we assign its "translation"* $\lambda\psi\in W.\psi\circ\phi$ *We indicate the result of this translation as* $\hat{\phi}$. *Clearly, for each* ϕ *in* U^U, $\hat{\phi}$ *will be an element of* W^+. *So the function* σ_2', *obtained by translation from* σ_2, *and defined by* $\sigma_2=\lambda\tau\in\bar{\bar{T}}.\hat{\sigma}_2(\tau)$, *gives rise to the* L_4*-model* (σ_1,σ_2').

Let (ζ_1,ζ_2) *and* (ζ_1',ζ_2') *be the* L_3*- resp.* L_4*-denotation function, associated to* (σ_1,σ_2) *and* (σ_1,σ_2') *respectively. It can be proven in a straightforward way that* ζ_1 *and* ζ_1' *are equal. So for each* L_3*-model, we may take a* L_4*-model that yields interpretations of formulae to the same semantic effect as the original* L_3*-model exercises on formulae.*

Apart from this use of terms, there are three rather important ways to apply the machinery involved in L_4.

These are what may be called universal reference to sets, existential reference to sets, and measurement on sets.

We encounter universal reference to sets in clauses like"during the past year". More or less informally, "during the last year" applied to a sentence S is true at the moment t if and only if S is true at all momentsindicated as "the last year." In this case the period itself referred to is dependent on the moment of utterance, like time point references as "last sunday at noon". Of course, it is possible to have constant references to periods too: "during the first millennium BC" might be such a constant reference.

Let ξ *be any function that assigns subsets of* U *to elements of* U.

We define $\forall_\xi$ *to be* $\lambda\phi\in W.\lambda t\in U \prod_{t'\in\xi(t)} \phi(t')$.

Here the iterated product sign is used to quantify univerally on truth valued functions.

Whenever ξ *is a mapping from* U *into* 2^U, $\forall_\xi$ *might be the denotation of some term, and behaves like a kind of universal quantifier.*

"Always", taken in its most literate sense can be considered as corresponding to $\forall_\Xi$, Ξ *being the constant function that assigns* U *to any moment of* U.

Existential reference to sets may occur in sentences like "yesterday I met Mary", which sentence is true at t, if at least at one of the moments of the day previous to t, the sentence "I meet Mary" is true.

Suppose again that ξ *assigns subsets of* U *to moments of* U.

We define $\exists_\xi$ *to be* $\lambda\phi\in W.\lambda t\in U. \sum_{t'\in\xi(t)} \phi(t')$. *(The iterated sum is used here for existential quantification on truth valued functions).*

"Sometimes" in the most unrestricted sense of the word corresponds to $\exists_\Xi$.

It must be stressed here that often in natural language references to sets of moments are made, while it remanains unclear if the kind of quantification is universal or existential, or, perhaps, different from both types. Particularly references to periods as "last year" and so on seem essentially unclear in this respect. In the systems still to be explained we will try to handle these cases.

Existential and universal quantification on a set coincide if and only if the set referred to is a singleton. In this case we have a kind of reference to moments like in L_3.

The third important application of L_4*-like clauses consists in the use of time measurement phrases like "half an hour". This applied to a sentence seems to be true at a moment t if and only if there exists an interval of length half an hour during which the sentence is true. But here we feel that something goes wrong: if we are going to explain the logical status of time meausurement clauses this way, then the truth value of sentences with this sort of clauses would become independent of time. It seems that this conclusion would be completely absurd. Evidently, the truth value of "yesterday I sang half an hour" is dependent on time. We may infer that we were too simple-minded when we assumed that "half an hour" just changes sentence meanings into different sentence meanings. Considerations of this kind will lead us to the complex semantics of* L_6.

Finally, it is easy to see that negation can be expressed in L_4 *too. Consider the mapping* $\lambda\psi\in W.\lambda t\in U. 1-\psi(t)$. *Obviously, this is an element of* W^+ *, and it acts like negation in proposition calculus.*

We may use it in combination with "always" and "sometimes" in order to get "never". It may seem strange that such an "untemporal" element as negation enters into the system. On the other hand, it does no harm.

The System L_5

1 Alphabet

The alphabet of L_5 *is the union of the alphabet of* L_4 *and the set* $\bar{\bar{V}}_0$, *which is the set of conjunction variables* $\{v_0, v_1, v_2, \ldots\}$.

2 Syntactic Categories

Syntactic categories are those of L_4 , *in addition, there is the syntactic category* $\bar{\bar{V}}$, *the set of conjunctions.*

3 Syntactic Rules

The first three syntactic rules of L_5 *are identical to those of* L_4. *In addition:*

5-Synt*(iv)* $\bar{\bar{V}}=\bar{\bar{V}}_0$;

5-Synt*(v)* *if* $v\in\bar{\bar{V}}$ *and* $\alpha\in\bar{\bar{F}}$, *then* $v\alpha\in\bar{\bar{T}}$.

4 Models

An L_5*-model is a triple* $(\sigma_1,\sigma_2,\sigma_3)$, *such that* (σ_1,σ_2) *is an* L_4*-model and* $\sigma_3:\bar{\bar{V}}\to(W^+)^W$.

5 Semantic Rules

Let $\mathfrak{M}$ *be an* L_5*-model* $(\sigma_1,\sigma_2,\sigma_3)$. *We associate to it the interpreation function* $(\zeta_1,\zeta_2,\zeta_3)$, *the rules of which coincide with the semantics of* L_4. *In addition, the following rules hold:*

5-Sem*(iv)* $\zeta_3=\sigma_3$

5-Sem*(v)* *if* $v\in\bar{\bar{V}}$ *and* $\alpha\in\bar{\bar{F}}$, *then* $\zeta_2(v\alpha)=\zeta_3(v)(\zeta_1(\alpha))$.

Remarks

In the first place, the way the semantic rules are presented presupposes that L_8 *is not syntactically ambiguous. In our earlier systems this fact was so obvious, that we did not find worth noticing it. Here, however, the situation is slightly different: the non-ambiguity of* L_8 *is not so evident. But fortunately* L_5 *is syntactically not-ambiguous. Here we omit the proof of this statement.*

About the logical structure of L_5, *the following fact may be interesting.*

Propositional 2-valued connectives like conjunction, disjunction, material implication can be expressed in the conjunctions of L_5 .

Let us take conjunction, for example. Now let $\mathfrak{M}$ *be a model that assigns to the conjunction variable* v *the value* $\lambda\phi\in W.\lambda\psi\in W.\lambda t\in U.\phi(t)\wedge\psi(t)$.
It is easily found that under the interpreation function $\mathfrak{M}$, $v\alpha_1\alpha_2$ *will be true at the moment* t *if and only if both* α_1 *and* α_2 *are true at this moment.*

Of course, we are not only, and even not in the first place, interested in logical connectives. We may construct interpretations of our conjunctions that have something to do with time too. As an example, we might assign to a value that can be understood as "half an hour before ----".

We must admit that the semantics of this kind of expression is not so clear as it might seem. In the sentence "half an hour before it started raining, he went away" there is some strange point. Evidently, it soes not say just "There is a moment t, such that at t it started raining, and half an hour before t he went away." More adequate solutions for the semantics of this kind of expressions will be given in the ultimate system L_8 .

Finally, L_5 *is both semantically and syntactically a proper extension of* . *This can be seen directly from the rules.*

The system L_6

Some preliminaries

The semantics of L_6 is quite complex. So we will introduce some new concepts before we will present the system as such .
First, we observe that quantifiers with respect to a given set S *can be thought of as truthvalued functions with the power set of* S *as their domain. The universal quantifier with respect to* S *is the function that assigns* 0 *to all parts of* S, *except* S *itself; the existential quantifier with respect to* S *assigns* 1 *to all parts of* S, *except the empty set, which gets value* 0 *(or falsity). In similar ways other known quantifiers (e.g. "exactly one", "exactly two", "infinitely many", "more than hundred") can be easily defined.*

Now we introduce the concept of complex quantifier.
This is a function that assigns to all pairs of subsets of a given set a truth value, under the condition that whenever Q *is a complex quantifier with respect to* S, *and* X *and* Y *are subsets of* S, *the equality*
$Q(X,Y)=Q(X,X\cap Y)$ *must hold.*
By the latter restriction it is easy to relate each complex quantifier with respect to S to a function that assigns to subsets of S *quantifiers with respect to that subset.*
The universal complex quantifier with respect to S *is a function*
$Q:2^S\times 2^S\rightarrow\{0,1\}$
such that for $X,Y\subset S$, $Q(X,Y)=1$ *if and only if* $Y\subset X$;
similarly, the existential complex quantifier with respect to (S) is a function such that for $X,Y\subset S$, $Q(X,Y)=0$ *if and only if* $X\ Y=\emptyset$.
Now it is possible to costruct functions with domain and range the set of complex quantifiers with respect to S. *Functions of this kind will be called indexed complex quantifiers with respect to* S .

A complex quantifier with respect to S *is said to select* s *from* X *whenever* $s\in S$, $X\subset S$ *and for all* Y S*:* $Q(X,Y)=1$ *if and only if* $s\in X\cap Y$.

We call an indexed complex quantifier nearly constant if it satisfies the following condition:
let Q *be the indexed complex quantifier with respect to* S *; now for* $s_1,s_2\in S$, $X,Y\subset S$:
if $Q(s_1)$ *does not select* s_1 *from* X, *and* $Q(s_2)$ *does not select* s_2 *from* X ,
then $Q(s_1)(X,Y)=Q(s_2)(X,Y)$.
The reason for the definition of "nearly constant" may seem mysterious. We motivate it as follows. In the present system and in L_8, "sentences" without explicit temporal clauses are supposed to refer to a given set of moments by some contextual or extralinguistic information. If the moment t *is included in the set referred to at* t, *the truth value of the sentence is supposed to be equal to the truth value of the same sentence, prefixed by a reference to the moment* t. *On the other hand, whenever* t *and* t *are both outside the set referred to (by extralinguistic information or by a temporal clause) the way of quantification over this set must be equal for both "outsiders".*
Now we define the standard indexed complex quantifier with respect to S .
This is the unique indexed complex quantifier that satisfies
for $s\in S$, $X\subset S$, $Y\subset S$:
a $Q(s)$ *selects* s *from* X , *if* $s\in X$;
b $Q(s)(X,Y)=0$ *if and only if* $X\cap Y=\emptyset$, *if* $s\notin X$

The standard complex quantifier with respect to the specific set $\mathbf{S}$*, will be indicated by the symbol* STAND_6.

The symbol $\mathbf{Q}$ *will be used for te set of allthose indexed complex quantifiers that are nearly constant.*

We will write Λ *for the set of pairs* $\mathbf{Q}\times(2^U)^U$ *and* Λ^+ *for* Λ^{Λ} .

Now we define a "transformation", i.e. a function that assigns elements of W^+ *to elements of* Λ.

This transformation is given the name Tr*, and is defined by:*

for $Q\in\mathbf{Q}$*, and* $\xi:U\to 2^U$,

$$\mathrm{Tr}(Q,\xi)=\lambda\psi\in W.\lambda t\in U.Q(t)(\xi(t),\{t'\,|\,\xi(t')=1\}).$$

After these introductory definitions and conventions, we exhibit the system.

1 Alphabet

The alphabet of L_6 *is like that of* L_5 *, except for the presence of two additional bracket signs: "*[*" and "*]*".*

2 Syntactic categories

L_6 *has the same syntactic categories as* L_5*, and in addition the category* $\bar{\bar{B}}$ *(clauses)*

3 Syntactic Rules

For L_6*, the syntactic rules (i), (iv) and (v) are the same as in* L_5. *The remaining rules are:*

6-Synt*(ii)* $\bar{\bar{T}}_0\subset\bar{\bar{T}}$;

6-Synt*(iii)* *if* $\tau_0\in\bar{\bar{T}}_0$ *and* $\tau_1\in\bar{\bar{T}}$*, then* $\tau_0\tau_1\in\bar{\bar{T}}$;

6-Synt*(vi)* *if* $\tau\in\bar{\bar{T}}$*, then* $[\tau]\in\bar{\bar{B}}$;

6-Synt*(vii)* *if* $\beta\in\bar{\bar{B}}$ *and* $\alpha\in\bar{\bar{F}}$*, then* $\beta\alpha\in\bar{\bar{F}}$.

4 Models

An L_6*-model is a quadruple* $(\sigma_1,\sigma_2,\sigma_3,\xi)$ *, such that*

$\sigma_1:\bar{P}\to W$, $\sigma_2:\bar{\bar{T}}\to\Lambda^+$, $\sigma_3:\bar{V}\to(\Lambda^+)^W$ and $\xi:U\to 2^U$.

5 Semantic Rules

Let $\mathfrak{M}$ *be the* L_6*-model* $(\sigma_1,\sigma_2,\sigma_3,\xi)$. *We associate to it the interpretation function* $\tilde{\mathfrak{M}}$*, which is a quadruple* $(\zeta_1,\zeta_2,\zeta_3,\zeta_4)$ *, such that*

$\zeta_1:\bar{\bar{F}}\to W$, $\zeta_2:\bar{\bar{T}}\to\Lambda^+$, $\zeta_3:\bar{V}\to(\Lambda^+)^W$,$\zeta_4:\bar{\bar{B}}\to W^+$

and satisfying the following semantic rules:

6-Sem*(i)* $\sigma_1\subset\zeta_1$;

6-Sem*(ii)* $\sigma_2\subset\zeta_2$;

6-Sem*(iii)* *if* $\tau_0\in\bar{\bar{T}}$ *and* $\tau_1\in\bar{\bar{T}}$*, then* $\zeta_2(\tau_0\tau_1)=\sigma_2(\tau_0)\circ\zeta_2(\tau_1)$;

6-Sem*(iv)* $\sigma_3\subset\zeta_3$;

6-Sem*(v)* *if* $v\in\bar{V}$ *and* $\alpha\in\bar{\bar{F}}$ *, then* $\zeta_2(v\alpha)=\zeta_3(v)(\zeta_1(\alpha))$;

6-Sem*(vi)* *if* $\tau\in\bar{\bar{T}}$ *, then* $\zeta_4([\tau])=\mathrm{Tr}(\zeta_2(\tau)(\mathrm{STAND}_6,\xi))$;

6-Sem*(vii)* *if* $\beta\in\bar{\bar{B}}$ *and* $\alpha\in\bar{\bar{F}}$ *, then* $\zeta_1(\beta\alpha)=\zeta_4(\beta)(\zeta_1(\alpha))$.

Remarks

First, we consider the syntax of L_6. *All formulae consist of a (possibly empty) sequence of "clauses", followed by one proposition variable. Clauses are terms in brackets, and terms are non-empty sequences of "atomic terms", i.e. either term variables or formulae prefixed by a conjunction variable.*

On the semantic level, the structure of L_6 *is quite more complex. We notice that the clauses in* L_6 *have the semantic function of terms in* L_4 *and* L_5, *being of the semantic type* W^+.

The terms form the ingredients for clauses, in that their meaning by enclosing in brackets, is transformed into a meaning of type W^+.

One may wonder how term meanings must be interpretated intuitively. The following remarks are intended to clarify this point.

We have met in L_4 *some difficulties with combining clauses like "yesterday" and "during half an hour".*

If we take "during half an hour" semantically as an L_4*-term, i.e. as a modifier of sentences, then inevitably sentences like "yesterday she was singing during half an hour" are bound to get absurd interpretations.*

Indeed, let us ignore, for the sake of simplicity, the complications involved with the use of the past tense. Then after having been translated into L_4, *with the use of an appropriate model, we may obtain these two interpretations (it depends on the order of the terms which interpretation we actually obtain):*

a "Yesterday it was the case that there is a period of half an hour during which she will be/ is / has been singing";

b "Once in the future, the present or the past the sentence 'she was singing yesterday' will be/ is/ has been true during half an hour."

In L_4 *the problem consists in the impossibility to combine "yesterday" and "half an hour" in such a way, that the quantifier "during half an hour" is applied to the period indicated with "yesterday" only. And this, we feel, is the way "half an hour" shuold be interpreted.*

At this point, we might think of an alternative system with two kinds of temporal clauses:

a References to sets (which may depend on the moment of utterance);

b Indicators, how the set referred to, must be referred to(e.g. existentially, or universally, or in some still different way.)

Evidently, clauses like yesterday would belong to class a, and clauses like "during half an hour" to class b.

Actually, for several reasons we do not construct a system like the one sketched above, but on the other hand, the distinction between reference (a) and complex quantifier (b) is preserved in L_6: *terms have both properties at once, so to say. The semantic effect of a term (at a given moment* t*) is that it changes a given set of moments referred to into a new one, and a given complex quantifier into a new one. "Yesterday", for instance, taken as an* L_6*-term, restricts a given set referred to, at the moment* t *to the intersection of the set referred to with the set of moments of the day before* t. *It leaves complex quantifiers unchanged, while "during half an hour" leaves the set referred to unchanged, but changes a complex quantifier into the one that could be defined as*

$$\lambda T_1 \subset U . \lambda T_2 \subset U . \theta(T_1, T_2),$$

such that θ *is truth valued, and takes on value* 1 *for* T_1 *and* T_2, *if and only if there is an interval* $T_3 \subset T_1 \cap T_2$, *that lasts half an hour.*

(We are not concerned here with the question whether this definition is accurate).

It is important to notice that at a moment t terms are not interpreted as pairs of complex quantifiers and sets of moments; they rather change such a pair into another one. So there must be sets referred to, to start with, and there must be an (indexed) complex quantifier to start with.

Now the set referred to "from the start" cannot be expressed by linguistic means, but must somehow be clear from the context. The knowledge about "unstated references" at each moment t *is given in the model by the fourth coordinate.*

At each moment t *, if we want to evaluate the truth value of a formula* $\beta\alpha$ *at* t*, we proceed as follows. First evaluate that term* τ*, that satisfies the equation* $\beta = \tau$ *. The result of this operation is applied to the pair* $(STAND_6, \xi)$ *. This yields a pair, consisting of an indexed complex quantifier and a function that assigns sets of moments to moments. We evaluate both parts at* t. *The result is a complex quantifier* Q, *and a set of moments* T_1. *Now we evaluate (not necessarily at* t*, but at all moments of* T_1.*) Let* T_2 *be the set of those moments in* T_1*, at which* α *is true. Finally, we calculate* $Q(T_1,T_2)$. *This ultimate result is the truth value of* $\beta\alpha$ *at the moment* t.

One of the weaker points in the present system is the definition of the standard indexed complex quantifier. The underlying assumption here is that, if no indication of the contrary is given one must quantify existentially over sets referred to. So "I saw you" is regarded as true at the moment t*, if and only if the part of the past of* t *referred to contains a moment at which "I see you" is true.*

One can take a different point of view. One could argue that, if no indication is given how to quantify over a set referred to, one specific "prefered" moment of that set is particularly referred to. Formalization of this line of thought leads us into accepting so called choice functions (i.e. mappings that "choose" from sets "preferred" moments) as a semantic tool.

Before enriching L_6 *with these choice functions, we are going to construct a system with choice functions that is in some respects poorer than* L_6.

The system L_7 (1 Alphabet)

The alphabet of L_7 *consists of* $\overline{\overline{P}}$ *, the so called term constants* τ_{ov}, τ_{bv}, τ_{ot}, τ_{bt} *and the conjunction constants* v_{ov}, v_{bv}, v_{ot}, v_{bt}, v_o, v_b.

2 Syntactic Categories

Syntactic categories of L_7 *are* $\overline{\overline{F}}$, $\overline{\overline{T}}$ *and* $\overline{\overline{V}}$.

3 Syntactic Rules

Rules (i) and (v) are like in L_6. *The remaining rules are*

7-Synt*(ii)* $\{\tau_{ov},\tau_{bv},\tau_{ot},\tau_{bt}\}\subseteq\overline{\overline{T}}$;

7-Synt*(iii)* $\overline{\overline{V}}=\{v_{ov},v_{bv},v_{ot},v_{bt},v_o,v_b\}$;

7-Synt*(iv)* *if* $\alpha\in\overline{\overline{F}}$ *and* $\tau\in\overline{\overline{T}}$ *, then* $\tau\alpha\in\overline{\overline{F}}$.

4 Models

An L_7*-model is a pair* (σ_1, K)*, such that* $\sigma_1:\overline{\overline{P}}\to W$ *and*

a $K: 2^U \setminus \{\emptyset\} \to U$,

b $K(T)\in T$ *whenever* $T\subset U$ *and* $T\neq\emptyset$.

5 Semantic Rules

Let $\mathfrak{M}$ *be the* L_7*-model* (σ_1,K). *We associate to it the interpretation function* $\hat{\mathfrak{M}}$, *which is a triple* $(\zeta_1,\zeta_2,\zeta_3)$, *such that* $\zeta_1:\bar{\bar{F}}\to W$, $\zeta_2:\bar{\bar{T}}\to W^+$, *and* $\zeta_3:\bar{\bar{V}}\to (W^+)^W$.

We are going to give some temporary definitions, before fixing the values of the coordinates of $\hat{\mathfrak{M}}$.

First, for any $\psi\in W$, *and* $t\in U$, *we write:*

$\psi^{<t}$ *for the characteristic function of* $\{t'\in U \mid t'<t \text{ and } \psi(t')=1\}$;

$\psi^{>t}$ *for the characteristic function of* $\{t'\in U \mid t'>t \text{ and } \psi(t')=1\}$

In the second place, for $\psi_1,\psi_2\in W$ *and* $t\in U$ *we write:*

$Ex(\psi_1,\psi_2)$ *for* $\sum_{t\in U}\psi_1(t)\wedge\psi_2(t)$;

$Keuze(\psi_1,\psi_2)$ *for* $\psi_2(K(\{t\in U \mid \psi_1(t)=1\}))$.

Finally, for $T\subset U$ *we write* $T^{\dagger}$ *for the characteristic function of* T.

We are able now, to present the semantic rules in a comprehensible fashion. Rules (i) and (v) are identical to the corresponding rules in L_5 *and* L_6. *The remaining rules are:*

7-Sem*(ii)* a $\zeta_2(\tau_{ov}) = \lambda\psi\in W.\ \lambda t\in U.Ex(\psi,(T^{<t}))$ *(Here* T *stands for* $U^{\dagger}$*)*

b $\zeta_2(\tau_{bv}) = \lambda\psi\in W.\ \lambda t\in U.Keuze(\psi,(T^{<t}))$

c $\zeta_2(\tau_{ot}) = \lambda\psi\in W.\ \lambda t\in U.Ex(\psi,(T^{>t}))$

d $\zeta_2(\tau_{bt}) = \lambda\psi\in W.\ \lambda t\in U.Keuze(\psi,(T^{>t}))$

7-Sem*(iii)* a $\zeta_3(v_{ov}) = \lambda\psi_1\in W.\lambda\psi_2\in W.\lambda t\in U.Ex(\psi_1,\psi_2^{<t})$

b $\zeta_3(v_{bv}) = \lambda\psi_1\in W.\lambda\psi_2\in W.\lambda t\in U.Keuze(\psi_1,\psi_2^{<t})$

c $\zeta_3(v_{ot}) = \lambda\psi_1\in W.\lambda\psi_2\in W.\lambda t\in U.Ex(\psi_1,\psi_2^{>t})$

d $\zeta_3(v_{bt}) = \lambda\psi_1\in W.\lambda\psi_2\in W.\lambda t\in U.Keuze(\psi_1,\psi_2^{>t})$

e $\zeta_3(v_o) = \lambda\psi_1\in W.\lambda\psi_2\in W.\lambda t\in U.Ex(\psi_1,\psi_2)$

f $\zeta_3(v_b) = \lambda\psi_1\in W.\lambda\psi_2\in W.\lambda t\in U.Keuze(\psi_1,\psi_2)$

7-Sem*(iv)* *if* $\alpha\in\bar{\bar{F}}$ *and* $\tau\in\bar{\bar{T}}$, *then* $\zeta_1(\tau\alpha)=\zeta_2(\tau)(\zeta_1(\alpha))$.

Remarks

The system L_7 *has been built up for just one purpose: to show the difference between "definite" vs. "indefinite" references. The indices "o" and "b" stand for indefnite (onbepaald) and definite (bepaald) respectively. The indices "v" and "t" stand for past (verleden) and future (toekomst).*

First we shall consider the distinction between indefinite and definite past. As far as I know, it is easier to find clear examples of this distinction in Dutch and in French, than in English. So I will take some examples from the former two languages.

Sentences like "Ze vertrokken" in Dutch or "Ils partirent" in (litterary) French won't be easily understood in the sense of "once they left." They rather presuppose some moment in the past that is referred to. Whenever in Dutch one would not refer to such a moment in the past whose identity is tacitly assumed to be known, one would say in the perfect past "Ze zijn vertrokken"; in French one could say "Ils sont partis" for this existentially referring to the past.

I believe that to a certain extent, in a large number of languages, the distinction between existentially or indefintely referring and referring definitely or by choice function (which is fixed at beforehand), can be formally expressed. On the other hand, in no language known to me, the formal distinctions conincide completely; for example, the distinction between simple and perfect past in Dutch serves a lot of purposes, among which the distinction definite vs. indefinite is, to my opinion, of great importance.

The indefinite past could very well be circumscribed with an expression like "at a moment in the past", while the definite past might be thought of as equivalent to "at the moment in the past". In the latter case, there must be agreement between partners in the communication process about the time point referred to.

Such an agreement must exist too, in those cases where the past is restricted to a set of moments in the past.

With this I mean sentences of the kind "when he came in, silence fell." This sentence, I suppose, says, that at the moment of reference silence fell, while the moment of reference is defined by "when he entered". But how do we know which moment of his entering is referred to? Somehow, during communication acts, we do know. This knowledge (often extralinguistic) is necessary for correctly understanding a lot of vaguenesses in natural languages. This knowledge is tacitly assumed whenever we want to refer to objects by mentioning the sets to which they belong.(The use of the definite article in e.g."the book", even while there is not necessarily one unique book, is another example of this use of knowledge.) Such knowledge can be described by a so called choice function. In reality the choice function offers far more than we can ever know: in this respect the constructions in are highly unrealistic. As to that point, it would have been better to use partial choice functions. For a lot of purposes they would equally well do too; but using them would have required some technical changes whose consequences are still unknown to me.

contains two rather strange looking conjunctions, and , whose senses are more or less equivalent to "at a moment that ..." and "at the moment that". Although they look at first sight natural, a second look convinces us of their unusualness. Formulae, constructed with these conjunctions are always "timeless": i.e. their truth values do not depend on time.

It does not seem very likely that existential or choicewise reference is possible, without any indication of either past or future. So, the last two conjunctions, even if formally alright, are no candidates for translations of conjunctions in natural languages.

Most things I said about definite and indefinite past, hold for definite and indefinite future too, I believe. But in practice, so it seems to me, it is difficult to employ definite future, unless for highly predictable parts of the future, or for communication between psychics. So I completed with definite future, more for the sake of symmetry than for practical matters.

The system L_8

Preliminaries

In L_6 *we defined the standard indexed complex quantifier* $STAND_6$. *This can be regarded as a construction that at a given moment* t *and for a given set referred to* T *behaves as a quantifier that selects* t *from* T *if possible (i.e., if* $t \in T$*), and refers existentially to* T *in other cases.*
In L_8 *similar standard indexed complex quantifiers are present, but here the existential or indefinite referring has been changed by choicewise or definite referring.*
Formally, whenever K *is a choice function (i.e.* $K: 2^U \setminus \{\emptyset\} \rightarrow U$ *and for each* $T \in 2^U \setminus \{\emptyset\}$: $K(T) \in T$*),*
we write $STAND_K$ *for the function* Q, *defined by*

(i) Q *is a indexed complex quantifier;*
(ii) for $t \in U$ *and* $T_1, T_2 \subset U$:

$$Q(t)(T_1, T_2) = 1 \ , \text{ if } \ t \in T_1 \cap T_2 \ \text{ or } t \notin T \ \text{ and } K(T_1) \in T_2;$$
$$= 0 \ , \text{ in other cases.}$$

Now we give the formal definition of L_8.

1, 2, 3 Alphabet, Syntactic Categories, Syntactic Rules

Like in L_6 .

4 Models

An L_8*-model is a quintuple* $(\sigma_1, \sigma_2, \sigma_3, \xi, K)$ *that satisfies the following conditions:*
$\underline{a}$ $(\sigma_1, \sigma_2, \sigma_3, \xi)$ *is an* L_6*-model;*
$\underline{b}$ (σ_1, K) *is an* L_7*-model.*

5 Semantic Rules

Let $\mathfrak{M}$ *be the* L_8*-model* $(\sigma_1, \sigma_2, \sigma_3, \xi, K)$. *We associate to it the* (L_8)*-interpretation function* $\tilde{\mathfrak{M}}$, *which is a quadruple* $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)$.
The semantic rules of L_8 *are equal to the corresponding rules of* L_6, *with the exception of:*

8-Sem *(vi)* *if* $\tau \in \bar{\bar{T}}$, *then* $\zeta_4([\tau]) = Tr(\zeta_2(\tau)(STAND_K, \xi))$.

Of course, the difference with 6-Sem*(vi) is minimal, from a superficial point of view: the only change consists in substitution of* $STAND_6$ *by* $STAND_K$.

Remarks

Anything that can be expressed in one of the systems L_0 through L_7 can be expressed in L_8 too. This may seem obvious, but the proof of this statement requires the proof of some "translation lemma's" which I am not going to carry out here.

So I shall restrict myself to giving a few examples of temporal constructions in natural languages that cannot be expressed in one of the former languages, while they are "translatable" into L_8.

The treatment of these cases will be rather informal; the calculations that show how the "translations" work exactly are omitted.

First, in the remarks after the presentation of L_6, we stated that clauses like "yesterday", referring to (time dependent) sets, must be considered as functions that leave complex quantifiers unchanged, while they restrict the the set referred to, to its intersection with the set of moments of the day before. We keep this position here, but as the semantics of L_8 is different, "yesterday" behaves differently too. Indeed, if we prefix a sentence with a bracketed "yesterday" in L_6, the semantic effect will be that the prefixed sentence is true if and only if its unprefixed part is true at least at one moment of the day before, insofar such moment is included in the set implicitly referred to.

Now, in L_8 no existential, but choisewise referring will determine the possible truth values of a sentence with a bracketed "yesterday" at head. So the sentence will be true if and only if its unprefixed part is true at that particular moment that is chosen to be the representant of the intersection of the current set referred to with the set of moments of the day before.

My second point concerns the behavior of time measurement clauses like "during half an hour." Both in L_5 and in L_6 there are difficulties to translate it properly. In L_6 we took the position that it is to be regarded as a function of indexed complex quantifiers that changes quantifiers into the "half an hour"-quantifier of L_5 and leaves the set referred to unchanged.

In L_8 we prefer a different, and, as I hope, more appropriate definition. The sense of "during half an hour" must be taken to affect the current complex quantifier in the following way:

for indexed complex quantifier Q, and $\xi \in \Lambda$, the first coordinate of "half an hour"(Q,ξ) is Q', defined by

$$Q'=\lambda t \in U.\lambda(T_1,T_2) \subset U \times U.\Gamma(T_1,T_2,t)$$

Here $\Gamma(T_1,T_2,t)=1$ if and only if there is a moment t' such that $Q(t)$ selects t' from T_1, while $T_1 \cap T_2$ contains an interval T_3 of length half an hour, such that $t' \in T_3$.

So, "half an hour", at a given moment, will refer to a period of half an hour in the neighborhood of a moment referred to in the operand sentence.

If a term does not refer to a specific moment (at a given time point), then prefixing the term with "half an hour" will yield a term that, taken as a clause, makes formulae trivially false. So no sentence, prefixed with [half an hour always] can be true.

Sentences like "during half an hour she coughed sometimes" seem to contradict this statement. This need not be so, however. If we translate the sentence into L_8-like terms in the following way:

"[half an hour definite past] [sometimes] she coughs"

and "sometimes" is assigned a meaning like "repeatedly", "from time to time" "frequently, more or less, in the environment of the specific moment referred to" no doubt the sentence above will get its proper interpretation.

In this translation "half an hour" is prefixed to "definite past", and these together, as a clause, are prefixed to the sentence

"[sometimes] she coughs."

One feature that could have been explained in the remarks on the system L_7 too, is the possibility of constructions of the "past perfect past" tense in sentences like "I had come".

This tense refers existentally to the past of the choisewisely determined moment of the "actual past". Thus, in order to know if "I had come is true" one need to establish whether at the moment of the past the sentence "I have come" is true, whereas this later sentence is regarded to be true if and only if "I come"is true at some moment in the past (or rather that part of the part which is under discussion).

Finally, I would like to touch the problem of the so called consecutive conjunction. The sentence "He walked out of his house and crossed the road" is clearly no logical conjunction: in societies like ours it would always be false, as the two events, described in the successive parts of the sentence, cannot occur simultaneously. No one has ever found any difficulty in finding an explanation for this fact: the two events described occur one after another. In L_8 it is rather easy to define a semantics of this consecutive conjunction. This conjunction, applied to formulae in L_8, yields modifiers of sets referred to, to the effect that the set referred to originally is restricted to those moments in the set that follow the moment of the evbnt described in the subordinated formula.

So in the sentence mentioned above, first a given moment of the past is given us by the choice function, that applies to the past part of the set referred to. Then we look, if at that moment "He walked out of his house" is true. If not so, the sentence is false. If the first part is true, we restrict the set referred to to the moments after our first choice. We make a new choice, and this will give us a moment after the first one. We look if at this second moment "he crossed the road" is true. If not so, the sentence is false, in spite of the truth of the first part. If so, the sentence is true.

The semantics of the consecutive conjunction must be organized so that the subordinate formula may contain no universal quantifiers(and so on) of time, and if it does contain them, the whole sentence becomes trivially false (in fact meaningless).

Final Remarks

In a number of ways, even the latest system, L_8, is incomplete.

The first kind of incompleteness could be called "ontological": one could make the objection that only one possible structure of time is taken into consideration, and that moments instead of short open intervals are the primitive points of reference.

The second kind of incompleteness regards the fact that, in the preceding pages, all structure in sentences, apart from the temporal constructions, has been ignored. This omission is very serious indeed: even the use of proper names without any temporal designation may affect the choice of the time referred to. For example: we all know immediately and intuitively the difference in reference to time when we compare two simple sentences:"Julius Caesar wrote a letter" and "Aldo Moro wrote a letter".

Thirdly, one would like to see temporal constructions and words systematically translated into the systems we treated.

And finally: the reader may wonder why I never mentioned any author on the subject of time and tenses.

As to the first point: the elaboration of the systems only to a minor degree depends on the actual structure of time. So I do not think it will require a great effort to adapt them to different concepts of time. The second and the third form of incompleteness are to be coped with: I hope the futurewill enable me to pay themthe attention they deserve.

The forth omission can hardly be apologized for.

I am indebted to all time logicians and linguists for the suggestions I found in their texts. By no means I want to give the impression that to my opinion my problems, and their solutions are original.

I believe there is some originality in the latest three systems, but I do not know if this originality will not turn out to be a waste of time.

A FORMAL SEMANTIC ANALYSIS OF MASS TERMS AND AMOUNT TERMS

Harry Bunt
Institute for Perception Research
Eindhoven, Holland

1. Introduction

In recent years there has been quite some interest in the semantic problems of mass terms[1], both from a philosophical and from a linguistic point of view[2], but to my knowledge no adequate proposal for the formal semantic analysis of mass terms has been worked out so far. The most elaborate and explicit proposal comes from Parsons (1970), but unfortunately his framework can be shown to be logically inconsistent (Bunt, 1976; 1977). In the present paper I propose an analysis of mass terms that has a certain similarity with Parsons' analysis, but that does not suffer from the same logical deficiencies. In order to deal with quantified mass terms I also include a formal analysis of amount terms, which has a certain interest in its own right.

I want to focus on those aspects of mass terms in which they differ semantically from count terms; therefore I restrict myself e.g. to non-intensional contexts. Mass terms are, semantically as well as syntactically, similar in some respects to plural count terms and also in some respects to singular count terms. A mass term is, informally speaking, sometimes similar to a singular count term in the sense that both seem to refer to one single entity. For instance, in the sentence

(1a) The ice cream in the blue dish comes from Italy

the mass term "The ice cream in the blue dish" can very well be thought of as referring to an unstructured entity, just like the corresponding singular count term in:

(1b) The statue in the right corner comes from Italy

On the other hand, a difference between the mass term noun "ice cream" and the count noun "statue" is brought about by changing the predication "comes from Italy" into "weighs 5 kilogrammes" and recognizing that the resulting sentence (2a) has the implication (3a), whereas the resulting sentence (2b) does not have a similar implication (3b).

(2a) The ice cream in the blue dish weighs 5 kilogrammes
(3a) There is 5 kilogrammes of ice cream in the blue dish
(2b) The statue in the right corner weighs 5 kilogrammes
(3b) *There is 5 kilogrammes of statue in the right corner

If we compare sentence (2a) with a sentence where instead of the mass noun we have a plural count noun, such as:

(2c) The oranges in the blue dish weigh 5 kilogrammes

we notice on the syntactic level that the verb must have the plural form and on the semantic level we notice for instance that (2c) is ambiguous with respect to the type of quantification between a "total" reading:

(4c) The total weight of the oranges in the blue dish is 5 kilogrammes

and a "distributive" reading:

(4c') Each of the oranges in the blue dish weighs 5 kilogrammes

(which is of course not a very plausible reading, but one that cannot be excluded in view of the possibility to inject oranges with mercury).

Such an ambiguity does not exist in the mass noun case; there is no reading (4a') corresponding to (4c'):

(4a') *Each of the ice cream in the blue dish weighs 5 kilogrammes

Why is (4a') wrong? Why, in general, is it wrong to say "each of the m" with m a mass term? I think the reason must be that it would be quite unclear what such a phrase would refer to.

Comparing the noun phrases in (2a) and (2c), "The ice cream in the blue dish" and "The oranges in the blue dish", we have the modifying phrase "in the blue dish" functioning in both cases in the same way, viz. as a restriction of the extension of the head noun to the extension of the noun phrase. In the count noun case this notion of restriction is naturally treated in a formal analysis as the restriction of the set of all oranges to the subset of those oranges that are in the blue dish; in the mass noun case there is not such an obvious counterpart in terms of sets and subsets. The difference is that the singular form of the count noun indicates the elements of which the set consists that corresponds with the plural form, whereas in the mass noun case there is nothing in the language indicating what would be the elements of the set corresponding with the noun. This difference is often expressed by saying that a count noun 'individuates' or 'divides' its reference but a mass noun does not (see e.g. Quine, 1960, p.91).

A mass noun is, informally speaking, often similar to a plural count noun in the sense that its reference must be assumed to be a structured entity. Consider, e.g. the following sentences:

(5a) All the ice cream in the blue dish comes from Italy

(5b) All the oranges in the blue dish come from Italy

(5c) *All the statue in the corner comes from Italy

In order to understand the quantifier "all" in (4a) we must make some assumption as that the ice cream in the blue dish consists of certain parts, all of which are said to come from Italy. This might suggest the view that the mass term "the ice cream in the blue dish" refers to the set of all objects that count as ice cream parts in the blue dish. However, this view leads to a fundamental difficulty.

The difficulty is this. Suppose that a mass noun m is analysed formally as corresponding to a set S. Then the combination of m with a restrictive modifier corresponds to a subset of S. A set consists of elements, and thus has minimal, atomic subsets: subsets of one element. These subsets would correspond to minimal parts of the extension of m. Therefore, this commits us to the view that there are minimal parts of what a mass noun refers to. Indeed, Quine (1960) and others have expressed the view that an analysis of mass nouns should take into account that for every mass noun there are minimal parts of a size, characteristic for that noun. I think this view is mistaken. For one thing, though it may be true that e.g. gold, water and ice cream ultimately consist of molecules, the minimal parts of e.g. time, music, motion, phlogiston, love and nonsense still have to be discovered. But more importantly, such things as the physical knowledge of the molecular structure of matter is irrelevant to a formal linguistic study. Formal semantics is a study of aspects of language as a system in its own right. From a linguistic point of view, the particular feature of a mass noun is precisely that it provides a way of speaking as if its referent is not built up from atomic elements ("it seems that coffee <u>ought</u> to be infinitely divisible, though in fact it is not", Cartwright, 1970, p. 41)[3)].

In a way, one would like to use sets in the formal analysis of mass terms while forgetting about their elements and thereby about their atomic subsets, but this is of course not possible. Therefore, I have developed a set-like concept that I call an <u>ensemble</u>, having the same part-whole properties as sets and subsets but not necessarily having atomic parts. I will now proceed to outline the formal mathematical theory of ensembles; a more extensive introduction can be found in Bunt (1978).

2. Ensembles

2.1. The general theory

Ensembles are characterized by their parts: two ensembles are identical if they have the same parts. The part-of relation $\subseteq$ is a primitive notion in the theory of ensembles and its transitive nature must therefore be established by an axiom:

(6) $(\forall x, y, z)\ ((x \subseteq y \ \&\ y \subseteq z) \rightarrow (x \subseteq z))$

On the basis of this part-of relation it is possible to define operations corresponding to the union and intersection of sets. First, I define equality between ensembles as mutual inclusion:

(7) $x = y$ iff $x \subseteq y \ \&\ y \subseteq x$

The overlap (intersection) of two[4] ensembles is then defined as their largest common part, largest in the sense of the sense of the part-of relation:

(8) $z = x \cap y$ iff $z \subseteq x \ \&\ z \subseteq y \ \&\ (\forall z')\ ((z' \subseteq x \ \&\ z' \subseteq y) \rightarrow z' \subseteq z)$

The overlap of two ensembles can be empty. The empty ensemble $\emptyset$ is defined by the property that it has no other parts than itself:

(9) $(\forall x)\ (x \subseteq \emptyset \rightarrow x = \emptyset)$

The axioms of ensembles theory (listed in the appendix) guarantee that (9) defines the empty ensemble uniquely and that the empty ensemble is part of every ensemble.

The merge (union) of two[5] ensembles is defined as the smallest ensemble including both these ensembles:

(10) $z = x \cup y$ iff $x \subseteq z \ \&\ (\forall z')\ ((x \subseteq z' \ \&\ y \subseteq z') \rightarrow z \subseteq z')$

A fundamental property of the merge is that the merge of a number of parts of an ensemble is again a part of that ensemble.

A theorem of central importance in ensemble theory is the following.

Theorem 1. Given an ensemble x and a property P there exists an ensemble y which is the smallest ensemble containing all parts of x with the property P.
In formula:

(11) $$(\forall x)\ ((\exists y)\ ((\forall z)\ ((z \subseteq x\ \&\ P(z)) \rightarrow z \subseteq y)\ \&\ (\forall u)\ ((\forall z)\ (z \subseteq x\ \&\ P(z)) \rightarrow z \subseteq u) \rightarrow y \subseteq u))$$

I designate the ensemble y, uniquely determined by (11), by

(12) $$[z \subseteq x \mid P(z)].$$

From the definition of merge it follows immediately that Theorem 1 entails the following corollary.

Corollary 1. For any ensemble x and property P, the ensemble $[z \subseteq x \mid P(z)]$ is equal to the merge of those parts of x that have the property P.
In formula:

(13) $$[z \subseteq x \mid P(z)] = \cup\{z \mid z \subseteq x\ \&\ P(z)\}$$

Theorem 1 and corollary 1 form the basis for the application of ensembles to the formal semantic analysis of mass noun constructions, as we will see in section 3.

2.2. Continuous and discrete ensembles

Two particularly interesting kinds of ensembles are those that I call continuous and discrete, respectively.

It is useful to introduce first the relations of proper part ($\subset$) and genuine part ($\bar{\subset}$). The proper part-relation is defined in terms of the part-of relation and the equality relation (and thus, ultimately, in terms of the part-of relation only) in a way familiar from other parts of mathematics (See e.g. Kleene 1952, p.10).

(14) $x \subset y$ iff $x \subseteq y\ \&\ x \neq y$

Genuine part is then defined as non-empty proper part:

(15) $x \bar{\subset} y$ iff $x \subset y\ \&\ x \neq \emptyset$

An ensemble is called <u>continuous</u> if each of its non-empty parts has a genuine part. In formula:

(16) x is continuous iff $(\forall z)((z\subseteq x \ \& \ z \neq \emptyset)\rightarrow(\exists y)(y\subset z))$

In a continuous ensemble one can continue ad infinitum to take ever smaller parts. I believe that this kind of ensembles is the appropriate concept for formalizing the logical properties of mass terms of which the referents do not have minimal parts.

To introduce the notion of a discrete ensemble, I must first introduce <u>atomic ensembles</u>. An ensemble is called atomic if it is not empty and has no genuine parts. In formula:

(17) x is atomic iff $x \neq \emptyset \ \& \ \neg(\exists y)\ (y\subset x)$

To indicate that an ensemble x is atomic, I will write ATOM(x) as an abbreviation for the right-hand side of (17). An equivalent, alternative definition of atomic ensembles would be that an atomic ensemble has no other parts than the empty ensemble and itself. Thus:

(18) $\text{ATOM}(x)\rightarrow(\forall y)\ (y\subseteq x\rightarrow(y = x \vee y = \emptyset))$

Note that, if x, y and $\emptyset$ in (18) would denote sets rather than ensembles and $\subseteq$ would denote the subset relation, the sets satisfying (18) would be those sets that have exactly one element (atomic sets).

An ensemble is called discrete if it is equal to the merge of a number of atomic ensembles:

(19) x is discrete iff $x = a_1 \cup a_2 \cup a_3$ ... with $\text{ATOM}(a_1)$, i=1,2,3..

Note that, if x, a_1, a_2, a_3,... in (19) would denote sets rather than ensembles and $\cup$ would denote set-union, then this formula would be satisfied by every set x: a set is always equal to the union of its atomic subsets.

Discrete ensembles play a very special role in the formal theory. In order to see that clearly I introduce the second primitive relation of ensemble theory, the <u>unique-element</u> or <u>'unicle'</u> relation $\in$. The significance of this relation is expressed by the axiom saying that every atomic ensemble has exactly one unique element and every other ensemble has no unique element. In formula:

(20) $(\forall x)((ATOM(x)) \rightarrow (\exists! y)\ (y \in x)\ \&\ (\neg ATOM(x) \rightarrow \neg(\exists y)(y \in x)))$

In terms of the two primitive relations $\subset$ and $\in$ we can define a new relation, designated by ϵ, as follows:

(21) $y \epsilon x$ iff $(\exists z)(z \subseteq x\ \&\ y \in z)$

In other words, $y \epsilon x$ iff y is the unique element of an atomic part of x. Note that, since a continuous ensemble has no atomic parts, it has no elements.

The ϵ-relation defined by (21) has exactly the same properties as the element-relation in set theory. If we add to the axiom system by which the ensemble concept is formally defined (see the appendix) one more axiom saying that all ensembles are discrete, i.e. we restrict the ensemble concept to the discrete case, this axiom system can be proved to be equivalent to the Zermelo-Fraenkel axiom system for set theory (see e.g. Fraenkel et al., 1973, ch.II). In other words, discrete ensembles are sets. It follows that ensemble theory is a logically consistent generalisation of set theory. I will therefore call the ϵ-relation, defined by (21), the element-relation and I will also speak of the union of ensembles instead of the merge and of their intersection instead of their overlap.

3. Mass terms and ensembles

Let us see how ensembles could be used in the formal semantic analysis of mass noun expressions such as those considered in section 1.

My analysis will consist of giving a translation of such expressions in a formal language such that their formal (i.e. non-lexical) semantic properties are brought out precisely. For the points that I want to analyse in this paper it is sufficient to take for this formal language a simple extension of the language of first-order predicate logic, the extension consisting of the use of the $\subseteq$-, ϵ - and = - relations, the set notation $\{z \epsilon x | ...\}$ and the ensemble notation $[z \subseteq x | ...]$ I call this language Ensemble Language (EL)[6].

Consider sentence (2a):

(2a) The ice cream in the blue dish weighs 15 kilogrammes

In terms of ensembles there is a simple way of referring to the totality of all the ice cream in the blue dish as follows. Let IN BLUE DISH be a 1-place predicate constant, expressing that its argument is " in the blue dish", and let ICE be a constant referring to (naming) the totality of all ice cream. According to theorem 1, there exists a uniquely determined ensemble

(22) $[x \subseteq \text{ICE} \mid \text{IN BLUE DISH}(x)]$

which, according to corollary 1, is the union of all the ice cream parts in the blue dish. The ensemble denoted by (22) would thus correspond to what is meant by the English phrase: "the whole (totality) of all the ice cream in the blue dish", or simply: "the ice cream in the blue dish". Notice the striking structural similarity between (22) and the corresponding translation of the count noun phrase "the oranges in the blue dish":

(22a) $\{x \in \text{ORANGES} \mid \text{IN BLUE DISH}(x)\}$

The analysis of the complete sentence (2a) would involve, besides the translation of the mass noun phrase as (22), a translation of the amount phrase "15 kilogrammes". This must be postponed until section 5; in the next section I will first develop a formal analysis of amount terms.

4. Amounts

4.1. A definition of amounts

So-called amount terms occur in English and in many other languages in two kinds of context: in isolation, as in sentence (2a), or followed by "of" plus a noun phrase, as in (3a):

(2a) The ice cream in the blue dish weighs 5 kilogrammes

(3a) There is 5 kilogrammes of ice cream in the blue dish

Following Parsons (1970) I call these kinds of occurrence <u>isolated</u> and <u>applied</u>, respectively. I will present a formal treatment of both kinds, taking as my point of departure that in both cases an amount term can be viewed as denoting an abstract entity, an 'amount' (though for isolated amount terms I think it would be better to speak of 'magnitudes').

One of the most conspicuous features of amounts is that some amounts can be compared, e.g. 7 kilometres is more than 2 miles (but 7 kilometres is not more, less or equal to 7 kilogrammes), and can be added and subtracted, e.g. 7 kilometres minus 300 metres is 6 kilometres and 700 metres. Numbers of kilometres, miles and metres can be compared, added and subtracted because these units belong to the same dimension (length) and units of the same dimension are numerically related: a kilometre is 1000 metres, a mile is 1.6 kilometres, etc. Because of these numerical factors, 'conversion factors', it is possible to convert from one unit to another.

For a given set U of units of a certain dimension, the conversion factors relating the units can be specified by means of a function assigning to each pair of units the conversion factor that must be taken into account upon conversion from the first of these units to the second. For instance, for the dimension length the conversion function F_L would say: F_L(miles, kilometres) = 1.6. A conversion function must take into account that it is always allowed, when converting from a unit u_1 to a unit u_2, to first convert from u_1 to an arbitrary unit u' and subsequently from u' to u_2. More precisely, a conversion function F for a set U of units, i.e. a function from UxU into the set R^+ of positive real numbers[7)], must satisfy:

(24) $F(x,y) \cdot F(y,z) = F(x,z)$ for all $x, y, z \in U$

where the dot represents ordinary multiplication of numbers. A function satisfying (24) I call transitive, and I denote the set of all transitive functions from UxU into R^+ by $\tau r(U)$.

A dimension is now formally defined as follows.

Definition. A dimension D is a pair (U, F) with

(i) U a set, the elements of which are called D-units,

(ii) F a function belonging to $\tau r(U)$.

The conversion function F of a dimension D = (U, F) can be used to define a relation $=_F$ between pairs consisting of a positive real number and a D-unit as follows:

(25) $(n_1, u_1) =_F (u_2, u_2)$ iff $n_1 . F(u_1, u_2) = n_2$

The relation $=_F$ is easily proved to be an equivalence relation. Therefore, the relation $=_F$ determines a partitioning of $R^+ xU$ (see e.g. Birkhoff & MacLane, 1967, p.21) into equivalence classes. I will designate the set of equivalence classes in $R^+ xU$ determined by the relation $=_F$, defined by the dimension D = (U, F), by A_D. The elements of A_D will be called D-amounts.

Example. In the British system of lenghts, some of the units and their conversion factors are; yards, foot and inches with 1 yard being 3 foot and 1 foot 12 inches. This can be modelled by the following dimension $= (U_L, F_L)$:

(26) U_L = { yards, foot, inches }
F_L ={ ((yards, foot), 3),
((foot, yards), $^1/_3$),
((foot, inches), 12),
((inches, foot), $^1/_{12}$),
... }

Now the following equivalence class is an example of a weight-amount:
{(5, yards), (15, foot), (180, inches)}.

I will say that the elements of a D-amount are representations of that amount, and I will use the notation $\langle$ 3, inches $\rangle_D$ to designate the amount represented by (3, inches).

4.2. A calculus of amounts

In this section the comparison (i.e. ordering), addition and subtraction of amounts of the same dimension are formally defined.

Let D be the dimension D = (U, F).

Let zero-D-amount $\emptyset_D$ is defined by:

(27) $\emptyset_D = [\![o, u]\!]_D$ for an arbitrary $u \in U$.

It is easily proved that $[\![o, u]\!]_D = [\![o\ u']\!]_D$ for arbitrary $u, u' \in U$, in other words that $\emptyset_D$ is uniquely defined by (27), and also that for all pairs (n,u) with $n \in R^+$, $u \in U$ $(n,u) \in \emptyset_D$ iff $n = o$ (the number zero).

Between pairs (n_1, u_1), (n_2, u_2) with $n_1, n_2 \in R^+$ and $u_1, u_2 \in U$ an ordering relation $\leqslant_F$ is defined by:

(28) $(n_1, u_1) \leqslant_F (n_2, u_2)$ iff $n_1 . F(u_1, u_2) \leq n_2$

An <u>ordering</u> relation between D-amounts is now defined as follows:

(29) $a_D < a'_D$ iff $(\exists x \in a_D)(\exists x' \in a'_D)(\ x \leqslant_F x'\ \&\ \neg(x =_F x')\ )$

For any two D-amounts a_D and a'_D it is either the case that $a_D < a'_D$ or that $a'_D < a_D$ or that $a_D = a'_D$.

An addition of pairs (n_1,u_1) and (n_2,u_2) is defined as follows:

(30) $(n_1,u_1) +_F (n_2,u_2) = (n_3,u_3)$ iff $n_3 = n_1 . F(u_1,u_3) + n_2 . F(u_2,u_3)$

In terms of this operation the <u>addition</u> of D-amounts can be defined by:

(31) $a_D + b_D = c_D$ iff $(\exists x_a \in a_D, \exists x_b \in b_D, \exists x_c \in c_D)(x_a +_F x_b = x_c)$

<u>Subtraction</u> of D-amounts can be defined similarly, provided that we take care to subtract from a D-amount only D-amounts that are smaller in the sense of (29) or equal, in which case the result is $\emptyset_D$.

The <u>multiplication</u> of a D-amount with a non-negative number (which is useful for analysing expressions like "twice as much", "half as much") can be defined as follows. For any number $k \in R^+$ and D-amount $a_D \in A_D$, their product $k.a_D$ is defined by:

(32) $k.a_D = \{(n,u) \mid (\exists m \in R^+)\ (n = k.m\ \&\ (m,u) \in a_D)\}$

It is easily proved that $k.a_D$ is again a D-amount, i.e. $k.a_D \in A_D$ and that for any $k, n \in R^+$, $u \in U$:

(33) $k.[\![n,u]\!]_D = [\![k.n,u]\!]_D$

It is also possible to define the multiplication of a unit and a number (think of the unit gramme multiplied with the number 1000, resulting in the unit kilogramme) and to define operations on units for forming complex units such as "dollar per gallon", "metre per second^2", etc. However, for the problems considered in this paper we don't need these extensions.

4.3 Measure functions

D-amounts are meant to be useful in particular for analysing certain types of quantified mass noun expressions; we want to measure such things as icecream parts along as certain dimension D. This means that in the formal semantic analysis we must relate ensembles and D-amounts, and a natural way to do so is by means of functions, defined on ensembles and having D-amounts as values. I call functions, suitable for this purpose, measure functions.

Consider a certain ensemble E and a dimension D along which parts of E can be measured. Not every function from E-parts to D-amounts is suitable for that purpose; in order to be a suitable measure function we must impose on a function the requirements of being additive, which means that, applied to the union of two (or more) disjoint ensembles its value equals the sum of the values for these ensembles and of having the value ϕ_D only if applied to the empty ensemble. Formally, a measure function μ is defined as a function from the set of E-parts into the set A_D, i.e. satisfying the requirements:

(34) $(\forall x,y \in E)\ ((x \cap y = \phi) \rightarrow \mu(x \cup y) = \mu(x) + \mu(y))$

$(\forall x \subseteq E)\ (x \neq \phi) \rightarrow \mu(x) > \phi_D)$

This concept of a measure function is a straightforward extension of the usual mathematical concept of a measure (see e.g. Halmos, 1958).

Two useful theorems concerning measure functions, that follow from (34) and the logic of ensembles, are:

Theorem 2. If μ is a measure function that assigns to at least one ensemble in its domain a finite value[8], then $\mu(\emptyset) = \emptyset_D$.

Theorem 2. Let A and B be two ensembles and μ a measure function defined for A and B, with $\mu(A)$ finite. If $A \subseteq B$ and $\mu(A) = \mu(B)$ then $A = B$.

4.4 Amounts of something

D-amounts and measure functions, as developed above, express how much there is of what the measure function is applied to in terms of a certain dimension. When we consider a sentence like:

(35) The amount of gold in John's ring is the same as the amount of gold in Mary's ring,

there is a strong suggestion that we should make an analysis in terms of a concept 'amount of gold' which is not tied to a particular dimension. It should not matter, for (35) to be true, along which dimension the quantities of gold in question are measured. Cartwright (1970, p.39) expresses this view when she says that for a mass noun B the notion 'amount of B' should be such that:

(36) "Given any two (appropriate) measures of B, m_1 and m_2, the values of m_1 are the same for amounts of B, x and y, just in case the values of m_2 are the same for x and y."

It may be noted that when we are considering 'amounts of B' for a certain B, it is possible to compare amounts along different dimensions: 2 kilogrammes of gold is less than 2 m^3 of gold. I think this possibility exists because, although kilogrammes and cubic metres in general are independent notions, when applied to e.g. gold they are related since gold has a specific weight: a m^3 of gold weighs a certain number of kilogrammes which is characteristic for gold[9]. In general, when M is a mass term and $D_1, D_2, D_3, \ldots D_n$ are dimensions along which M-parts can be measured, we can speak of 'amounts of M' in a way that cuts across the dimensions $D_1 \ldots D_n$ because M determines a relation between the units of the different dimensions.

This notion 'amounts of M' can be formalised as follows.
Let $D_1, D_2, D_3, \ldots D_n$ be dimensions along which M-parts can be measured, with $D_i = (U_i, F_i)$ and $U_i \cap U_j = \emptyset$ for $i \neq j$. Let the relation between the units of $U_1, U_2, U_3, \ldots U_n$, determined by M be described by a function S_M, having as its domain a set of pairs (u_{ik}, u_{jl}) from $U_i x U_j$ such that for each U_i there is one element u_{ik} occurring at the left hand in a pair and one element u_{im} occurring as the right-hand element of a pair, and having non-negative real numbers as values.

On the basis of S_M and the conversion functions $F_1, F_2, F_3, \ldots F_n$ a function F_M belonging to $\tau r(U_1 \cup U_2 \cup U_3 \cup \ldots \cup U_n)$ can be defined as follows:

(37) 1. Let the pair (u_{ik}, u_{jl}) from $U_i x U_j$ belong to the domain of S_M.
Then for an arbitrary pair (u_{im}, u_{js}) F_M is defined as:
$F_M(u_{im}, u_{js}) = F_i(u_{im}, u_{ik}) . S_M(u_{ik}, u_{jl}) . F_j(u_{jl}, u_{js})$
2. Let, moreover, the pair (u_{jm}, u_{kp}) belong to the domain of S_M.
Then for an arbitrary pair $(u_{ir}, u_{kn}) \in U_i x U_k$, F_M is defined as:
$F_M(u_{ir}, u_{kn}) = F_M(u_{ir}, u_{jm}) . S_M(u_{jm}, u_{kp}) . F_k(u_{kp}, u_{kn})$
3. And so on.

In much the same way as a conversion function this function F_M can be used to define an equivalence relation $\underset{M}{\equiv}$, in this case an equivalence relation between D_i-amounts and D_j-amounts, as follows:

(38) $$a_{D_i} \underset{M}{\equiv} b_{D_j} \quad \text{iff } (\exists n_1, n_2 \in R^+)(\exists u_{ik} \in U_i, u_{jl} \in U_j)(n_1 . F_M(u_{ik}, u_{jl}) = n_2)$$

For instance, if M is CHEESE and S_{CHEESE} (litres, kilogrammes) = 1.5 then F_{CHEESE} (litres, pounds) = S_{CHEESE} (litres, kilogrammes). F_{WEIGHT} (kilogrammes, pounds) = 3, so the volume-amount $\langle 4$, litres $\rangle_V$ is equivalent to the weight-amount $\langle 12$, pounds $\rangle_W$.

The equivalence relation $\equiv_M$ determines a partitioning of all the different kinds of D-amounts, $A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n$. I denote the set of these equivalence classes by A_M and consider these classes as the formalisation of amounts of M. So for example $[\![4, \text{litres}]\!]_V$ and $[\![12, \text{poinds}]\!]_W$ are instances of the same amount of cheese. I indicate such an amount as $[\![4, \text{litres}]\!]_{\text{CHEESE}}$.

The relation between the units of the various dimensions, established by S_M, entails also a relation between the measure funtions for the dimensions in question. A cheese-part, measured as to its volume giving the value $[\![3, \text{litres}]\!]_V$, must give the value $[\![12, \text{pounds}]\!]_W$ when measured as to its weight. Generally speaking, if μ and μ' are the measure functions defined on M-parts for the dimensions D and D', the for any part $x \subseteq M$:

(39) $\mu_1(x) \equiv_M \mu_2(x)$

This means that, if $\mu_1, \mu_2, \mu_3 \ldots \mu_n$ are the measure functions defined on M-parts, a general 'amount function' μ_M from the set of all M-parts into A_M can be defined by:

(40) $\mu_M(x) = \{\mu_i(x) \mid i = 1, 2, 3, \ldots n\}$ for any $x \subseteq M$

Instead of $\mu_M(x)$ I will also write $|x|_M$.

It is easy to see that this general amount function inherits the properties (34) from the measure functions in terms of which it is defined and also satisfies the theorems 2 and 3.

It may be noted that, according to (38), the formal notion 'amount of M' defined here satisfies Cartwright's requirement (36). Russell (1937), who also discussed the requirements that a notion 'amount of something' (which he calls 'magnitudes of a kind') should meet, remarks that of any two amounts of M either the one must be less than the other or else they must be identical (rather than equal). This requirement is also fulfilled by the notion amount of M', formally defined here.

5. The analysis of mass terms and amount terms.

Let us now turn to some examples of the use of ensembles, amounts and measures in the formal semantic analysis of sentences with mass terms and amount terms. To this end I extend the syntax (and semantics, accordingly) of the analysis language EL with the expressions $\langle$ n, u $\rangle_D$ for D-amounts and $|x|_M$ for amounts of M.

Let us take first the sentences containing mass terms considered in section 1.

Concerning sentence (1a):

(1a) The ice cream in the blue dish comes from Italy

It was already noted in section 3 that the mass-NP "the ice cream in the blue dish" is analysed in terms of ensembles as (22).

(22) $[x \subseteq$ ICE | IN BLUE DISH $(x)]$

If we agree with the view, expressed in section 1, that "the ice cream in the blue dish" can well be taken to refer to a single entity, sentence (1a) is simply analysed in EL as (38), using the predicate constant ITALIAN for "coming from Italy":

(38) ITALIAN ($[x \subseteq$ ICE | IN BLUE DISH $(x)]$)

Sentence (2a):

(2a) The ice cream in the blue dish weighs 5 kilogrammes

only differs from (1a) in the predicate. The isolated amount term "5 kilogrammes" is analysed in EL as:

(39) $\langle$ 5, kilogrammes $\rangle_W$

and the ensemble (22) and the weight-amount (39) are tied together by the measure function in question, μ_W. So the analysis of the complete sentence is:

(40) μ_W ($[x \subseteq$ ICE | IN BLUE DISH $(x)]$) = $\langle$ 5, kilogrammes $\rangle_W$

Sentence (3a):

(3a) There is 5 kilogrammes of ice cream in the blue dish

is analysed as: there is an ice cream part in the blue dish weighing 5 kilogrammes, which corresponds in EL with:

(41) $(\exists x \subseteq ICE)\ (IN\ BLUE\ DISH(x)\ \&\ \mu_W(x) = \Sigma\ 5,\ kilogrammes\ \Sigma_W)$

Intuitively, sentence (3a) expresses a logical consequence of sentence (2a). A requirement on the formal semantic analysis of these sentences is that this logical relation comes out. Indeed, formula (41) is a consequence of formula (40) according to ensemble theory. The proof comes down to showing that the ice cream part called "x" in (41) is the ensemble denoted by (22). Since this ensemble is the union of a number of parts of ICE (see (13)), it is itself a part of ICE. Since, as we may assume, the union of ice cream parts that are in the blue dish is again in the blue dish, this ensemble itself is in the blue dish[10]. Moreover, according to (40), it weighs 5 kilogrammes.

Sentence (5a):

(5a) All the ice cream in the blue dish comes from Italy

Can be analysed simply as a distributive quantification over all the ice cream parts in the blue dish, i.e., formally, as:

(42) $(\forall y \subseteq [x \subseteq ICE \mid IN\ BLUE\ DISH(x)])\ (ITALIAN(y))$

It was noted in section 4.4 that a sentence like (35):

(35) The amount of gold in John's ring is the same as the amount of gold in Mary's ring.

Seems most appropriately analysed in terms of a dimension independent notion 'amount of gold'. Using the formalisation of this notion, developed in 4.4, and using EL predicate constants IJR, IMR for "in John's ring" and "in Mary's ring", resp., we get the following analysis of (35):

(43) $|[x \subseteq GOLD \mid IJR(x)]|_{GOLD} = |[y \subseteq GOLD \mid IMR(y)]|_{GOLD}$

For this analysis to be adequate, it should e.g. bring out that (35) plus (44) implies (45):

(44) The gold in John's ring weighs 10 grammes

(45) The gold in Mary's ring weighs 10 grammes

In terms of EL, (43) and (46) should imply (47):

(46) μ_W ([$x \subseteq$ GOLD | IJR(x)]) = ⌶ 10, grammes $⌶_W$

(47) μ_W ([$y \subseteq$ GOLD IMR(y)]) = ⌶ 10, grammes $⌶_W$

That this is indeed the case is proved as follows. An amount of gold contains at most one weigh-amount; according to (46) the weight-amount in the amount of gold at the left-hand side in (43) is ⌶ 10, grammes $⌶_W$. According to (43), this is also the weight-amount in the amount of gold at the right-hand side. Since, according to definition (40), this weight-amount is the value of the function μ_W, applied to the gold-part in question, (47) follows.

Notes

1. See Bunt (1978) for a definition of 'mass terms'. The discussion in this paper is completely restricted to mass noun phrases. Bunt (1978) also contains a tentative proposal for defining count - and mass adjectives.
2. See Pelletier (1975) and (1978).
3. This matter is discussed in more detail in Bunt (1978) section 2.3.
4. The definition can easily be extended to more than two.
5. Again, the definition can also be given for more than two ensembles. I will write the merge of a_1, a_2, a_3,...as: $a_1 \cup a_2 \cup a_3$,....; also I write the merge of all ensembles that are the elements of a set S as $\cup$S, see e.g. (13).
6. The formal language EL is defined precisely in Bunt (1978) section 4.2. I have no pretensions with EL as a semantic representation language; it is only meant to be illustrative for the use of ensembles in a formal representation language . EL is in fact a highly simplified version of the semantic representation language used in the PHLIQA 1 language understanding system (see Medena et al.,1975). The work described in this paper was conducted as part of the PHLIQA 1 project.
7. More accurately, by R^+ I mean the set of all non-negative real numbers and the symbol + ∞, that is the non-negative half of the so-called <u>extended real numbers</u> (see e.g.

Halmos, 1958, p.2).

8. By a finite value I mean any D-amount $\langle n, u \rangle_D$ with $n < \infty$.
9. For many mass terms, such as "stuf=", "matter", "equipment ", etc. there are no such clear specific properties (see Cartwright, 1970, p. 37 ff. for a discussion of this point). In fact, this makes me doub+ about the usefulness of a dimension-indepedent notion 'amount of M'.
10. The role of the logical properties of predicates in the formal analysis of quantified mass noun expressions is discussed extensively in Bunt (1978), section 5.

<u>Appendix</u> . Axioms of ensemble theory

The following list of 10 axioms constitutes an axiomatisation of ensemble theory. In addition to the notations, introduced in section 2 a special notation is used here for expressing that there is an ensemble x that is the smallest ensemble for which formula $\psi(x)$ holds, smallest in the sense of the part-of relation: $(\underline{\exists} x)\ \psi(x)$ abbreviates $(\exists x)(\psi(x)\ \&\ (\forall y)(\psi(y) \rightarrow x \subseteq y))$.

1. Axiom of Transitivity.

$(\forall x,y,z)((x \subseteq y\ \&\ y \subseteq z) \rightarrow (x \subseteq z))$

2. Axiom of Extension.

$(\forall x,y)(((\forall z)((z \subseteq x\ \&\ z \neq \emptyset) \rightarrow (z \cap y \neq \emptyset)) \rightarrow (x \subseteq y))\ \&$

$ATOM(x) \rightarrow (x \subseteq x))$

3. Axiom of Unicles.

$(\forall x)((ATOM(x) \rightarrow (\exists! y)(y \in x))\ \&\ (\neg ATOM(x) \rightarrow \neg(\exists y)\ (y \in x)))$

4. Axiom of Substitutivity.

$(\forall x,y,z)((x \in z\ \&\ x = y) \rightarrow (y \in z))$

5. Axiom of Unions.

$(\forall x)((\exists y)\ (y \in x) \rightarrow (\underline{\exists} u)(\forall z)(z \in x \rightarrow z \subseteq u))$

6. Axiom of Replacement.

Let ψ be any well-formed expression with 2 free variables, such that $(\forall x,y,z)((\psi(x,y)\ \&\ \psi(x,z)) \rightarrow (y=z))$ i.e. ψ represents a function. $(\forall x)(\underline{\exists} y)(\forall z)(z \in y \leftrightarrow (\exists u)(u \in x\ \&\ \psi(u,z))$

7. Axiom of Powers.
$(\forall x)(\exists p)(\forall z)(z \subseteq x \rightarrow z \in p)$

8. Axiom of Pairs.
$(\forall x,y)(\exists p)(x \in p \ \& \ y \in p)$

9. Axiom of Infinity.
Let $\{z\}$ denote the atomic ensemble with unicle z.
$(\exists x)((\exists y)(y \in x) \ \& \ (\forall z)(z \in x \rightarrow \{z\} \in x))$

10. Axiom of Regularity.
$(\forall x)((\exists y)(y \in x) \rightarrow (\exists z)(z \in x \ \& \ z \cap x = \emptyset))$

References

G. Birkhoff, S. MacLane (1967) - Algebra.
MacMillan, London.

H.C. Bunt (1976) - The formal semantics of mass terms.
Papers from the IIIrd Scand.Conf. of Linguistics, Hanasaari.

H.C. Bunt, (1977) - Some recent developments in semantics
IPO Annual Progress Report 11.

H.C. Bunt (1978) - Ensembles and the formal semantic properties of mass terms.
In: Pelletier (1978).

H.M. Cartwright (1970) - Quantities
Philosophical Review 79.

A.A. Fraenkel, Y. Bar-Hill, A.Levy, D. van Dalen (1975) -
Foundations of set theory.
North-Holland, Amsterdam.

P.R. Halmos (1958) - Measure theory, 5th ed.
Van Nostrand, Princeton.

S.C. Kleene (1952) - Introduction to metamathematics
North-Holland, Amsterdam

P. Medema, W.J. Bronnenberg, H.C. Bunt, S.P.J. Landsbergen, R.J.H. Scha, W.J. Schoenmakers, E.P.C. van Utteren -
PHLIQA1: Multilevel semantics in question answering.
Am. J. of Comp. Linguistics, microfiche 32.

T. Parsons (1970) - An Analysis of mass terms and amount terms.
Foundations of language 6..

F.J. Pelletier (1975) - A bibliography of recent work on mass terms.
Synthese 31 nos. 3/4

F.J. Pelletier (ed) (1978) - Mass terms: some philosophical problems.
Reidel, Dordrecht.

W.V.O. Quine (1960) - Word and object.
M.I.T. Press, Cambridge, Mass.

B. Russell (1937) - The principles of mathematics, 2nd ed.
Allen and Unwin, London.

Ergative languages, reflexives and questions of formal grammar

by

Jerold A. Edmondson, Technical University Berlin

1. Introduction

Just as the title indicates, this paper is going to have three parts.* I will open with a general discussion of some features of language typology, i.e. the properties of the so-called ergative-absolutive languages. Of particular interest will be the traits of such languages as they cluster along the following two dimensions: (a) the encoding properties, e.g. how are the grammatical relations, subject, direct object, indirect object, etc. indicated; (b) the syntactic and control properties, e.g. what kinds of operations--conjunction reduction, complementation, etc.--do the languages allow? The second part of the paper will contain a discussion of the reflexive constructions in a number of the world's languages. Once again the analysis will be seen from the perspective of the two dimensions, encoding vs. syntactic/control properties. The third part of this paper turns to the implications of cross-linguistic variation for a universally valid descriptive apparatus of the type envisaged by Montague in "Universal grammar" (1974a).

The term universal grammar as employed by Montague lays the emphasis on the construction of a mathematically precise theory that can comprehensively capture the formal semantic and syntactic features of natural languages and logical languages simultaneously. Paramount in Montague's program is comprehensiveness. Keenan and others, while acknowledging the requirement of universality, have claimed that the language specific realizations of the effect of a universally valid principle may differ; zero variation must be regarded as only one possible realization of linguistic universals, cf. (1975a). Most importantly, Keenan has pointed out that a formal treatment of language phenomena can lead to predictions about what operations are possible in various kinds of languages, Keenan (1975b):

> ...that the task of formally defining logical structures (LS) for natural languages (NL) has a linguistic interest beyond the immediate one inherent in representing logical notions like entailment, presupposition, true answer, etc. The reason is that LS can be used as a basis for describing and in some cases explaining, certain kinds of syntactic variation across NL.

One of the problems that Keenan (1975b:410-13) uses to exemplify "logically expressive power" in natural language is the reflexive. After discussing a number of examples, he concludes (1975b:413):

> NL then are logically different in that they vary with regard to the NP positions that are naturally stipulated as being co-referential. But for the moment we can offer no general explanation for this variation.

This paper addresses an explanation of this question.

2. Ergative-absolutive languages

Most of the grammar fragments for which an explicit semantics has been developed are oriented on English and to a lesser extent on some West European languages. There can be no doubt, however, that these culturally and genetically related languages will yield categories of analysis that are strongly biased. In particular, these languages fail to exhibit very pronounced ergative-absolutive features. Among the group of clear ergative languages are: Eskimo, Chukchee, Burushaski, Tongan and many of the native languages of Australia and the Americas. West European languages tend for the most part to exhibit rather strong nominative-accusative features, which may be considered the opposite end of the spectrum. For example, PTQ makes implicit use of the direct object position in the derivation of intensional contexts, the direct object of <u>seek</u> can have <u>de dicto</u> readings. But, the grammatical relation direct object in some ergative-absolutive languages is rather different than in English. Should we not expect that this difference could influence the semantic outcome in such languages?

Having talked about ergativity for some time without defining it, I want to continue by stating how one recognizes an ergative-absolutive language. With respect to the encoding dimension, ergative-absolutive principles are at work if in that language the subject of an intransitive verb and the object of a transitive verb are indicated identically. Nominative-accusative languages, on the other hand, express the subject of a transitive verb and the subject of an intransitive verb identically. The surface realization or encoding of ergativity is typically: (a) an absolutive <u>case marker</u> on the intransitive subject and the transitive object with a different case marker (the ergative case) for transitive subjects, or (b) agreement in the verb that employs one marker for TV-object and

IV-subjects and a different marker for TV-subjects. Tongan and Yucatec illustrate ergative-absolutive encoding:

(1) Tongan (Malayo-Polynesian VSO language)

a.

Na'e	taak'i	'e	Sione	'a	e	puaka.
past	hit	erg	John	abs	def	pig

'John hit the pig.'

b.

'Oku	lahi	'a	Sione.
pres	big	abs	John

'John is big.'

(2) Yucatec (Mayan VSO language)

a.

Bin	in	kins-	oobe.
fut	1sg erg	kill	3pl abs

'I will kill them.'

b.

Bin	kink-	oobe.
fut	die	3pl abs

'They will die.'

If ergativity ended with the encoding traits, it might be regarded as a peculiar but not particularly interesting feature of some off-beat languages. More recent work, however, has demonstrated that ergativity can figure decisively in the permitted syntactic operations that stand at the disposal of a particular language. As is well-known from more familiar languages, being subject or object can determine what structures are potential candidates for some transformation (I use the concept transformation for the purposes of convenience only). In English, for example, conjunction reduction demands not only that the NP's involved be identical (co-referential) but that both also be nominatives:

(3) John hit the pig and died.
⇐John hit the pig and John died.
⇍John hit the pig and the pig died.

3 unambiguously asserts that John and not the pig has died. If the pig dies, then no conjunction reduction is possible, a pronoun must be retained. In Tongan, however, the identity of case is between absolutive NP's. If different cases are involved, then, just as in English, a pronoun is retained.

(4) a.

Na'e	taak'i	'e	Sione	'a	e	puaka	pea	mate.
past	hit	erg	John	abs	def	pig	and	die

'John hit the pig and the pig died.'

b.

Na'e	taak'i	'e	Sione	'a	e	puaka	pea	ne	mate.
past	hit	erg	John	abs	def	pig	and	he	die

'John hit the pig and he died.' ne could mean he or it.

This kind of example makes it look as if encoding of one sort entails automatically syntactic/control properties of the same type. This assumption, however, would be false. Basque, according to Lafitte (1962:668), allows conjunction reduction even when the subjects differ in case.

(5) a. Ama ethorri da eta erran du.
mother come has and told(me)has
abs

'Mother has come and told me.' (da='be' auxiliary with intransitive verbs and du='have' auxiliary with transitive verbs.)

b. Amak gakoa hartu du eta joan da.
mother key taken has and dissappeared has
erg

'Mother has taken the key and disappeared.'

An ergative NP and an absolutive NP can be simultaneously involved in conjunction reduction. The different auxiliaries reveal that the transitivity of the combined forms was different and nevertheless the deletion is possible. Basque thus exhibits ergative-absolutive encoding but with respect to this operation subject and not nominative decides.

I was also able to find a case of the opposite type, i.e. nominative-accusative encoding but ergative-absolutive control properties. Modern German is not normally considered to be ergative. Yet, when intransitive and transitive sentences are conjoined below the causative verb lassen, the deletion takes place according to ergative-absolutive principles.[1]

(6) Der König liess den Diener kommen und töten.
'the king had the servant come and kill

'The king had the servant come and be killed.'

The NP den Diener serves as the subject of the infinitive kommen but as the object of töten. The translation shows that in English the reduction is permitted only if the second infinitive is passivized.

3. Reflexives cross-linguistically

The terminology we have been employing in discussing the effects of ergativity can also be fruitfully applied to describing the reflexive construction. I will therefore consider reflexives from the viewpoint of encoding; how does the language indicate that the initiator of an action is also the sufferer of the action, i.e. patient.

I will then turn to the syntactic properties that reflexives display.

There are, first of all, a few languages that indicate the presence of a controller-reflexive link only in a very weak manner. Typically the encoding of the non-reflexive is used. These languages do not have sentence positions that are especially marked when there is identity between two terms, cf. Keenan (1975b:411) for other examples of this.

(7) Burmese (Tibeto-Burman SOV language)

a. John ha thu kou tha' te. 'John kills him/her.'
John subj. him obj. kill fin
her

b. John ha thu kou thu tha' te. 'John kills himself.'
John subj. him obj. him kill fin
her her

te is the final particle for assertions.

(8) Tongan (VSO Malayo-Polynesian language)

a. Na'e alo'i ia 'e Nuku. 'Nuku taught him/her.'
past teach him erg Nuku
her

b. Na'a ku ako'i pē au 'a au. 'I taught myself.'
past I teach only I abs I

The reflexive is apparently viewed in these languages as a kind of strengthened non-reflexive, the added strength being imparted by repeating the pronoun. In both of these languages the strengthening part may be omitted so that there is no encoding difference between John kills him and John kills himself. This latter situation appears to have also been the case in Anglo-Saxon.

For languages in which reflexive encoding is differentiated from non-reflexive encoding, I was able to confirm the taxonomy of types proposed in Faltz (1977): (a) noun reflexives, often a noun meaning 'core', 'head', 'body' or 'marrow' sometimes accompanied by a possessive pronoun; (b) pronoun reflexives, having no clearcut interpretation out of construction, i.e. synsemantic and not autosemantic expressions; and (c) verbal reflexives, a special form of the verbal paradigm is employed. There may, of course, be more than one strategy in a language. Some examples of the first kind of encoding are: English himself, German selbst, Basque burua 'head', Buginese alena 'body', Okinawan du:na 'body', !Xũ l'esi 'body' and possibly Japanese jibun. Some clear cases of pronoun reflexives are: German sich, Swedish sig, sin(a), sitt, Russian sebja, the

Romance forms *se*, *si*, *soi*, Finnish *itse* and Hungarian *magan*. It is not always easy to decide what forms are nouns and what forms are pronouns. Verbal reflexives are probably less familiar than the preceeding types. They are, however, quite widespread in Africa, Australia and the New World.

(9) Swahili (SVO Bantu language)

a.	u-	li-	mw-	ona.	'You saw him.'
	you	past	him/her	see	
b.	a-	li	ji-	ona.	'He saw himself.'
	he	past	self	see	

(10) Malagasy (VOS Malayo-Polynesian language)

a.	Isaky	ny	olona	dia	manasa	lamba.	'Everyone washes clothes.'
	Every	def	person	def pred	wash	clothes	
b.	Isaky	ny	olona	dia	misasa.		'Everyone washes himself.'
	Every	def	person	def pred	wash himself		

(11) Quechua (SOV Quechumaran language)

a.	mayλa-		n-	Ø	maki-	s-	n- ta.	'He washes his (another's) hands.'
	wash		3	3	hand	pl	acc	
b.	mayλa-	ku-	n-	Ø	maki-	s-	n- ta.	'He washes his own hands.'
	wash	ref	3	3	hand	pl	acc	

Faltz (1977) points out that it is most likely that reflexives arise in the nominal form and become in the course of historical change more and more grammaticalized, first becoming pronouns and then clitic pronouns and finally merging with the verb altogether. The history of English, at least, suggests this development. In the earliest stages the non-reflexives were used. Sometimes an intensifying *self* is present. (Examples from Traugott 1972:88.)

(12) a. *hu hi hi behealdan sceolden.* 'How they should direct themselves.'
how they them direct should.

b. *þa adrencte he hiene selfne.* 'Then he drowned himself.'
then drowned he him self

In Modern French the reflexive *se* is always clitic, with the exception of reflexives following indefinite controllers such as *on*, *nul*, *personne* and *quiconque tel* where *soi* may be found in non-clitic position, e.g. after prepositions. Latin, however, permitted *sē* after prepositions, *Dumnorīgem ad sē vocat* 'he calls Dumnorix to him'. Generally speaking, nominal reflexives have a larger set of permitted contexts in which they appear than do pronominal reflexives. Adding the form *selbst* in German, for example, enables controller-reflexive links to be established that are not normally found.

(13) a. ?Der liebe Gott überliess die Menschen sich.
'God abandoned mankind to himself (i.e. God).'

b. Der liebe Gott überliess die Menschen sich selbst.
'God abandoned mankind unto himself (i.e. mankind).'

German normally demands that the controller of a reflexive be a subject NP. With the strengthening selbst in 13b a direct object may also be possible.

Though much more could be said about the encoding of reflexives, this short introduction should suffice for orientation. I now turn to some of the syntactic properties of reflexives. Of particular interest will be the positions, i.e. subject, object, prepositionally governed NP's and genitives, in which a reflexive may appear.

It has been widely reported in the literature that various languages permit reflexives in various NP positions. Lees and Klima (1963) puzzled about the fact that English sometimes has self-pronouns after prepositions and sometimes the non-reflexive forms. The test sentence has even become famous.

(14) a. The soldiers ignored the smokescreen around them.

b. The soldiers threw up a smokescreen around themselves.

German and Swedish are more highly constrained than English as they demand reflexives after prepositions. Kuno (1973) reports that Japanese can even tolerate intersentential controller-reflexives linkages. Faltz (1977) has investigated this phenomena in a large number of languages. I tabulate now some of the data he unearthed supplemented by my own informant work.

Language	Permissible position of reflexive				
	Same subj.	obj.	prepositional	genitive	InterS
Gã		x	x	x	x
Japanese		x	x	x	x
Okinawan		x	x	x	x
Swedish		x	x	x	?x
Russian		x	x	x	?x
Latin		x	x	x	?x
German		x	x		
Dutch		x	x		
English		x	?x		
Dholuo		x	?x		
Jacaltec/ Yucatec		x			
French		x			

TABLE 1

The cases of intersentential controller-reflexive linkages in Swedish and Latin are highly marked and at some stage have disappeared. In Modern Swedish the Clausemate Constraint is usually observed. As Mona Lindau and I found, however, there are 18th century examples with intersentential reflexivization, cf. Edmondson/Lindau (1972). Classical Latin permitted intersentential reflexivization, the so-called "indirect reflexive", when the lower sentence expressed words or thoughts of the subject, cf. Allen/Greenough (1931:181). This feature, however, also disappeared in Late Latin.

Table 1 clearly shows an implicational pattern, direct objects are most accessible to reflexivization, followed by prepositionally governed NP's, genitives and finally intersentential terms. Same subject positions are not accessible to reflexivization. Even in cases for which the conditions would be given, reflexives are not found in this sentence position.

(15) Swedish
Jan och *sin/hans flickvän gick till stationen. sin=reflexive
'Jan$_i$ and his$_i$ girlfriend went to the station.'

But:

(16) Jan gick till stationen med sin flickvän.
Jan went to the station with his girlfriend.

And similarly in Japanese:

(17) Kichi to *jibun/kare no tsuma wa ikimasu. (acceptable jibun= speaker)
'Kichi$_i$ and his$_i$ wife are going.'

This observation also holds true for German and Bengali.

(18) Peter und die Freunde von *sich/ihm kommen mit. Sich=reflexive
'Peter$_i$ and the friends of him$_i$ are coming along.'

(19) Še o *nijer/tar istri ekta šap kete-silo. nijer = reflexive
'He$_i$ and his$_i$ wife a snake killed.

It must be seen as a surprising fact that languages that consistently require reflexives in the predicate, disallow them when such reflexive forms appear as a genitive or beneath a preposition in the same subject as their controller. This regularity appears extremely wide spread. I am aware of only one clear counterexample, Kurdish.

(20) Hêns û jina xwe çûn mal. xwe= reflexive form.
'Hans and wife his go home we = non-reflexive.

The reflexive form of the pronoun is acceptable in this case as is also the non-reflexive.

I'd now like to address the status of reflexive sentences with respect to transitivity. Some philosophers, notably Geach (1962, 1965 and 1972), have been arguing that on semantic grounds sentences containing reflexives are in fact intransitive despite their outward appearance. Geach observes that the sentences 5 is divisible by 5 and 1 and 3 is divisible by 3 and 1 share in common the structure -- is divisible by -- and 1, which is usually paraphrased by the one-place predicate --is prime. To put this claim in a more testable form, reflexives are not terms, that is noun phrases, and do not refer to things; they are instead mere place reducers of transitive verbs. Semantically reflexives would then be functions mapping two-place properties onto one-place properties. In more recent work Partee (1975) has argued that reflexives in English behave like bound variables and not like full NP's. This claim, in my opinion, is a further manifestation of the non-terms status of reflexives. The question of interest, however, is what cross-linguistic evidence is there that supports non-termhood for reflexives.

First let us consider some wellknown but still puzzling facts. In the Scandanavian and Slavic languages unstressed forms of the reflexive pronoun have become attached to verbs to create a kind of medio-passive, i.e. Han kalla sig Jan 'He calls himself Jan' has become Han kallas Jan 'He is called Jan'. Many West European languages possess reflexive verbs that obligatorily co-occur with a form identical with the reflexive pronoun. This pronoun has, however, only a kind of ghost status since it can usually not be questioned or emphasized. In some languages it must also remain in close proximity to the verb. Thirdly, in French verbs with the reflexive pronoun require the auxiliary être and not avoir just as most of the common intransitives, e.g. rester, venir and aller. All of these cases speak for a one-actant status for reflexive forms.

Ergative languages provide us with some very useful evidence in this issue. Of those ergative languages I have consulted many with verbally encoded reflexives have subjects with absolutive markers, indicating that these languages regard TV + reflexive as intransitive. Typical of this type of language is Yanyuwa, an Australian aboriginal language, cf. Kirton (1976)

(21) a. Ganj-ilu-wuduřuma-Ø nja-miṇiŋiya-Ø nju-mařalŋudji-
him he feed past masc man abs masc doctor

lu. 'The doctor fed the man.'
erg

b. Gumba-wuduřuma-Ø nja-miṇiŋiya-Ø
refl. feed past masc man abs

'The man fed himself.'

Nearly identical structures are found in Chukchee, Eskimo, Yidiɲ and Dyirbal. In Navajo the verbally encoded reflexive is always accompanied by a detransitivizer, cf. Young/Morgan (1976:50,62):

(22) a. yoo- 'į́. 'He sees him.'
he-him see

b. 'ádoo- t- 'į́. 'He sees himself.'
he-self detr. see

When the reflexive is encoded as a noun or pronoun, ergative languages generally regard the verb as transitive, i.e. the subject takes an ergative marker. This is, however, not the entire story. There is some evidence that despite encoding traits of a transitive verb sentences containing reflexives treat such sentences syntactically as if they were intransitive.

Craig (1976:147,158) remarks on an exception to a rule of Jacaltec, an ergative Mayan VSO language spoken in Guatemala. Whenever the subject of a transitive verb undergoes Relative Deletion, Clefting or Question, i.e. is moved or deleted, a special marker is attached to the verb and the ergative pronoun in the verb is deleted. This special marking occur, however, only if the moved or deleted NP is a transitive subject, i.e. is marked ergatively, and does not occur to transitive objects or intransitive subjects, i.e. absolutively marked NP's. Notice that the NP's themselves contain no indication of case.

(23) a. x-Ø-w-il naj x-Ø-ul ☐ ewi.
asp-A3-E1-see cl asp-A3-arrive DEL yesterday.
'I saw the man who arrived yesterday.' (embedded S is intransitive.) naj and te' are classifiers and pronouns.

b. x-Ø-w-il te' tx'at x-Ø-s-watx'e naj ☐
asp-A3-E1-see cl bed asp-A3-E3-make cl DEL
'I saw the bed he made.' (embedded S is transitive.)

c. x-Ø-w-il naj x-Ø-☐-watx'e-n ☐ te'
asp-A3-E1-see cl asp-A3-DEL-make-suffix DEL cl
tx'at. 'I saw the man who made the bed.' (embedded
bed
S is transitive and the subject has been deleted.)

As can be seen in 23c, when the embedded subject is relativized upon, the suffix n(i) is added to the verb and the ergative pronoun is deleted. Craig argues that the addition of the suffix disambiguates possible confusion about the status of the following NP; it could be subject or object of the underlying embedded sentence.

But, if the NP relativized upon in the lower sentence is subject of a reflexive sentence, then the ergative marker in the verb is not deleted and the n(i) is not appended to the verb.

(24) Caw meba naj x-Ø-s-potx' s-ba [ ]
very poor cl asp-A3-E3-kill self DEL
'The man who killed himself was very poor.'

Even though the verb in 24 possesses ergative agreement, i.e. is transitive, the behavior of the constuction is like an intransitive. While Craig's disambiguation analysis is compatible with these data and also with my conception--clearly reflexives cannot be subjects and therefore ambiguity does not arise--, it seems to me that disambiguation is more an epiphenomenon than a phenomenon. Evidence for this claim is found in a passing remark that the n(i) suffix "is an intransitivizer suffix as shown by the intransitive future suffix -oj which must follow it." Craig (1976:146). This fact is difficult to accommodate in a disambiguation analysis, but follows clearly from the hypothesis that verbs + reflexives behave syntactically as if they were intransitives. A second intransitivizer would be superfluous.

4. Questions of formal grammar

In the preceeding sections I have tried to show that facts of language typology can influence not only the encoding of grammatical relations but also the behavior of structures with respect to syntactic operations. I hope I also have demonstrated that position is important in picking a form, reflexive or non-reflexive. Some positions are more accessible than others. Beyond that, reflexives when combined with a transitive verb yield structures that are fundamentally different than a transitive verb plus direct object, for example. I would now like to turn to questions of what influence these observations must have on a grammar with an explicit semantics.

I will begin somewhat circuitiously by looking at the question of how categorial grammars express grammatical relations. Relational grammarians such as Johnson (1976) have correctly criticized transformationalists for not placing the relations subject-of, direct object-of and indirect object-of in a central enough position in the theory. These categories are derivatively expressed in phrase structure grammars of the type used in TGT. Definitions like subject is the NP directly dominated by the S node encounter difficulty in VSO languages since it seems questionable that there is a VP constituent at all. If there is no VP constituent, however, then both subject and object would be dominated by the S node and no NP could be singled out as subject.

In categorial grammar choosing categories determines what can combine; dominance is not a meaningful concept. Nevertheless, it is easy to conceive of ways to relate category and grammatical function. For example, Montague selected the functor category t/IV for terms, leaving no elements in the basic category, e. A term is to be that functor category that turns intransitives into sentences. The important point about this choice with respect to grammatical relations is that the subject position is used to define the category of terms. This development must be considered crucial. Although Montague's motivation for this choice is his concern for an adequate semantics, with this step he provides us with the foundations of linking category and function. Putting it more pointedly, in categorial grammars where terms belong to the category t/IV, an architypical term is a subject.

Such categorial grammars differ both from relational grammars and from phrase structure grammars in interesting ways. First of all, categorial grammars of this type employ a grammatical function to define the category and not the reverse as in transformational grammar. The choice of category t/IV automatically gives the subject widest scope of any term and the most independence of the accompanying verb. And, most importantly, these properties are inherent in the category and are not superimposed upon it by means of some special configurational definition. On the other hand, the grammatical functions subject-of, direct object-of, etc. are not primitives of the descriptive system as in relational grammar. In this paper I will be arguing that one may want to consider categorial grammars instead of relational grammars for universal syntactic description because they contain fewer primitives and possess a more transparent semantics. Set theoretical counterparts of subject and direct object might ultimately be required in a truly universal grammar but it is unclear what sort of entities would be appropriate.

One difficulty for categorial grammars of the type we have been discussing is that terms clearly can occur in relations other than the subject relation. And, in order to account for just such cases, categorial grammarians have had to resort to various devices to enrich the syntax. Cresswell (1973) introduces lambda conversion as a syntactic mechanism, while Montague opts for altering the category of transitive verbs. Instead of picking the category t/(e, e), he chooses a category of higher order, i.e. IV/T.

As I am now going to argue, this complication of order of categories can be seen as the natural result of choosing one sentence position as definatory. If it can be shown that the resulting positions fall along the accessibility heirarchy: subject > direct object > indirect object > prepositionally governed > genitive, then categorial grammar will provide a close reflection of natural language phenomena.

Let us first consider direct objects. Since terms operate on intransitives, one solution to employing terms as objects is to make out of the transitive verb an intransitive; this is Cresswell's solution. Hit, a 2-place verb, is remodeled by means of lambda conversion to a 1-place verb, $\lambda y(\underline{hit}(x, y))$. Montague selects a category for the transitive verb that makes a term its argument. This choice accords well with the general tendency of objects to be much more dependent upon their verbs than subjects. The things objects denote can be denied reference if the verb is intensional or existence if the verb is negated, objects can be created or destroyed by their verbs. The price paid, however, is an elevation of the order of the category from a functor containing only basic categories, i.e. t/(e, e) to one of higher order, i.e. (t/e)/(t/(t/e)). By counting parentheses, we see that the latter expression is of the third order.

Under the assumption that prepositionally governed terms along with their prepositions are adverbs, we can also count the order of the category needed to operate on a term, i.e. a preposition is then ((t/e)/(t/e))/(t/(t/e)). The numerator of this expression is one level more complex than that of direct objects.

Finally, let us consider genitives. I shall assume that genitives function as functors that make a term into a determiner. In most languages John's car is just as definite as the car. The genitive operator is realized in English as 's but may not always have an overt morpheme in other languages. The category for the genitive would then be ((t/e)/(t/(t/e)))/(t/(t/e)) = (T/CN)/T.

As I have just demonstrated, the complexity of the functors that must operate on terms to make out them connex strings varies in a manner parallel to the accessibility of the terms to syntactic operations. The reason for the fanning out of categories in such a grammar is that terms must be made into something that is combinable with other sentence parts. Metaphorically speaking, terms are located at varying depths in the structure, cf. Bartsch/Vennemann (1972) on argument raising.

What category do we want to assign to reflexives? I have tried to show that there are good reasons to believe that reflexives in some languages or reflexives in some uses differ from terms, e.g. direct objects. This occurs whenever the reflexive loses its semantic independence as a true referring expression and becomes a grammatical particle dependent on the verb. The test for such grammatical uses I take to be whether a sentence with a reflexive in the environment of only is ambiguous; if one can still interpret the reflexive as a proxy of a full NP, then it possesses more nominal features; if only sloppy identity readings are permitted, then it has altered its status, cf. Partee (1975). As Partee points out, himself can go proxy for a full NP only if it is stressed. Unstressed forms in English behave like bound variables.

(25) Only Lyndon voted for himself.

25 with an unstressed reflexive is compatible with the interpretation that Lyndon received more than one vote but was the only self-voter. I would like to propose that this kind of case is the typical use of reflexives and demands a category IV/TV, a functor mapping transitive verbs onto intransitive verbs. This choice of category can account for many of the cross-linguistic traits of reflexives we have discovered.

The first prediction this choice of category would make is that the unmarked position for reflexives is the object position, i.e. an operator on the 2-place verb. For those categorial grammars like Montague's, the reflexive would appear in a similar position as the direct object although the function-argument relation would be different. As indeed the cross-linguistic facts have demonstrated, the object position can always be reflexivized upon even when any other position is excluded. This choice also can account for the tendency of reflexives to cliticize onto the transitive verb.

The second prediction this choice makes is that the accessibility hierarchy is derivable and not a primitive. Since the architypical reflexive is in the category IV/TV, other positions must be associated with special functors that can join the reflexive to other sentence members. Prepositions, therefore, belong not only to the category of operators on terms but also to the category of operators on IV/TV, i.e. would be in the second category (IV/IV)/(IV/TV). Genitives map reflexives onto a kind of determiner just as terms are mapped onto a determiner. Reflexive genitives, however, have a peculiarity not found in non-reflexive genitives. As Cresswell (1973) points out, a quantifier phrase in the genitive position can react with the subject to produce scope ambiguity, i.e.

(26) Everyone loves someone's wife.

But, determiners that contain reflexives cannot be ambiguous in this way.

(27) a. Only John loves his own wife.

b. Only John loves his wife.

27a unlike 27b permits only the sloppy identity reading "John was the only one that loves own wife", cf. the discussion in Edmondson (1976). The conclusion to draw is that his own wife is itself an expression of the category IV/TV. With the internal structure (Gen(Refl))(CN), the category for genitive must be ((IV/TV)/CN)/(IV/TV). A comparison of the three categories, direct object, prepositionally governed reflexive and genitive shows that the order of the functors needed to operate on reflexives mirrors the accessibility hierarchy.

A third and most important prediction of this category choice is that reflexives should not be part of the subject. If I have made correct analyses of the preposition and genitive, then these functors always produce something that demands a verb to form connex strings. But, if such an expression winds up in the subject, then there is no way for it to combine with a verb since subjects take verbs as arguments.

This choice of category also more or less rules out intersentential reflexivization. In languages where such a linkage is permitted there is nearly always some additional non-reflexive sense attached to this form. In Latin and Greek, for example, these are the cases of the "indirect reflexive". If I understand Kuno (1973) correctly, the same must be said for Japanese. For such cases the reflexives present the situation from the perspective of the subject and not from the perspective of some observer, cf. Cantrall (1974).

Finally, this theory can explain why reflexives are to be associated with intransitives. The category itself intransitivizes. It is also certainly no accident that Cresswell (1973) assigns the passive operator to the same category as the reflexive, i.e. IV/TV.

Ergative languages must also suggest to us that some of the well-known semantic properties and their concomitant syntactic structures may require a different analysis than Montague gave to English. PTQ provides that direct objects are to be treated as arguments of the verb and subjects as functors on the verb. Yet, in ergative languages the encoding and syntactic/control properties of subject and object have a different distribution than in English. This fact must lead to the question whether this difference extends to the semantic properties of the type: (a) de dicto vs de re readings, and (b) preferentially narrow scope for the object term of a two quantifier sentence. In other words, does the encoding reflect a basically different stance as to what sentence positions can be influenced by the verb?

Ioup (1975:97) found that the grammatical relation entered into by a particular term was more of a factor in assigning scope preference than linear order. Subjects tend to receive a wide scope interpretation in preference to objects. Thus, in 28:

(28) Every sailor on the ship is in love with an attractive girl.

most speakers have definite priorities; every sailor has the wide scope reading more naturally than an attractive girl. But, in a passive sentence like 29:

(29) An attractive girl is loved by every sailor on the ship.

the priorities usually reverse. This point is important because ergative encoding has often been viewed as a case of passive in disguise. Might one then not expect basically opposite preferences in sentences with scope ambiguity? My investigation of ergative-absolutive traits and scope preference has only begun and therefore must be considered preliminary. But, in Tongan, at least, there appears to be no clear deviation from the pattern found in English.

(30)	Na'e	tamate'i	'e	he	ongo	tulimanu	e	lapisi	'e	tolu.
	past	kill	erg	def	two	hunter	def	rabbit	?	three

'Two hunters killed three rabbits.'

30 was interpreted by my informant to mean that a total of three rabbits were killed, which is also the usual reading in English.[2] The results are similar for putative cases of scope ambiguity between negation and a quantifer phrase as subject, e.g. All the men didn't come.

(31) a.	Na'e	'ikai	ha'u	kātoa	e	kau	tangatá.
	past	NEG	come	all	def		man

b.	Kātoa	e	kau	tangatá	na'e	'ikai	ha'u.
	all	def		man	past	NEG	come

kau accompanies indefinite quantities, (lit.) 'a score'.

31a and 31b represent unmarked word order and clefted word order respectively. Both mean 'no one came'. If anything, this result contradicts the prediction that subjects of intransitive sentences in ergative languages tend to be more under the influence of the verb, i.e. are more patient-like, than subjects in nominative-accusative languages. As a matter of fact, the interpretation 'not all the men came' is possible when the quantifier is floated to preverbal position:

(32) Na'e 'ikau kātoa mai e kau tangatá.
past NEG all here def man
'The men didn't all come.'

Let me now try to sum up some of the conclusions that one must make about the logical expressive power of reflexive forms. It appears that reflexives typically begin as nominal expressions and tend to become grammaticalized. Nominal reflexives are, therefore, somewhat overexpressive since they can have independence of the verb; pronominal reflexives and especially verbal reflexives have a great deal less freedom. This unused freedom isn't needed to express the intransitivity of reflexive sentences. It is, therefore, not surprising that reflexives end up a part of the verb or even as zero.

Notes

*Part of the data cited here was collected within the frame of the project "Universals and language data" sponsored by the Deutsche Forschungsgemeinschaft, Ed 9/1.

[1]There are a number of factors at work in such constructions in German beyond those of interest here. It seems to be necessary that only one term appear explicitly with the transitive verb.

(i) Der König liess den Diener kommen und jemand töten.

This sentence can mean only 'the king had the servant come and kill someone (not be killed by someone).'

[2]The particle 'e sometimes occurs in front of cardinal numbers, cf. Churchward (1953:172).

Bibliography

Allen, J. H./J. B. Greenough. 1931. New Latin grammar. Boston: Ginn and Co.

Bartsch, R./T. Vennemann. 1972. Semantic structures. Frankfurt a.M.: Athenäum Verlag.

Cantrall, W. 1974. Viewpoint, reflexives, and the nature of the noun phrase. The Hague: Mouton.

Churchward, C. M. 1953. Tongan grammar. Oxford Univ. Press.

Craig, C. G. 1976. Disambiguation and hierarchies in Jacaltec. Mayan linguistics, ed. by M. McClaran, 141-160. American Indian Studies Center, University of California, Los Angeles.

Cresswell, M. J. 1973. Logics and languages. London: Methuen and Co.

Edmondson, J. A. 1976. Strict and sloppy identity in lambda categorial grammar. Indiana Univ. Linguistics Club.

Faltz, L. 1977. Reflexivization: A study in universal syntax. Univ. of California, Berkeley, Diss.

Geach, P. 1962. Reference and generality. Ithaca: Cornell Univ. Press.

_____. 1965. Logical procedures and expressions. Ratio 7.2. Reprinted in 1972. Logic matters. Oxford: Blackwell's.

_____. 1972. A program for syntax. Semantics of natural language, eds. D. Davidson/G. Harman. Dordrecht: Reidel.

Ioup, G. L. 1975. The treatment of quantifier scope in a transformational grammar. Ann Arbor: Xerox Univ. Microfilms.

Johnson, D. E. 1976. Toward a theory of relational grammar. Indiana Univ. Linguistics Club.

Keenan, E. L. 1975a. The logical diversity of natural language.

_____. 1975b. Logical expressive power and syntactic variation in natural language. Formal semantics of natural language, ed. E. Keenan, 406-21. Cambridge Univ. Press.

Kirton, J. 1976. Yanyuwa nominative and ergative-allative cases. Pacific linguistics series A .1-11.

Kuno, S. 1973. The structure of the Japanese language. MIT Press.

Lafitte, P. 1962. Grammaire basque. Édition des "Amis du Musee Basque" et "IKAS".

Lees, R./E. Klima. 1963. Rules for English pronominalization. Modern studies in English, eds. D. R ibel/S. Shane. Englewood Cliffs.

Montague, R. 1974a. Universal grammar. Formal philosophy, ed. by R. Thomason, 222-246. Yale Univ. Press.

_____. 1974b. The proper treatment of quantification in ordinary English. Formal philosophy, ed. by R. Thomason, 247-270. Yale Univ. Press.

Partee, B. 1975. Deletion and variable binding. Formal semantics of natural language, ed. by E. Keenan, 16-34. Cambridge Univ. Press.

Traugott, E. 1972. The history of English syntax. Holt, Rinehart and Winston.

Young, R./W. Morgan. 1976. The Navaho language. Salt Lake City: The Desert Book Co.

MONTAGUE GRAMMAR AND PROGRAMMING LANGUAGES

by

Peter van Emde Boas

ITW/VPW - University of Amsterdam

and

Theo M. V. Janssen

Mathematical Centre Amsterdam

0. INTRODUCTION.

In this paper we discuss the interface between Montague grammar and the research field of semantics of programming languages. The existence of such an interface was realised back in 1976 by the second author. It appeared that certain problems in the semantics of programming languages could be solved by applying Montague's ideas in this field.

We will not present here a complete technical proposal; this has been done elsewhere (see [9,10]). Instead we explain for the non computer scientist the nature of the problems involved, and the way Montague grammar is used to solve them. This application is a prime example where the use of a possible-world-semantics is above all criticism. Moreover, this application inspired the development of some new tools which may be useful for natural languages as well.

The paper is organised as follows. Section 1 presents an introduction to programs, and the need for some formal semantics for programs. Some approaches to semantics are sketched in section 2. In section 3 we consider a particular aspect of programs : variables and assignment; their semantical treatment and the problems arising in their treatment are discussed in section 4. Section 5 presents some features of our proposal for a Montague-style semantics for variables and assignment. In section 6 we illustrate by examples how the problems explained before are solved in our system. Section 7 contains some tentative suggestions how the tools developped before may be applied for natural languages.

1. PROGRAMS AND THEIR SEMANTICS.

Programs are pieces of text written in some programming language. These languages were designed for the special purpose of instructing computers, or, indirectly, instructing human beings how to instruct computers. Sometimes they are used by humans to communicate algorithms among themselves.

There exist nowadays several thousends of mutually partly incompatible programming languages. Each has its private set of strange ideosyncrasies, design errors, perfectly nice ideas and clumsy conventions. A few type of instructions are present in most of the languages; the present paper deals with one of those instructions - assigning a value to a variable.

We discuss programs on base of some examples. These examples are written in the programming language ALGOL 68 .

Our first example computes by use of a well known formula the solutions of the quadratic equation $a\ x^2 + b\ x + c = 0$: $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4\ a\ c}}{2\ a}$

The program reads as follows :

```
program 1  :  begin  real  a,b,c,disc,d,x1,x2 ;
              read((a,b,c)) ;
              disc := b * b - 4 * a * c  ;
              d  :=  sqrt(disc);
              x1  :=  -b + d ;  x1 := x1/(2*a) ;
              x2  :=  -b - d ;  x2 := x2/(2*a) ;
              print((a,b,c,x1,x2,newline))
              end
```

The program illustrates the fact that the computer, in order to perform some usefull computations, must obtain data from outside (input) and communicate the results to the outside world (output). It also illustrates that the mathematical expression above looks much compacter than the program, but this compactness is made possible by the use of hidden mathematical conventions which have to be made explicit to the computer. For example b x denotes b times x but in the program we must write b * x . In the formula we use some 2-dimensional features which are eliminated in the program (sqrt(...) in stead of $\sqrt{....}$) ; such a linear characteristic is enforced by the fact that programs have to be communicated by way of a sequential channel , like the wire connecting the computer with a card reader. The program contains some words written in a peculiar character set like ***begin*** ; in the program these words are considered to be individual symbols just like a, b , ; , + , = , etc. .

Programming languages are formal languages with a complete formal definition for the syntax of the language. Such a definition specifies exactly when a string of symbols from the alphabet of the language is a program and when it is not. All tools from formal language theory are applicable, and , in fact, designers of programming languages form the largest user group of Chomskian-style rewriting rules since they were invented.

The definition of a programming language specifies also how a string of symbols which has the correct syntactical structure for being a program, should be executed on a computer, or formulated more generally, what the program is intended to do. The formal definition of a programming language should specify both the syntax and the semantics of the programs in the language.

There exist a large confusion among programmers, theoreticians, and designers what we should understand by the semantics of a programming language. First one should look for the relevant properties of programs to be dealt with by a formal semantics. For some of these properties there is a larger dergee of agreement ; these properties are :

correctness: a program should perform the task it is intended to perform. For example program 1 is incorrect: it does not account for a=0 or disc < 0 .

equivalence: two different programs may yield the same results in all circumstances; we like to explain or even prove this behavior. For example, in program 1 we may interchange the order of computation of x1 and x2 but we cannot compute d before we compute disc.

termination: if we start the execution of a program will it ever stop ?

In order to illustrate this aspect we need another example of a program. It is known in mathematics that the sequence defined by $x_1 = 2$, $x_{n+1} = (x_n + 2/x_n)/2$ is a decreasing sequence with limit $\sqrt{2}$.

In the program below the elements in this sequence are computed until the value gets sufficiently close to the limit.

```
program 2 :   begin  real  x := 2 ;
              while  ( x * x - 2 )  >  1.0E-14
                do   x := ( x + 2/x)/2   od  ;
              print(x)
              end
```

On base of the mathematical result that the sequence monotonically decreases to $\sqrt{2}$ we might conclude that the program terminates by achieving the required accuracy in the approximation of the limit.(Whether this prediction is justified when the program is executed on a concrete computer depends on the computer involved . Due to numerical inexactness the mathematical properties of the sequence yield no guarantee for termination.)

Each of the above aspects, correctness, equivalence and termination, tells us something on the possible computations invoked by the program when provided with input data which are presently not specified. We want to make predictions on what may happen in case that ... ; more specifically, we like to prove that our predictions are correct. How can we achieve this goal ?

Clearly it is impossible to try out all possible computations of the program. In stead one is tempted to run the program on a "representative" set of input data - this activity is known under the name *program debugging*. This way one may discouver errors, but one can never prove the program to be correct. Still, in practice, most programs being used in our world have been verified in this way.

One might try to understand the program by reading its text. Again this is not of great help since the mistakes made by the programmers can be remade by the reader.

The only way out is the invention of a mathematical theory for proving the correctness and other aspects. Therefore we need a formalised semantics on which such a theory can be based.

2. SOME APPROACHES TO SEMANTICS.

What does a formal semantics look like ? The most common approach is a so-called *operational semantics*. We define the meaning of a program by first describing some abstract machine (which is in fact a mathematical model of the real computer) and next specifying how the program is to be executed on the abstract machine.

Needless to say that in this way the problem is transfered from the real world to some formal idealistic world; the possibly infinitely many computations of the program remain as complex as before. On the other hand, it is by the use of an operational semantics that most of the existing programming languages have obtained their meaning.

For some 14 years there have been given so-called *denotational semantics* for programming languages (see EG. [16]). The meaning of a program is represented by a mathematical object in a model. Mostly this object is a function which describes the input-output relation of the program. In denotational semantics we abstract from the intermediate stages of the computation, and the model has far less resemblance to a machine than the abstract machines used in operational semantics.

In denotational semantics programs are not so much considered as transforming values into values, but transforming the entire state from the computer from some initial state to some final state. States are highly complex descriptions of all information present in the computer. In general we are only interested in a small part of the information in the computer (EG. the values of the input and output variables). This leads to the "dual" approach to semantics which uses the so-called *predicate transformers*. A predicate is a (mostly highly incomplete) desription of states ; a predicate specifies a set of states - all states for which the predicate holds true. We like to correlate desriptions of the state before execution of the program with descriptions of the state afterwards or vice versa. Given a partial description of the state before we like a description of the state afterwards, or conversily, given some description of the state we like to achieve, we want to find out from which initial states this goal state may be reached.

Both directions have been investigated. The meaning of a program is expressed using a forward or backbard predicate transformer.

For example consider program 1. We may describe the initial state by specifying that on the input channel there are present three numbers a, b and c such that $a \neq 0$ and $b^2 - 4ac \geq 0$. From this description we conclude that execution of program 1 leads to a state where x1 and x2 contain the solutions to the equation $a x^2 + b x + c = 0$. Note that both descriptions are incomplete in so far that the initial and final values of disc and d remain unspecified.

Conversily we may observe that, in order that after execution of the program x1 and x2 represent the solution of the equation $a x^2 + b x + c = 0$, it suffices to require that the coefficients a, b and c are present on the input channel (in this order !) before execution of program 1 and that moreover $a \neq 0$ and $b^2 - 4ac \geq 0$.

In the formalization we will forget about the input and output instructions in the program, and restrict ourselves to the real computation. In program 1 we will omit the 2nd and 7th line. The above observations are now formalised by the formula :

$$a \neq 0 \ \& \ (b^2 - 4ac) \geq 0 \ \{ \text{ program 1 } \} \ a x_1^2 + b x_1 + c = 0 \ \& \ a x_2^2 + b x_2 + c = 0 \ .$$

Here we use a convention introduced by Hoare [7] . The expression $\Phi \{ \Pi \} \Psi$ means that if we execute program Π starting in an initial state where predicate Φ holds true, and if the execution of the program terminates, then predicate Ψ will hold true in the resulting state.

The convention can be used in both directions: given Ψ one likes to find a formula Φ which specifies the correct precondition without specifying too much; this precondition is called the *weakest precondition* and is denoted by $wp(\Pi,\Psi)$. Conversily, given Φ one likes to obtain a formula Ψ which expresses as much as can be said of the resulting state; this so-called *strongest postcondition* is denoted $sp(\Pi,\Phi)$.

Mathematically $wp(\Pi,\Psi)$ is characterised by :

I $\quad wp(\Pi,\Psi) \{ \Pi \} \Psi$

II $\quad$ from $\Phi \{ \Pi \} \Psi$ we can derive $\Phi \Rightarrow wp(\Pi,\Psi)$.

The problem of program semantics now may be rephrased as :

Given Π *and* Ψ *compute* $wp(\Pi,\Psi)$ *based upon the structure of* Π *and* Ψ .

The phrase "based upon the structure of Π and Ψ " requires some further explication.

It would be useless to have a semantics which attaches to each program Π and each postcondition Ψ a weakest precondition by an ad-hoc method. We must use the fact that the class of programs (as well as the class of predicates) is formed in a structured way by way of the grammar used for the syntax of the programming language. For example, if Π_1 and Π_2 are programs then $\Pi_1 ; \Pi_2$ is a program as well ; it is executed by first executing Π_1 and next executing Π_2 beginning in the final state of the first computation. This program is called the *sequential composition* of Π_1 and Π_2 . Clearly the structure of this composition must be retrievable from the semantics. This is indeed the case as can been concluded from the following rules :

$$\Phi \{ \Pi_1 \} \Xi \ \& \ \Xi \{ \Pi_2 \} \Psi \ \Rightarrow \Phi \{ \Pi_1 ; \Pi_2 \} \Psi$$

or

$$wp(\Pi_1;\Pi_2 , \Psi) = wp(\Pi_1, wp(\Pi_2,\Psi))$$

or

$$sp(\Pi_1;\Pi_2 , \Phi) = sp(\Pi_2, sp(\Pi_1,\Phi))$$

So what is practised here is nothing but Frege's principle of semantical compositionality : the meaning of a compound expression must be a composition of the meanings of its constituent parts.

3. VARIABLES AND ASSIGNMENTS.

In order to give the reader some intuition about the subjects treated in the sequel we use a model for an Abstract Computer. Our model consists of an infinite number of *registers*, each capable of storing an integral number. The registers are numbered by integers; the number of a register is called its *address*, whereas the number stored in the register is called its *value*.

The register with address j is denoted M[j] (M comes from the anthropomorphic term Memory). In many circumstances we will denote the value of register M[j] by M[j] as well. This ambiguity wich repeats itself in most existing programming languages is the origin of some of the problems dealt with in the present paper.

An essential feature in the behaviour of a computer is the possibility to modify the values of the registers by executing an instruction. Suppose that in a particular state it is the case that M[5] = 1 , M[6] = 3, M[7] = 5, and M[8] = 7 . If we now execute the instruction :

M[5]:= M[7] + M[8]

we have to replace the value of M[5] by the sum of the values of M[7] and M[8] , so M[5] now contains the value 12 . Next execute the instruction :

M[5]:= M[5] + 1

The meaning of this instruction is that the value of M[5] is increased by one, so afterwards the value of M[5] equals 13 .

The above instructions may be represented in the programming language by statements like : x := y + z and x := x + 1 or alternatively x+:=1 . the identifiers x,y, and z denote individual registers which remain fixed during the execution of the program. This method of addressing is called *direct addressing.* On most computers there exist other ways of addressing. We discuss also the methods of *incremental addressing* and *indirect addressing* . For the method of incremental addressing one of the general registers is used for the following special purpose: its value is added to the address given explicitly in the instruction, and the resulting integer is taken as address of the instruction. The special register is called an *index register* and is denoted by B . Assume by way of example that B is the register M[6] , and consider the instruction :

M[B+2] := 15

Since the value of B equals 3 this instruction is in the state considered equivalent with M[5]:= 15 ; after execution the value of M[5] has become 15 and all other registers keep their former value.

In the programming language incremental addressing is indicated by the use of *arrays* and *subscripted variables* . The former instruction may be denoted by : a[3]:= 15 . The array a denotes an infinite list of variables indexed by use of an index which is an integer. In executing the instruction the value of the index is stored in the index register B ; the constant 2 is added to the value of B indicating that the array a starts at address M[3] in the computer. In this way all instructions involving array a are of the form M[B+2] := .... ; for some other array this may look like M[B+44] := ..... .

In the programming language the index of the array element may be a complicated expression; in this way statements like a[a[x+4]-5]:= y+a[14] can be obtained.

For the method of indirect addressing the value of a general register is used as the address of an instruction. For example in the instruction :

M[M[7]] := 6

the value of M[7] (which equals 5) is taken as address for the instruction; the result is that the value of M[5] becomes again equal to 6 .

Indirect addressing is represented in the programming language by the use of *pointers* . Pointers are variables whose value don't represent plain integers but addresses within the computer. In our abstract computer it is impossible to see from the value which is stored in some particular register whether it represents an integer or an address. (the same is true for most real computers as well). In order to prevent the unintended use of variables for adresses or vice versa, some programming languages require for each identifier a specification of the kind of values to be stored in the corresponding register. This way there is created an artificial distinction between the identifiers occurring in the program.

In the programming language ALGOL 68 this typing has been worked out completely. We have among others the types ***int*** for plain integers, ***ref int*** for integral variables, ***ref ref int*** for pointers referring to integral variables, ***ref*** [ ]***int*** (read ref row int) for array variables, etc. .

The general *assignment instruction* in ALGOL 68 has the form :

$\chi := \tau$

The *destination* χ denotes some variable of a specific type ***ref*** Σ whereas the *source* τ denotes some expression of the type Σ . The instruction requires that the value of χ is made equal to the value of τ . It is allowed that the type of τ is not equal to Σ but of type ***ref*** Σ or even ***ref ref*** Σ etc.. In these circumstances the superfluous ***ref***'s are removed by the action of *dereferencing* ; we deliver the value of a variable instead of the variable itself. In programs like x:=x+1 , or x:=y+z this dereferencing will occur on the source side of the assignment statement automatically. More examples of dereferencing can be found in our next program example :

```
program 3 :    begin   int x,z ;   ref int p ;
                 x := 5; p := x ; x:=x+1 ; z := p
               end
```

The specifications in the first line explain the types of the variables x, z and p . ALGOL 68 uses the (deplorable) convention that the type of a variable always is one ***ref*** higher than indicated in the specification ; so x , being an ***int***egral variable is of type ***ref int*** , etc. .

The first two assignments require no dereferencing ; in the third assignment the variable x is dereferenced once, but in the fourth assignment the pointer p must be dereferenced twice. The value of p which equals x is not an integer ; we use the value of the value of p, which equals the value of x which equals 6 (since the value of x is increased by one by the third instruction).

Program 3 is a typical example for the type of problems encountered when we want to define the semantics for the assignment statement. The fact that the twofold dereferenced value of p is modified by the third assignment which does not involve p visibly, has escaped a correct treatment by previous semantics for the assignment statement.

4. THE HOARE PREDICATE TRANSFORMER FOR THE ASSIGNMENT STATEMENT.

Let us investigate the semantics for the assignment $x := t$. In the Hoare formalism we must provide a precondition Φ describing the set of states σ from which we may arrive in states σ' where a given postcondition Ψ holds by execution of the assignment $x := t$. Now the states σ and σ' are equal except for the fact that the value of x in state σ' equals the value of t in state σ. This has inspired Hoare and Floyd for proposing the following rules:

$[t/x]\Psi \ \{ x:=t \} \ \Psi$

or equivalently

$wp(x:=t,\Psi) = [t/x]\Psi$ (Hoare) [7] ,

and

$sp(x:=t,\Phi) = \exists z[[z/x]\Phi \ \& \ x = [z/x]t]$ (Floyd) [2] .

Here $[t/x]$ denotes the substitution operator which replaces all occurrences of the identifier x by the expression t ; the variable z which occurs in Floyd's rule is used to "remember" the old value of x before the assignment.

Some examples may illustrate the use of these rules :

$1 > 0 \ \{ x:=1 \} \ x > 0$ or

$wp(x:=1, x>0) = 1 > 0$, which reduces to *true* .

This example formalizes that always after execution of $x:=1$ it will be the case that $x > 0$.

$x = 4 \ \{ x := x+1 \} \ x = 5$ or

$wp(x:=x+1,x=5) = [x+1/x](x=5) = (x+1 = 5)$, which reduces to $x=4$.

$sp(x:=x+1,x=4) = \exists z[[z/x](x=4) \ \& \ x = [z/x](x+1)]$ which reduces to $\exists z[z=4 \ \& \ x = z+1]$, which reduces to $x = 5$.

It has been realised during the last five years that the above rules lead to incorrect results when applied to more complex situations. Again some examples :

wp(a[i]:= 13, a[i] = a[j]) = [13/a[i]](a[i] = a[j]) , which reduces to a[j] = 13 .

this is a correct precondition but it is not the weakest one ; i=j is a correct precondition as well ! An even more serious problem is encountered in the treatment of program 3. In the computation we use the rule for sequential composition explained in section 2. We like to compute

wp(x:=5 ; p:=x ; x:=x+1 ; z := p , z=6) ,

and according to our intuition the result should reduce to ***true*** . Applying Hoare's rule we obtain however :

[5/x][x/p][x+1/x][p/z] (z=6) , which reduces as follows :

[5/x][x/p][x+1/x] (p=6) = [5/x][x/p] (p=6) = [5/x] (x=6) = (5=6) =
= ***false***

This formal computation has yielded the absurdity that it is impossible that z=6 becomes true. The problem is caused by the substitution operator which does not replace the value of some identifier but this identifier itself. Grace to the fact that programmers have been trained to live with the ambiguity between identifiers and their values, these problems have not been discouvered before when Hoare's rule was originally formalised.

At this point we should indicate a link with Montague grammar. We claim that the above problem with assignments may be considered to be an instance of the problem with an opaque context. This may be illustrated by comparing the assignments :

x := y+7	x := x+7
y := y+7	y := x+7

If we have in the initial state that both x and y have the value 4 , then the assignments are row-wise equivalent, but column-wise distinct. In the context of the source identifiers with equal values may replace each other (referentially transparancy), whereas such replacement is strictly forbidden on the destination side (referentially opaqueness). Identifiers behave differently on both sides.

This observation is not new. Strachey [15] observed it back in 1964 , and the entire "name" or "reference" concept in the definition of

ALGOL 68 [17] is based upon this distinction. Pratt [12] even uses Quine's terminology for describing the situation. However, the solutions offered to the problems are far from ideal. Strachey introduces L- and R-values for identifiers where L-values represent abstract locations. The ALGOL 68 semantics introduces a highly abstract machine where names may refer to values or may be the value refered to by other names. Pratt evades the problem by restricting himself to first order references only - in this context the consequences of the opaqueness are not visible. According to Hoare's opinion the rule should be considered to be an axiom - if it is violated by some program then this program is outside the realm of programs under consideration.

In our proposal we attack the problem using the same tools which were introduced by Montague for dealing with opaqueness in natural language: Intensional logic. Variables will be considered to represent intensions of their values. Dereferencing becomes simply taking the extension of some intensional value, and pointers and other higher order references become intensions of intensions.

On base of this idea we have constructed a Montague grammar for a fragment of ALGOL 68 which contains the above assignments. Clearly Montague grammar introduces more than the use of Intensional logic; there is also the strict correspondence between syntactic and semantic rules. This idea is not revolutionary in the theory of semantics for programming languages. Most designers of programming languages have used some form of structural compositionality for both syntax and semantics, in order to make it possible that a computer program may understand and translate or interprete a program. Never the less, the use of this principle has inspired in our treatment an unconvential method for describing the assignment to subscripted variables like a[i] := y . We obtain a three-placed rule which differs from the rules in the ALGOL 68 report.

5. FEATURES FROM A MONTAGUE SEMANTICS FOR ASSIGNMENT STATEMENTS.

The programs we like to treat are simple sequences of assignments involving variables, array elements and pointers of arbitrary complexity. Program 1 and program 3 are examples. Note that this is only a very restricted class of programs, which, however, forms the building blocks of more complex programs. As such our theory is relevant for other areas of program semantics, where problems involving conditionals, iteration or recursion are discussed. Our treatment deals with problems which have been overlooked or taken for granted in this area of flow-of-control semantics. Related work can be found in de Bakker [1] or Gries [5] .

5.1. Grammar.

In Montague semantics sentences are represented by tree-structures which indicate the way the sentences are generated by the rules of the grammar. This is also common in programming language theory, with the additional restriction that the syntactical structure must be effectively and efficiently computable from the program text. The algorithm performing this analysis is called a *parser* ; its output is called the *parse tree*. This tree may become input to another program, called *code generator*, which generates a sequence of machine instructions which are capable to execute the program on a real computer. The combination of a parser and a code generator is called a *compiler*.

In Montague grammar the syntactic tree structures are input to a translation procedure which translates the sentence into a formula in IL. The same can be done with the parse tree of a program; its translation will be some meaningful expression in IL (more specifically an IL expression denoting a predicate transformer).

An essential feature of a Montague grammar are the Categories. Each phrase belongs to some particular category, which corresponds to the type of its translation in IL. Since the grammar contains only finitely many rules at most finitely many categories will be non-empty.

This typing mechanism is present in many of the grammars used to describe programming languages. The most complete typing can be found in the definition of ALGOL 68, where a *van Wijngaarden grammar* is used.

A van Wijngaarden grammar may be considered to be a categorical grammar dealing with infinitely many categories. The categories themselves are generated using a context-free *meta-grammar*. The rules of the grammar actually are rule-schemes which only become rules by replacing occurrences of *meta-variables* by some categories which are terminal productions of these meta-variables in the meta-grammar.

We illustrate this mechanism by way of an example. Our meta-grammar is given by :

MODE : ***int*** ; ***ref*** MODE ; [ ] MODE . (A)

This is a regular meta-grammar involving the single meta-variable MODE ; terminal productions are : ***int*** , ***ref int*** ,[] ***int*** , ***ref*** [] ***int*** , etc. .

Next consider a rule scheme like :

<***ref*** MODE assignment> :< ***ref*** MODE destination> := <MODE source> . (a) .

A rule is obtained by replacing in (a) the three occurrences of MODE by a production of MODE in (A) . This way we obtain among others the rule:

< ***ref int*** assignment >:<***ref int*** destination> :=<***int*** source> .

The same production of MODE must be used for each of the three occurrences, so the following utterance is not a rule :

< ***ref int*** assignment> : <***ref*** [] ***int*** destination > :=<***ref ref int*** source> .

Note that (a) expresses the structure of the assignment statement to integer variables or pointers as explained in section 3 : if the destination is of type ***ref*** Σ then the source must be of type Σ .

In our proposal we have used a van Wijngaarden grammar ; the above rules are taken from our proposal. This grammar can thus be considered as a grammar with infinitely many rules and an infinite number of non-empty categories.

5.2. Interpretation - types .

An interpretation of intensional logic requires a set S of possible worlds. In our proposal these worlds must represent computer states. Since these states can actually be constructed no ontological problems arise.

The modes in our grammar correspond to the type in IL. In fact there exist types in IL which do not correspond to a mode in the grammar, but are used for the interpretation of complex expressions. Below we present a list of types with their corrsponding modes, values and IL-expressions.

type	MODE	values	IL-expressions
t	----	{*true* , *false* }	*true* , *false* , 0=1 , ˇx=5, ...
e	*int*	ℤ = {...,−1,0,1,..}	1, 3+4, ˇx, ˇˇp, ...
s	----	S (states)	s_0, s, ...
<s,e>	*ref int*	integer variables	x, y, z, ˇp, a[5], ...
<s,<s,e>>	*ref ref int*	integer pointers	p
<e,e>	[ ]*int*	integer functions	ˇa, λn[*if* n=7 *then* 5 *else* n *fi*]
<s,<e,e>>	*ref* [ ]*int*	array variables	a, b, q[5]
<s,<e,<e,e>>>	*ref* [ ][ ]*int*	two dim array var's	q
<s,t>	----	predicates	^(ˇx=5)
<<s,t>,<s,t>>	----	predicate transformers	λ*P*[^(ˇ*P* *and* (ˇx=5))]

The last two lines in the list will be discussed in section 5.4. Note that we have extended IL by introduction of the *if then else fi*-construct.

5.3. Interpretation - states and state-switchers.

For each of the types there exists a set of objects of this type. Objects of the type <s,e> will be functions from the set S to the set ℤ. We should however not be naive and take for the set of objects of type <s,e> the full set of all functions from S to ℤ. Instead we need a general model in the sense of Henkin (see [6] & [4]). This restriction is motivated by the following requirement :

If we like that our worlds describe the states of a computer it is reasonable to require that a variable of type *ref* Σ may obtain each legitimate value of type Σ as value in at least one state. Now take Σ equal to the type *ref int* . Then the above requirement implies that for each legitimate function f from S to ℤ there exists a state s ∈ S such that the value of p in state s (i.e. the interpretation of ˇp with respect to world s) equals f . This would imply that there exists a surjective mapping from the set of states S to the set of legitimate functions so, due to cardinality reasons, this can not be the full set of functions.

The general model which we have constructed is described in [9] . It has the following property : Let χ denote the interpretation of some ***ref*** Σ identifier in the grammar, and let β denote an acceptable value of type Σ (i.e. a value which is required to exist within our model). Then there exist for each state s an unique state denoted $(\chi \leftarrow \beta)s$, such that the value of χ in $(\chi \leftarrow \beta)s$ equals β , whereas the value of all other ***ref*** Σ identifiers are equal in the states s and $(\chi \leftarrow \beta)s$. As before we mean by "the value of χ in state s" the interpretation of ${}^{\vee}\chi$ with respect to world s .

This property expresses that we can freely assign a value to a destination without leaving our universe or destroying the values of other identifiers.

The state transitions $(\chi \leftarrow \beta)$ are made expressible in IL by the use of a corresponding set of modal operators called *state-switchers* . The state-switcher corresponding to the transition $(\chi \leftarrow \beta)$ is denoted $\{\beta/{}^{\vee}\chi\}$. Its meaning is defined semantically by :

"the interpretation of $\{\beta/{}^{\vee}\chi\}\delta$ in state s equals the interpretation of δ in state $(\chi \leftarrow \beta)s$ " .

The state switchers resemble a substitution operator but they behave slightly different. We have the following reduction rules:

$\{\beta/{}^{\vee}x\}\ x = x$

$\{\beta/{}^{\vee}x\}\ {}^{\vee}x = \beta$

$\{\beta/{}^{\vee}x\}\ {}^{\vee}y = {}^{\vee}y$ if $y \neq x$

$\{\beta/{}^{\vee}x\}\ (t+u) = \{\beta/{}^{\vee}x\}\ t + \{\beta/{}^{\vee}x\}\ u$

$\{\beta/{}^{\vee}x\}\ \{\beta'/{}^{\vee}x\}{}^{\vee}x = \{\beta/{}^{\vee}x\}\ \beta'$

5.4. Predicate and predicate transformers.

The goal of our semantics is to obtain a predicate transformer which expresses the behaviour of a program. The predicates on which this transformer should operate clearly are functions from states to truth-values, i.e. they are objects of the type $\langle s,t \rangle$. Since all arithmetic is performed at the level of objects of the types t and e we use the intension operator for describing the predicates in IL .

For example , 5 is an integer and so is $^{\vee}x$ (note the use of the extension operator !) . Hence $^{\vee}x = 5$ is a truth value ; it is *true* iff the value of x in the present state equals 5 . Now the predicate "the value of x equals 5 " is denoted $^{\wedge}(^{\vee}x = 5)$.

A predicate transformer is of the type <<s,t>,<s,t>> . In IL these transformers can be described using λ-expressions where P denotes a variable of the type <s,t> . We can use these expressions to denote the meaning of programs . We illustrate this by some example.

Consider the predicate transformer :

$\eta = \lambda P[\ ^{\wedge}(\{5/^{\vee}x\}\ ^{\vee}P\ )\ ]$.

We compute the result of this transformer on the predicate $^{\wedge}(^{\vee}x{=}5)$.

$\lambda P[^{\wedge}(\{5/^{\vee}x\}^{\vee}P)](^{\wedge}(^{\vee}x{=}5)) = {}^{\wedge}(\{5/^{\vee}x\}^{\vee\wedge}(^{\vee}x{=}5)) = {}^{\wedge}(\{5/^{\vee}x\}(^{\vee}x{=}5)) =$

$= {}^{\wedge}(5{=}5) = {}^{\wedge}\,true$.

This simple computation shows that the transformer η transforms the predicate "the value of x equals 5 " into the predicate "true" . Note that by Hoare's rule we obtain :

$wp(\ x{:=}5, x{=}5\ ) = [5/x](x{=}5) = (5{=}5) = true$.

So for this particular combination of an assignment and a postcondition the transformer η has yielded the correct weakest precondition.

6. THE REVISED FLOYD AND HOARE RULES FOR THE ASSIGNMENT STATEMENT.

Suppose that we execute x := t in some state s then the resulting state should become (x ← t)s . So if we want that some predicate Ψ becomes true after execution of x := t it should be the case that the predicate Ψ should hold in state (x ← t)s . According to our definition of the state-switcher this is equivalent to requiring that the predicate $\{t/^{\vee}x\}\Psi$ holds in state s . If we remember that in our framework most predicates are of the form $\Psi = {}^{\wedge}\psi$ we obtain our revised Hoare rule :

$wp(\ x{:=}t, \Psi\ ) = {}^{\wedge}(\ \{t/^{\vee}x\}\psi\ )$ (Hoare')

The revised Floyd rule becomes :

$sp(\ x{:=}t, \Psi\ ) = {}^\wedge(\ \exists z[\ \{z/{}^\vee x\}\psi\ \wedge {}^\vee x = \{z/{}^\vee x\}t\ \ ]\ )$ (Floyd')

We may express the above by stating that the meaning of the assignment x := t is given in IL by the (backward) predicate transformer :

$\lambda P[\ {}^\wedge(\ \{t/{}^\vee x\}{}^\vee P\ )\ ]$

or the (forward) predicate transformer :

$\lambda P[\ {}^\wedge(\ \exists z[\ \{z/{}^\vee x\}{}^\vee P\ \ \wedge\ {}^\vee x = \{z/{}^\vee x\}t\ \ ]\ )\ ]$

By way of example we treat program 3. As postcondition we take ${}^\wedge(\ {}^\vee z{=}6 \wedge {}^\vee p{=}x\ )$. By applying rule (Hoare') we obtain the following calculation:

$wp(\ \text{program 3},\ {}^\wedge(\ {}^\vee z{=}6 \wedge {}^\vee p{=}x\ )\ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\{x/{}^\vee p\}\{{}^\vee x{+}1/{}^\vee x\}\{{}^{\vee\vee}p/{}^\vee z\}\ ({}^\vee z{=}6 \wedge {}^\vee p{=}x\ )\ \ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\{x/{}^\vee p\}\{{}^\vee x{+}1/{}^\vee x\}\ ({}^{\vee\vee}p{=}6 \wedge {}^\vee p{=}x\ )\ \ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\{x/{}^\vee p\}\{{}^\vee x{+}1/{}^\vee x\}\ (\ {}^\vee x{=}6 \wedge {}^\vee p{=}x\ )\ \ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\{x/{}^\vee p\}\ (\ {}^\vee x{+}1{=}6 \wedge {}^\vee p{=}x\ )\ \ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\{x/{}^\vee p\}\ (\ {}^\vee x{=}5 \wedge {}^\vee p{=}x\ )\ \ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\ (\ {}^\vee x{=}5\ \ \wedge\ x{=}x\ )\ \ ) =$

$\quad {}^\wedge(\ \{5/{}^\vee x\}\ (\ {}^\vee x{=}5\ )\ \ ) =$

$\quad {}^\wedge(\ 5{=}5\ ) =$

$\quad {}^\wedge\ \textit{\textbf{true}}$.

So this computation yields the correct result.

The correctness of the rule (Floyd') was proved in [9]; in [10] we have shown that the correctness of (Floyd') and (Hoare') are equivalent so (Hoare') is correct as well.

Next we consider the assignment to array elements, e.g. a[i]:= y . In this assignment both y and i may be replaced by expressions with the same value, but for a such a replacement is forbidden. So the phrase a[i] := y is referentially transparant in i and y, but referentially opaque in a . This indicates that each of the three sub-expressions a , i and y has an independent contribution to the meaning of the assignment. Due to the principle of correspondence between syntax and semantics we have decided to use a three-placed rule for describing these assignments.

The value of a is the intension of a function from $\mathbb{Z}$ into $\mathbb{Z}$. The semantics of assigning a value to an array element can easily be described by assigning a new function to the array-identifier as follows :
"the meaning of the assignment a[i] := y is equivalent to the meaning of the assignment a := λn[*if* n = i *then* y *else* a[n] *fi*] ".

This type of rewriting is repeated until the destination becomes a simple identifier.

By way of illustration we present the example:

$$wp(a[{}^{\vee}i] := 6,\ {}^{\wedge}({}^{\vee}a[{}^{\vee}i] = {}^{\vee}a[{}^{\vee}j])) =$$
$$wp(a := \lambda n[\textbf{\textit{if}}\ n={}^{\vee}i\ \textbf{\textit{then}}\ 6\ \textbf{\textit{else}}\ {}^{\vee}a[n]\ \textbf{\textit{fi}}],\ {}^{\wedge}({}^{\vee}a[{}^{\vee}i] = {}^{\vee}a[{}^{\vee}j])) =$$
$${}^{\wedge}(\{\lambda n[\textbf{\textit{if}}\ n={}^{\vee}i\ \textbf{\textit{then}}\ 6\ \textbf{\textit{else}}\ {}^{\vee}a[n]\ \textbf{\textit{fi}} / {}^{\vee}a\}\ ({}^{\vee}a[{}^{\vee}i] = {}^{\vee}a[{}^{\vee}j])) =$$
$${}^{\wedge}(\lambda n[\textbf{\textit{if}}\ n={}^{\vee}i\ \textbf{\textit{then}}\ 6\ \textbf{\textit{else}}\ {}^{\vee}a[n]\textbf{\textit{fi}}]\cdot[{}^{\vee}i] = \lambda n[\textbf{\textit{if}}\ n={}^{\vee}i\ \textbf{\textit{then}}\ 6\ \textbf{\textit{else}}\ {}^{\vee}a[n]\textbf{\textit{fi}}]\cdot[{}^{\vee}j] =$$
$${}^{\wedge}(\textbf{\textit{if}}\ {}^{\vee}j={}^{\vee}i\ \textbf{\textit{then}}\ 6=6\ \textbf{\textit{else}}\ 6={}^{\vee}a[{}^{\vee}j]\ \textbf{\textit{fi}}) =$$
$${}^{\wedge}({}^{\vee}j = {}^{\vee}i\ \vee\ {}^{\vee}a[{}^{\vee}j] = 6)$$

This is again the correct result.

The solution of considering an assignment to the array element as being an assignment to the array itself is not new. It can be found in Gries [5] , Hoare [8], or Pratt [12]. The solution given in de Bakker [1] may be derived from our solution, but was not obtained by use of assignments to the array identifier. Moreover, the use of intension and extension operators seems unnecessary as long as no pointers are involved.

Examples of use of the modified Floyd rule can be found in [9] .

7. APPLICATIONS TO NATURAL LANGUAGE

In this section we present some tentative suggestions indicating how the tools mentioned in the previous sections might be useful for natural languages as well.

7.1. Van Wijngaarden grammar

It has been demonstrated that the concept of a van Wijngaarden grammar is a very powerful tool (e.g. in the formal definition of ALGOL 68 [17]). On the other hand it is a very manageable kind of grammar, due to the principle of splitting up the rules in two levels: the meta rules and the rule schemes. A -restricted- use of such grammars seems fruitful for natural languages as well.

There are syntactic arguments for distinguishing among the nouns several subsets, e.g. mass nouns and count nouns. If these subsets are considered as independent categories, we would lose the information that they behave in much respects the same: they can all be combined with a determiner in order to form a term, they all may have a relative clause and in each we may substitue a term for occurrences of he_n. If we use the idea's of a van Wijngaarden grammar, the differences and relations between these subsets can be expressed in a short and readable way.

We might use the meta-rules

NOUN --> Mass Noun

NOUN --> Count Noun

and the rule schemes

S3: if $\alpha \in P_{NOUN}$ and $\beta \in P_t$ then $F_{3,n}(\alpha,\beta) \in P_{NOUN}$

S4: if $\alpha \in P_{NOUN}$ and $\zeta \in P_T$ then $F_{10,n}(\alpha,\zeta) \in P_{NOUN}$

The definition of a van Wijngaarden grammar requires that all occurrences of NOUN in e.g. S3 are replaced by the same terminal production of NOUN. So S3 expresses that a Mass noun with a relative clause still is a

Mass Noun and that a Count Noun with a relative clause is a Count Noun again. The same holds for rule S15.

The power of a van Wijngaarden grammar can also be used to deal with a restriction on the use of term substitution. Rule S14,n needs as input a term phrase and a sentence. The term has to be substituted for the first occurrence of he_n. A reasonable requirement is that the sentence indeed does contain an occurrence of he_n. This cannot be expressed in the rules of PTQ. Using a van Wijngaarden grammar with structured category symbols which contain information about the he_n's that have to occur, one can formulate the rules in such a way that it is guaranteed that the restriction is satisfied. In such a situation the rule scheme represents an infinite number of rules since there is no last he_n.

Notice that in fact PTQ [11] contains (very limited) rule schemes. A rule like S14,n is rule scheme with meta notion n , the terminal productions of this metanotions are 1,2,3,.. .

7.2 State Switchers

In PTQ [11] there are defined the modal operators for necessity, possibility, future and past. The state switchers provide a large extension of this class of operators. An example of a sentence which is difficult, if not impossible, to treat using only these four modal operators is the following (due to Saarinen [14])

Every man who ever supported the Vietnam war, now believes that he was an idiot then.

A translation like

$$\forall x\ [\mathit{man}(x) \wedge \mathbb{H}\ (\mathit{support}(x, \mathit{Vietnam\text{-}war}) \rightarrow \mathit{believe}(x, {}^\wedge \mathbb{H}(\mathit{idiot}(x))]$$

is incomplete in sofar that it does not express that the same moments in the past are involved. If we can use a state-switcher for modifying the time-coördinate, we may enforce coherence:

$$\forall x\ \forall t_0\ [\mathit{man}(x) \wedge \{t_0/t\}\ (\mathit{support}(x, \mathit{Vietnam\text{-}war}) \rightarrow \mathit{believe}(x, {}^\wedge(\{t_0/t\}\mathit{idiot}(x))]$$

7.3. Predicate Transformers

The basic expression in a programming language is the assignment. For the computer an assignment is a command to perform a certain action. We have demonstrated how the semantics of such commands is dealt with by means of predicate transformers. Inspired by this approach we might do the same for commands in natural language. Consider for instance the imperative

John, drink tea.

Its translation as a predicate transformer would becomes something like

$\lambda P\,[\,{}^{\wedge}(B({}^{\vee}P) \wedge \; \mathit{drink\text{-}tea}({}^{\wedge}j)]$

This expression describes the change of the state of the world if the command is obeyed. The operator B is a kind of state-switcher, it indicates the moment of utterance of the command.

A similar approach can be used to describe the semantics of actions. One might describe the semantics of performative sentences like

We crown Charles emperor.

by means of an predicate transformer.

Often a sequence of sentences is used to perform an action rather than to make some assertions: sentences can be used to give information to the hearer. Consider the text

Mary seeks John. John is a unicorn.

These sentences might be translated into the predicate transformers

$\lambda P[P \wedge \mathit{seek}_*(m,j)]$ and $\lambda P[P \wedge \mathit{unicorn}_*(j)]$

Suppose that the information the hearer has in the beginning is denoted by ϕ. Then by the first sentence this information is changed into

$\phi \wedge \mathit{seek}_*(m,j)$

and by the second sentence into

$\phi \wedge \mathit{seek}_*(m,j) \wedge \mathit{unicorn}_*(j)$

From the final expression the hearer may conclude that Mary seeks a unicorn.

8. REFERENCES

1 de Bakker, J.W., *Correctness proofs for assignment statements*, Rep. IW 55/76, Mathematical Centre 1976, Amsterdam.

2 Floyd, R.W., *Assigning meanings to programs*, in J.T. Schwarts (ed.) Proc. Symp. in Appl. Math. 19, *Mathematical aspects of computer sciences*, Am. Math. Soc. 1967, pp. 19-32.

3 Frege, G., *Ueber Sinn und Bedeutung*, Zeitschrift für Philosophie und Philosophische Kritik 100, 1892, pp. 25-50. Transl.: *On sense and reference*, in P.Geach & M.Black (eds.) *Translations from the philosophical writings of Gottlob Frege*, Basil Blackwell 1952, Oxford.

4 Gallin, D., *Intensional and higher order modal logic*, North. Holl. Publ. Comp. 1975, Amsterdam.

5 Gries D., *Assignment to subscripted variables*, Rep. TR 77-305, Dept. Comp. Sci. Cornell univ. 1977, Ithaca NY.

6 Henkin L., *Completeness in the theory of types*, J. Symb. Logic 15 1950, pp. 81-91.

7 Hoare, C.A.R., *An axiomatic base for computer programming*, Comm. ACM 12 1969, pp. 576-580.

8 Hoare, C.A.R. & N. Wirth, *An axiomatic definition of the programming language PASCAL*, Acta Inf. 2 1973, pp. 335-355.

9 Janssen, T.M.V. & P. van Emde Boas, *On the proper treatment of referencing, dereferencing and assignment*, in A. Salomaa & M. Steinby (eds.) Proc. ICALP 4 (Turku 1977), Springer Lecture Notes in Comp. Sc. 52, pp 282-300.

10 Janssen, T.M.V. & P. van Emde Boas, *The expressive power of intensional logic in the semantics of programming languages*, in J.Gruska (ed.) Proc. MFCS 6 (Tatranska Lomnica 1977), Springer Lecture Notes in Comp. Sc. 53, pp. 303-311.

11 Montague, R., *The proper treatment of quantification of ordinary English*, in J.Hintikka, J.Moravcsik & P.Suppes (eds.), *Approaches to natural language*, Reidel 1973, Dordrecht. Repr.: in R.H.Thomason (ed.) *Formal philosophy. Selected papers of Richard Montague*, Yale university press 1974, New Haven & London.

12 Pratt, V.R., *Semantical considerations on Floyd-Hoare logic*, Proc. IEEE FOCS 17 1976, pp. 109-121.

13 Quine, W.V.O. *Word and Object*, the MIT press 1960, Cambridge Mass.

14 Saarinen E., *Backwards-looking operators in tense logic and in natural language*, in J.Hintikka, I.Niiniluoto & E.Saarinen (eds.) *Essays on mathematical and philosophical logic*, Reidel 1978 Dordrecht.

15 Strachey, C., *Towards a formal semantics* in T.B.Steel (ed.) *Formal language description languages for computer programming*, North. Holl. Publ. Comp. 1966 Amsterdam, pp.198-220.

16 Stoy, J.E., *Denotational Semantics: The Scott-Strachey approach to programming language theory*, the MIT press 1977, Cambridge Mass.

17 van Wijngaarden, et al. (eds.), *Revised report on the algorithmic language ALGOL 68*, Tract MCT 50, Mathematical Centre 1976,

THE USE OF CATEGORIAL SYNTAX IN MONTAGUE GRAMMAR

Peter Eisenberg

This paper examines some aspects of the way, categorial syntax is used in Montague grammar. It tries to answer the question whether it is justified, suggestive or even necessary to take categorial syntax as the basis for grammatical description within this framework.

Our problem must be seen within the following context. One of the main issues in recent work on syntax has been the relatedness between properties of grammars and properties of language. This problem can be approached in many different ways. One way of putting it was the well known attempt of the generative school to find out which grammar - or class of grammars - is formally most restricted and still sufficiently rich for the description of natural language. For reasons which have been given elsewhere we consider this kind of question as misleading and not to be answered.[1]

Much less ambitious but indispensable is the requirement for the grammarian to know, whether the formal properties of a grammar - understood as the properties of a metalanguage - are such that the structural properties of a given object language can be adequately described. With respect to the use of categorial syntax in Montague grammar there are a number of obstacles to the answer of this question which force us to rely partly on speculations and suppositions. First of all, the fragments which are dealt with in Montague grammar are at present still relatively small. Therefore it is not always easy to imagine how certain phenomena would be handled within a certain fragment, what the result would be if certain fragments would be put together etc. So it might very well be the case that some of the suppositions which will be made in what follows do not lead to the best solution possible. I am quite sure that some of my criticisms could be avoided by simply giving a different grammatical solution for the phenomena in question.

Another difficulty is caused by the fact that if one wants to say something about the relation between properties of grammar and properties of language one has to know exactly what these properties are. The interest of this paper is mainly a syntactic one, i.e. we are interested in the compatibility of the structural properties of the object language with categorial syntax. In Montague grammar this syntax is not used for the description of the surface of natural language in a strict sense: in most cases our object language is a language somewhere between the surface and a semantic correlate language and has to be related to the surface by a transformational component or the like. So it is not quite clear to what extent we are talking about syntactic properties of natural language at all when we talk about Montagues desambiguated language, about the language which Bartsch describes on what she calls the categorial level etc., or whether we are talking at least in part about semantic properties of natural language. For this reason we will be cautious and mostly employ the term 'grammatical' to avoid a committment with respect to syntax and semantics. The difficulties with the language are augmented by those with the grammar. Categorial grammar is by no means the same thing everywhere in Montague grammar. On the other hand, categorial grammar is now widely used within this framework. For some people it is even part of the characteristics of modeltheoretic semantics to use a categorial syntax.

Despite these difficulties it seems possible to put forward some properties of categorial grammar (henceforth CG) as well as of the object language to be described which are common to most approaches and to relate these to each other. By doing this we will come to the conclusion that CG is not compatible with the properties of these languages, or that it is at leas not justified to use this highly restricted type of grammar.

To begin with an example, I would like to question the assumption that it really is an advantage, as some authors seem to imply,[2] if the potential categories of the grammar are recursively enumerable, provided the basic categories are given. This property of categorial grammars seems of questionable value, because up to now the recursive properties of natural language as well as of those artificial languages used for semantic description have been explicated by relating the infinite set of object language expressions to structural relations between elements of the finite metalanguage vocabulary. CG certainly does not claim that an infinite stock of categories is needed for the description of any language, but nevertheless it makes available such an inventory. By doing this it poses the question in which way and for what reasons a language does choose its categories from the infinite stock. This question is of great interest, but it can only be put as a sensible and concrete one if one knows something about the parameters by which the infinite set of categories is to be described. Infinitness is not a merit as such. Therefore one is again confronted with the question of whether the specific formalism of CG is at all adequate to capture the grammatical properties of natural language.

In what follows we will discuss this question by defining a CG as having the following properties: (a) the categories of the grammar are recursively enumerable, (b) the set of categories devides into the disjunctive subsets of basic categories and derived categories, and (c) all derived categories have to be understood as being functors or operators on other categories.

I would now like to make some remarks on the relation between CG and context free phrase structure grammar (PSG), since I believe that the extensive use made of CG at the moment can historically be explained by considering the relationship between these two types of grammar.

In his work on CG Bar-Hillel discussed in some detail the question of why this kind of grammar is not adequate for the description of natural language surface structure. Besides the repeatedly mentioned difficulty to handle discontinous constituents Bar-Hillel (1960a,81) hints at the problem that a more detailed categorization of nouns (sg. - pl., animate - inanimate) or of sentences (question, assertion, imperative) would enhance the number of basic categories without giving us the possibility to relate these categories to each other in an adequate way. Further notational complications which have nothing to do with the basic idea of CG would be required to formulate the rules for grammatical determiners, subject-verb agreement, and similar phenomena. Bar-Hillel concludes that it is "highly doubtful whether English grammar ... can at all be forced into the strait-jackett of the immediate constituent model and remain workable and revealing."(ibid.).

This formulation shows that Bar-Hillel is not concerned with the specific properties of CG, but with PSG in gerneral. He extensively discusses the relationship between the two types of grammar and concludes that CG is inadequate for the description of natural language because PSG is. Since Chomsky (1957) had shown PSGs to be inadequate, it was sufficient for him to show that CG and PSG are (weakly) equivalent to question the adequacy of CG (cf. Bar-Hillel 1960b, 103 and 114). This equivalence was then often an important argument when the question of adequacy was discussed. The specific properties of categorial grammar as compared to PSG have not been considered to be of great interest. So it is not surprising that Lewis (1972,173) remarks one should be concerned with categorially based transformational grammars because "we have become interested in the plan of using a simple phrase structure grammar as a base for transformational grammar". And for Lewis, a PSG is of course nothing but a notational variant a CG (ibid).

In the most recent discussion concerning the appropriateness of CG as surface grammar Bartsch et al. (1977) also draw our attention nearly exclusively to those difficulties which also hold as well for PSGs. Besides Bar-Hillel's criteria they mention the following ones. (a) "die morophophonemischen Veränderungen an Wörtern in ihrem syntaktischen Gebrauch" (216) as we find for instance as Ablaut in the forms of the past of German strong verbs (singe - sang). There are no specific difficulties for categorial grammar here, i.e. no difficulties which would not arise with PSG. (b) the fact that certain constituents can have multiple syntactic functions (215f.). In sentences like Maria ist schöner in Hosen als im Kleid, in Hosen has the same syntactic relation with respect to Maria as has im Kleid. Similarly, tanzt has a double syntactic function in Hans tanzt und Fritz auch. Here we also don't see special problems for CG. And if one, as Bartsch et al. (216) does, calls such expressions not well - formed with respect to CG "weil sich aus den Bedeutungen der Teile nicht die Bedeutung des Gesamtausdrucks ergibt", then this can of course be generalized. There is no grammar at all which could assign the meaning of Fritz tanzt auch to an isolated constituent Fritz auch. According to the logic of Bartsch et al. such expressions would not only be illformed with respect to CG but would not be wellformed at all. (c) the assignment of expressions to different categories, "die in ihrer Funktion gleich sind" (214), for instance the categorization of the adjective in weiße Rose as n/n and in rose blanche as n\n. Here we are in fact faced with a difference to PSGs. While in PSGs the distributional behavior of syntactic units may be a criterion for their categorization, this is necessarily the case for operator categories in CG. Yet it is questionable whether one can call CG "nicht adäquat für natürliche Sprachen" (215) because German and French have different syntax.

I hope that was has been said so far is sufficient to illustrate that the specific properties of CG have not played a central role when its adequacy was discussed. In most cases it was rejected because of its close relationship to PSG.

As long as one talks about syntactic surface structures it does not seem to cause much trouble if CG is just understood as a special type of PSG. But in my opinion this identification has had the unfortunate consequence that some important properties of CG have not been seen any more. In other words, it was assumed that CG would be adequate for the description of natural language if the problems could be solved which prevent its application as surface grammar. As a cosequence of this, we will have to show that this is not the case. In my opinion CGs are even inadequate for the type of language they are now used for.

Let us now first have a look at the relationship between the category of an expression and its distribution. Generalized linguistic statements, concerning classes of expressions or whole languages, can not be made without the use of grammatical categories, no matter what is considered to be the ontological status of these categories. The possibility to make generalized statements relies on the assumption that grammatical categories are sets or indices of sets of expressions which share certain properties. If - as it is explicitly or implicitly the case in most linguistic approaches - distribution is the most important criterion for categorizing expressions, then the simplest common property of a grammatical class of expressions is to occur or not to occur in a certain linguistic context. Of course this simplest case is not what linguists are looking for. Instead, expressions are classified with respect to all or no common contexts. Only by trying to do this can one hope to come close to the most generalized statements.

This is one of the points which can only be partly taken into account if we use a CG. The reason is that within CG every category has to be used either as an operator or as an argument. With respect to the use of categories as arguments no special problems show up. Arguments are usable more or less like 'normal' grammatical categories in that they can occur in every possible context. Those categories which occur exclusively as arguments can even be established by classical distributional methods, since for every possible context one can think of a suitable derived category. Derived categories, on the other hand, are bound to the occurrencein one single context as long as they are used as operators. Therefore, we are forced to assign to a class of functional or relational expressions a different category for each single context, even if these expressions share all of their contexts. So it is impossible to account for the familiar property of natural language and many artificial languages that expressions of this kind take formally (i.e. syntactically) different classes of expressions as arguments. Let us give some examples for what is ment by this.

In Montague 1970 (250) conceive and be are taken as transitive verbs (tv). Both of these verbs illustrate the difficulty of CG we are talking about. Besides transitive verbs Montague (ibid.) uses a verb class which maps sentences into intransitive verbs (iv/t) like believe that, assert that etc. Without doubt conceive that belongs also to this class, since we have John conceives that the plan is difficult as well as John conceives the difficulty of the plan. Montague has avoided the treatment of such cases, which really are nothing special in natural language. In his fragments a verb always belongs to the 'that-verbs' or to the transitive verbs. From a linguisitc point of view it is most important for the categorization of conceive that this verb takes nominal complements as well as certain sentential complements. It is this property which makes conceive syntactically different from tv-verbs and from iv/t-verbs, and of course there are semantic differences which correspond to the syntactic ones.

Most grammars express facts of this kind on the level of syntax by taking verbs like conceive as a class for themselves and giving them their own syntactic category. As opposed to this, in CG this verb has to be considered as belonging to several categories. Of course it is possible in every single case to bring up an argument as to whether a unit should be assigned to several categories or whether one should establish a new category. At the moment we are not interesting in claiming that we have found the right solution of this problem, not even in any single case. What we want to show is that CG brings discussions of this kind to an end simply because the format of its categories does not allow us to formulate things in different ways. The structure of the grammatical format prevents us from raising certain grammatical questions.

Consider now the difficulties which we run into when we want to handle the most simple grammatical phenomena involving the verb be. Montague's treatment of be as transitive verb is based on sentences with nouns in predicative position. The question then arises, to which category be should be assigned in copula sentences with predicative adjectives. If one is not to claim that is in John is teacher is the form of a completely different verb from is in John is clever, then one has to treat be uniformly as a transitive verb and to assign predicative adjectives to the same category as predicative nouns. But the attributive adjectives, which are normally considered to be of category n/n, would have to be categorized differently from predicative adjectives. Categorial grammar makes the best of this dilemma when it states that "schön in drei Kategorien, nämlich als Adnominal, als Adverbial un als Prädikatsadjektiv gebraucht wird." (Bartsch et al. 1977, 214). If one follows our argument, then adjectives like schön belong to several categories simply because sein should not belong to several categories. This is again a consequence of the fact that in CG operator categories cannot operate on different argument categories.

In most Montague grammars adjectives belong to those classes of expressions which can equally well be used as operators (attributive) and as arguments (predicative). Verbs also belong to this class, since they mostly work as operators when they bind their subject and objects, and as arguments when they are modified by adverbs. But while in the grammar of Bartsch et al. (1977, 212) as in many other CGs verbs do not change their category when they switch from operators to arguments, adjectives do change their category. What could be the reason for this difference? In my opinion, the claim that adjectives belong to several categories rests on a confusion of the concept of category and the concept of function, in other words, attributive and predicative adjectives belong to one single category (namely the category of adjectives), but they can have different grammatical function.

With respect to most types of grammar, much has been said about the danger of confusing the concepts of category and function. For CG of the kind which is presently used within the framework of Montague grammar almost nothing has been said about this problem, though there are special reasons here for mixing up the two without becoming aware of it.

In CG category and function coincide for those units which are used as operators. If an adjective is of category n/n and is used as operator, then this means that its function as attribute is given with its category. By this we don't want to say that this method of characterizing grammatical functions or relations will be generally sufficient for natural language. This method is restricted to the use of dominating and coordinated categories, since only these categories can show up in a derived category, and it is doubtful whether this will be enough for all cases.[3] Yet as far as this method works, category and function indeed coincide for the units in question, i.e. the function of a unit is given by its category alone. In a language which is described by a CG such an unit has, when used as operator, one and only one function.

The situation is different with respect to units which are not used as operators, no matter whether they belong to basic or derived categories. Nothing can be said from the category alone about the function of the unit. According to the grammar of Lewis 1972 (172) we would categorize grunts as t/n and piggishly as (t/n)/(t/n) when they are used as in Petunia grunts piggishly. Here the category of the verb does not tell us anything about the verb's grammatical function but only about a possible function as predicate. With respect to units which belong to basic categories, not even this will be the case. Their categories do not tell us anything about their function, as is generally the case with the other types of grammar. The same holds for units which belong to derived categories but whose category is never used as an operator category, like intransitive verbs (t/e) and common nouns (t//e) in Montague 1970 or infinitive phrases (t///e) in Thomason 1976 (78).

In CG the relationship between category and function is more complicated than in other types of grammar because there is the special difficulty that in certain cases categorial and functional decisions coincide. This restricts the flexibility of our metalanguage to some extent. It also means that there cannot be a unified method of establishing grammatical gategoeries. Whereas normal grammatical procedures can be applied in most cases for establishing argument categories, this does not hold for operators.

The partial identification of category and function has the further consequence that the grammarian is faced with the necessity to regard one certain occurrence of a unit as basic, i.e. responsible for the categorization of that unit, and all the other occurrences as secondary. The distinction between a 'primary', an 'underlying', or 'basic' use and 'derived', 'secondary', or 'nonbasic' uses of a unit is to a certain extent responsible for the fact that the empirical foundation of many grammars, especially those of the generative persuasion, has become more and more problematic during the past ten years or so. For this reason alone, one should try to avoid such a distinction as long as one does not have a satisfactory interpretation for concepts like 'primary' and 'secondary' use. The constraint that one has to decide between primary and secondary uses, which has its reason in the special format of the derived categories, often makes the distinction an arbitrary one. Moreover, it leads to technical difficulties in some cases.

As an example consider intransitive verbs as compared with nominals. For the verbs it seems to be quite natural to regard their use as predicates as primary and their use as arguments to adverbs as secondary, though it is not quite easy to give the precise reasons for this intuition. But with respect to nominals we cannot even rely on our intuition. Take the nominals which Montague (1970) calls terms. Their categorization as t/(t/e) tells us that their use as subject is regarded as determining the category. Now Cresswell (1973, 81) points out that this categorization of terms (or that as t/(t/n), since the special role, which Montague assigns to the category e, is of no importance here), does not allow for the grammar to be formulated in a natural way as long as one wants to have what Cresswell calls a pure categorial language as object language and CG as language of grammatical description. The reason for the difficulties is, according to Cresswell, that terms have to belong to the same grammatical category, regardless of whether they are used in subject position or in object position. But when categorized as t/(t/n), terms can not combine with transitive verbs (which are of category t/n,n in Cresswell's grammar). Therefore Cresswell proposes the introduction of variables, i.e. he changes the object language and makes an important step towards λ-categorial languages. Terms can now be uniformly assigned to category t/(t/n). For this one has to pay the price that the object language has diverged quite a lot from natural language. Grammar has had another victory over language.

Quite different consequences follow from a uniform categorization of terms in Montague 1970. While Cresswell uses an object language in which terms occur always as operators of the category t/(t/n), regardless of whether they are used as subject or as object in the corresponding natural language sentence, Montague uses terms as operators only in subject position. As direct objects and in prepositional phrases, terms are considered to be arguments of transitive verbs or prepositions respectively, but they are always of category t/(t/e), cf. Montague 1970, 251 (rules S4, S5, S6). As is well known, Montague wants to assign expressions like <u>John</u> and <u>every man</u> to the same category because they are closely related with respect to their syntactic properties. They both can occur as subject and as object (cf. Partee 1975, 236ff.). Since in Montague grammar a strict parallelism between syntax and semantics is required, syntactic kinship also has to mean semantic kinship. Therefore, Montague treats terms semantically always as sets of properties. From this in turn it follows that they syntactically must operate on expressions which denote properties (i.e. intransitive verbs).

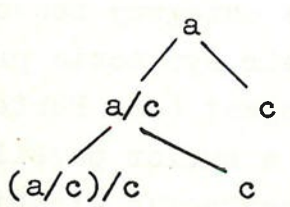
This example shows what it can mean for CG if in fact syntactic considerations are not neglected at some point. The assignment of terms to one single category has the consequence that subjects and objects have a different status with respect to the operator-argument-distinction. For the grammar of PTQ this consequence is even worse, because transitive verbs translate to simple two place expressions in intensional logic, whose arguments don't show any syntactic differences. With respect to the object language, syntactic unification at one point leads to irregularities at some other point, the reason for this being that category and function can not be related to each other in an uniform way.

One further restriction of CG, which also follows from the division of the categories into operators and arguments, becomes apparent when one looks at its possibilities to handle recursive properties of languages. No difficulties arise with argument categories as can be seen from widely used categories like n/n or iv/iv. On the other hand it is impossible to use operators as selfembedding categories since an operator always contains its dominating category. A structure like (1), which could occur in

(1) [a[a[b c] c]]

a verb phrase (a) dominating several noun phrases (c) as objects on different constituent levels, has no isomorphic correlate in CG. Instead one would have something like (2).

(2)



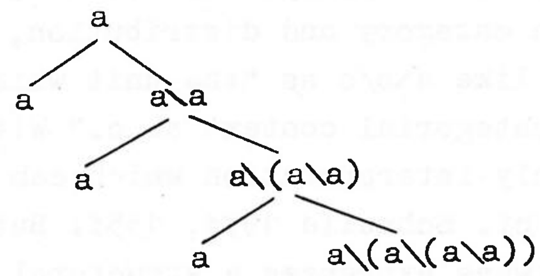
But in the present case one has to realize that the structure in (2) is more adequate than the one in (1), which is generated by the recursive rule a ⟶ a b . Since we always can tell what the number of objects can be for a given verb, no recursive rule should be applied here.[4]

Yet in natural language there can be constructions where a branching category is dominated by itself by virtue of recursive rules. Consider for instance nominal attributes in German like das Auto des Mannes der Schwester seiner Frau, to which a structure like (3) would be assigned in PSG.

(3) [a[a a[a a[a a]]]]

If we want to treat all attributes in the same way we treat attributive adjectives, i.e. as operators on nouns, then we would have to assume a structure like (4) in CG.

(4)



Of course this structure can do everything that could be done by the constituent structure (3). Nevertheless it is an extremely clear demonstration for the necessity to keep category and function distinct, since otherwise the category of a unit will be dependent of how deep that unit is embedded in a recursive construction. This would mean that we would have to give up the idea of assigning a unit to finitely many categories and instead to make use of the possibility to generate and to use infinitely many categories for CG.

Moreover, (4) exhibits a consequence from the property of categorial grammars to specify syntactic rules by specifying syntactic categories. Schnelle writes (1973, 157): "Zu jeder so eingeführten Kategorie ist im Prinzip eine Regel möglich ... Aus diesem unbeschränkten, aber definiten Reservoir wird aber jede Sprache ... nur endlich viele aussondern ... Zu jeder speziellen Sprache gibt es also, wie bei den klassifikatorischen Grammatiken, ein endliches System von Eigenschaften, nach dem die unbeschränkte Menge von Ausdrucksgestalten charakterisiert wird."[5]

Despite this, there is a significant difference with respect to PSG. In PSGs every category can be dominated by itself if there is a rule to this effect. Since there is - as far as I know - no theory which says that the recursive properties of natural language have to be explicated by means of certain categories, one should regard it as a necessary condition for grammars to allow for all categories to occur in endocentric constructions. In this respect every restriction needs special justification.

One concluding remark concerning the possible generality of our arguments seems to be necessary. Our starting point was the relationship between category and distribution, i.e. we have interpreted a category like a\b/c as "the unit which belongs to this category has the categorial context a__c." Without any doubt this is not the only interpretation which can be given to the categorial slash, cf. Schnelle 1973, 155f. But however it is interpreted, it always expresses a structural relation between constituents. And whenever this relation will be changed, the category has to be changed. Therefore category and function will always be bound together in the way we have criticised. Independance of category and function presupposes that the elements of pairs or n-tuples of constituents can be related to each other by the same grammatical relation even though their structural relations might be completely different.

Montague (1970, 249) made the guess that "It was perhaps the failure to pursue the possibility of syntactically splitting categories originally conceived in semantic terms that accounts for the fact that Ajdukiewicz's proposals have not previously led to a successful syntax." I do not think that this is quite correct but that there are more basic shortcomings of CG which prevent us from using it for the description of natural language.

NOTES

1 In Lieb 1974, 43ff. it is argued that transformational grammars do not contain any expressions "that can be understood as statements on the intended subject matter of the grammar". (43).

2 cf. e.g. Schwarz 1977, 69: "Mit zwei klaren Grundkategorien und einer einfachen Regel wird versucht, alle anderen Kategorien zu entwickeln ..."

3 A clear example where we have to "look down" into the structure are sentences like Egon zweifelt ob Paul kommt as compared to Egon zweifelt weil Paul kommt. The syntactic function of the subordinate clause becomes clear only when we look at its internal structure.

4 An other way of handling nominals in subject- and object-position, which also shows clearly how CG works, is proposed in Lambek 1958. Lambek (161f.) categorizes proper names uniformly as n whereas pronouns are assigned to different categories. One reason for this is the difference in form between he and him. For pronouns Lambek introduces categories very similar to those which much later were used by Montague and Cresswell, namely s/(n\s) for he and (s/n)\s for him. This differentiation is possible because there is a correlation between difference in form and difference in distribution in this case. The common properties of proper names and pronouns are then expressed by syntactic rules of the form $n \longrightarrow s/(n\backslash s)$. This is to say that basic categories can be expanded into derived categories within this grammar. If the difference between operators and arguments is to make any sense at all for the syntactic description, then rules of this kind certainly can not be permitted.

5 Roughly speaking, klassifikatorische Grammatiken in Schnelle's sense are grammars whose categories are to be one place predicates whose extensions are classes of expressions (op.cit., 143).

REFERENCES

Bar-Hillel, Y. (1960a): Some Linguisitc Obstacles to Machine Translation. In: Bar-Hillel, Y. (1964), 57-86.
Bar-Hillel, Y. (1960b): On Categorial and Phrase Structure Grammars. In: Bar-Hillel, Y. (1964), 99-115.
Bar-Hillel, Y. (1964): Language and Information. Reading (Mass.) 1964.
Bartsch, R./Lenerz, J./ Ullmer-Ehrich, V. (1977): Einführung in die Syntax. Kronberg.
Chomsky, N. (1957): Syntactic Structures. The Hague.
Cresswell, M.J. (1973): Logics and Languages. London.
Lambek, J. (1958): The Mathematics of Sentence Structure. American Mathematical Monthly 65 (1958), 154-170.
Lewis, D. (1972): General Semantics. In: Davidson, D./ Harmann, G. (eds.): Semantics of Natural Languages. Dordrecht. 169-218.
Lieb, H.-H. (1974): Grammars as Theories: the Case for Axiomatic Grammar (Part I). Theoretical Linguistics 1, 39-115.
Montague, R. (1970): The Proper Treatment of Quantification in Ordinary English. In: Thomason, R. (ed.)(1974): Formal Philosophy. New Haven. 247-270.
Partee, B.H. (1975): Montague Grammar and Transformational Grammar. Linguistic Inquiry VI, 203-300.
Schnelle, H. (1973): Sprachphilosophie und Linguistik. Reinbek.
Schwarz, C. (1977): Zur kategorialgrammatischen Klassifizierung von neufranzösisch <u>très</u>. Papiere zur Linguistik 17/18, 67-101.
Thomason, R.H. (1976): Some Extensions of Montague Grammar. In: Partee, B.H. (ed.) (1976): Montague Grammar. New York. 51-76.

SOME OPERATORS CANCELLING HYPERINTENSIONALITY

Michael Grabski

Hyperintensional contexts - in the sense that they are "established" by certain operators - are those in which substitution even by logically equivalent expressions does not preserve the truth value of the sentence in question.[1)] An example for an operator establishing hyperintensional contexts is believe that; believe that is more "sensitive" than, for instance, necessarily, in that the latter, but not the former, permits substitution by logically equivalent expressions *salva veritate*.

In the following, I discuss properties of German analogues of these operators (i.e. glauben daß and notwendigerweise) as well as of some other German words as my intuition, although probably strange enough, is more firm there.

The mentioned properties of glauben daß and notwendigerweise can be illustrated by the following sentences, where (1) and (2) entail (3), but (4) and (2) don't entail (5).

(1) Notwendigerweise (2 und 2) = 4.
(Necessarily (2 plus 2) = 4.)

(2) 2 = (5 minus 3)

(3) Notwendigerweise ((5 minus 3) und 2) = 4.
(Necessarily ((5 minus 3) plus 2) = 4.)

(4) Fritzchen glaubt, daß 2 und 2 = 4.
(Fritzchen believes that 2 plus 2 = 4.)

(5) Fritzchen glaubt, daß ((5 minus 3) und 2) = 4.
(Fritzchen believes that ((5 minus 3) plus 2) = 4.)

A situation (or context) in which (2) and (4) are true while (5) is false can easily be conceived: Fritzchen knows enough of arithmatics for the truth of (4), but he is far from accepting (2), although, for a more or less classical notion of "possible world", (2) should be true in all possible worlds. In other words, substitution of the term 2 by the logically

equivalent term 5 minus 3 within a sentence like (4) does not preserve its truth value.
A similar, although a bit different case is the relation between (6) and (7):

(6) Fritzchen glaubt, daß nicht der Fall ist, daß Franz nicht kommt.
(Fritzchen believes that it is not the case that Franz does not come.)

(7) Fritzchen glaubt, daß Franz kommt.
(Fritzchen believes that Franz comes.)

(6) does not entail (7), as it might be the case that Fritzchen is so "logically blind"[2] that he believes a twice negated sentence to be true without believing the sentence itself to be true.
The difference of the step from (6) to (7) to the one from (4) to (5) is that in the former no substitution is involved; instead, it is crucial how certain words occuring in (6) (especially nicht) can be "computed" with one another.
Now some words seem to interfere with glauben daß in that they change the hyperintensional contexts established by that operator into intensional ones when embedded under it.
Among these words is sozusagen (Engl. as it were, so to speak), others, which I shall not differentiate here, are gewissermassen, gleichsam, quasi.
I take sozusagen as a representative of these. It is plausible to regard it to be a 1-place sentence operator that can be inserted under glaubt daß in (4) and (5) which yields (4') and (5'):

(4') Fritzchen glaubt, daß sozusagen 2 und 2 = 4.
(Fritzchen believes that, as it were, 2 plus 2 = 4.)

(5') Fritzchen glaubt, daß sozusagen ((5 minus 3) und 2) = 4.
(Fritzchen believes that, as it were, ((5 minus 3) plus 2)= 4.)

Now, for me, (5') is entailed by (4'), together with (2) above; it is even entailed by (4), together with (2). As for (6) and (7), (7') is entailed by (6):

(7') Fritzchen glaubt, daß sozusagen Franz kommt.
(Fritzchen believes that, as it were, Franz comes.)

In the same way, (7') seems to be a consequence of (7) but not vice versa, for (7') might be true because (6) would be true and, at the same time, Fritzchen would not subscribe to the truth of Franz kommt.
To sum up the rôle of sozusagen and its mates in more pragmatical terms, it seems to permit reference to beliefs (and similar highly subjective propositional attitudes) of others by sentences which are no longer hyperintensional, that is, they are subject at least to some logical operations; these operators have a kind of "adapting function".
In (5') a logically equivalent expression has been substituted; does sozusagen permit substitution by contingently equivalent expressions? This seems not to be the case (although my intuition is weaker here), cf. (8) - (1o) (substitution of a term) and (11), (12) (substitution of a sentence):

(8) Fritzchen glaubt, daß Franz kommt.
(Fritzchen believes that Franz comes.)

(9) Franz = der Gewinner des Wettrennens.
(Franz = the winner of the race.)

(1o) Fritzchen glaubt, daß sozusagen der Gewinner des Wettrennens kommt.
(Fritzchen believes that, as it were, the winner of the race comes.)

(11) Franz kommt genau dann, wenn Lisa kommt.
(Franz comes if, and only if, Lisa comes.)

(12) Fritzchen glaubt, daß sozusagen Lisa kommt.
(Fritzchen believes that, as it were, Lisa comes.)

Sentences like (1o) and (12) show, that also when sozusagen has been inserted, all belief worlds of the believer have to be considered (as for Fritzchen, not in all of them (9) or (11) have to be true). Thus, in hyperintensional contexts, sozusagen does not re-establish extensional contexts for the sentence embedded under it, but hyperintensionality is changed to "normal" intensionality.

1. Writing down the semantical properties of sozusagen that have been discussed so far presupposes the existence of a semantical analysis of operators like glauben daß that takes into consideration the fact that these operators establish hyper-

intensional contexts.
Within the approaches along model theoretical lines I know there have been different successful solutions to the problem how substitution by logically equivalent expressions can be made impossible, but it seems to me, that the price that has been paid in all of these approaches is to high.
The price is in general, that the sentence embedded under glauben daß is not interpreted as a normal proposition but as an entity that is made incomparable to normal propositions in some way or other; the effect is always that substitutions of the indicated kind are made impossible. On the other hand, the reflections made on sozusagen above show that it is worth while not to change the semantical type of the interpretation of sentences embedded under operators like glauben daß.
A deep-going change of character is imposed on the interpretation of the embedded sentence in solutions like Hill (1976), where the embedded sentence is interpreted as a complex object that, in any case, contains a syntactical individual, i.e. the embedded sentence itself (or a sentence that is"innocuously equivalent"). It is not clear how semantical relations of the kind mentioned in the preceding paragraph could be treated in this approach; probably, what one would expect to be relations between propositions would have to be reconstructed as relations between individuals.
Moreover, this approach seems to have difficulties with sentences like

(13) Cäsar glaubt, daß Kleopatra ihn liebt.
(Caesar believes that Cleopatra loves him.)

In this approach, (13) comes out false for all situations in which Caesar does not know German (English).
As for that, better proposals have been made in Cresswell (1974) and Thomason (1974) as they try to come to grips with the hyperintensionality of the embedded sentence by purely semantical means.
In Cresswell's approach the sentence embedded under glauben daß is interpreted, according to its construction, as a function that is applied to certain arguments. Now, like in the normal case, i.e. when the sentence in question is not embedded under glauben daß, the function has the character of an intension; the arguments, on the other hand, are in general entities of a

different kind, i.e. they are meanings of the syntactical parts of the sentence in question.
This notion has been defined in Lewis (1972) as follows: For an expression α that is not complex, the meaning of α = the intension of α ; if α is complex, i.e. α consists of a functional expression δ that has been applied to argument expressions $\beta_1, \ldots, \beta_n$ intension and meaning are defined as something different:

(14) Intension (α) = Intension (δ)(Intension (β_1),..., Intension (β_n))

whereas

(15) Meaning (α) = <Meaning (δ), Meaning (β_1),..., Meaning (β_n)>

The advantages of Cresswell's approach are obvious: the notion of a meaning of an expression is defined in terms of its intension (together with the notion of sequence); thus, no part of a sentence combined with glauben daß is interpreted by entities of a principally non-semantic kind. And, for all expressions α the meaning of α is an unequivocal representation of α's syntactical construction. (6) and (7) above are sentences whose interpretation came out differently according to this solution.
Still, two objections can be raised:
For certain reasons, Cresswell wants the differentiation between the intension and the meaning of an expression to apply only to the argument expressions of a sentence embedded under glauben daß, that is, for a construction a glaubt daß ϕ, where ϕ itself is of the form $\delta\beta_1 \ldots \beta_n$, ϕ is to be interpreted as

(16) Intension (δ)(Meaning (β_1),..., Meaning (β_n))

There certainly occur cases where some β_m is not complex, but where there still is hyperintensionality, as in

(17) Fritzchen glaubt, daß Cicero großartig ist.
(Fritzchen believes that Cicero is great.)

This sentence is syntactically so simple that, when the prescription (16) has been applied for its interpretation, there are no meanings of complex expressions which would bring out the difference between (14) and (15) above.

In general, this solution seems to rest on the assumption that hyperintensional properties of sentences depend on the way the non-complex expressions occurring in them are reckoned up, not on the meanings of the non-complex expressions themselves.
Now, for this solution, the substitution of Cicero in (17) by Tullius (Tully) which possibly would change the truth value of the sentence cannot be taken account of, unless one would regard the two terms as only contingently equivalent; that is, one would make this case a matter of intensions; in Kripke (1972) at least strong philosophical reasons are given which point into another direction.
Secondly, it is not clear how entities within a universe of discourse can be referred to by a sequence like the one in the second half of (15); one would prefer a more natural interpretation of hyperintensional sentences, one in which not the possiblity of constructing consequences of these sentences was blocked at a certain depth by the sequences.
The approach in Thomason (1974) is about like this: Principally, expressions of the object language are assigned two different interpretations, an "objective" one according to which e.g. sentences are interpreted as functions from words into truth values, as well a "subjective" one according to which sentences are assigned functions from worlds into a set of sets of persons; intuitively, a sentence is interpreted, at a world, as the set of persons believing that sentence to be true at that world.
By this assignment for sentences, together with one for terms, assignments for further categories can be defined. Belief operators then involve the subjective interpretation for sentences embedded under them which makes it possible to stop certain substitutions that otherwise, i.e. within objective interpretations, would be possible.
This seems to be the most successful approach considered up till now; the logical properties of sentences containing glauben daß are met, at the same time the contributions made by the parts of the embedded sentence are taken account of.
Still I should prefer to modify this approach, such that the subjective and objective interpretation of an expression are allowed for by a "double indexing".
The idea is that hyperintensional operators, such as glauben daß

not only take into consideration other worlds (for a belief operator just those worlds that are compatible with everything a person believes), but also the way in which the facts in those other worlds are expressed and that this way of expression partly reflects the linguistic usage of the person to which the "propositional attitude" in question is assigned.[3)]
That (5) and (7) fail to be consequences of (4) and (6) respectively, can, along these lines, be explained by saying that Fritzchens linguistic usage is such that it does not permit substitution of 2 by 5 minus 3 and that the interpretation of the negation sign is different from that of classical logic.
Sozusagen then is an operator that leaves everything as it is, but for the "linguistic usage" aspect of the contexts taken into account by the hyperintensional operator; what sozusagen does is to copy the actual linguistic usage that has been preserved by the second place of the ordered pair of contexts onto the contexts where the embedded sentence has to be true.
As the actual context (here) is one where classical logic and arithmetic are compulsive for the interpretations of certain constants, hyperintensionality is removed from the embedded sentence, that is, the contexts where its truth is at stake are all "linguistically normalized".

2. Let an interpretation $\mathfrak{A}$ for an object language L that contains the operators discussed be a sequence with at least the following elements: sets U, C, W, SG, and functions f_W, f_{SG}, ..., a family of relations $\langle R_u^{Bel} \rangle$ for $u \in U$, and a function F.
U, C, W, SG are the sets of individuals, contexts, possible worlds, and linguistic usages, respectively; f_W and f_{SG} are such that for $c \in C$, $f_W(c) \in W$ and $f_{SG}(c) \in SG$, i.e. f_W and f_{SG} are functions that determine, for each context, certain "aspects" of that context.[4)]
As for SG, it can be conceived of as the set of all possibilities constants of L are assigned intensions of (categorially) appropriate kind, cf. Thomason (1973).
F is a function that assigns intensions to the constants of L; it is important that this is done dependent on elements of SG.
R_u^{Bel} is a relation $\subseteq W \times W$ and gives, in a classical way, for the worlds in its domain, the worlds that are, for u, "doxastic alternatives" of these worlds.

Then, within the exaples given above, what is affected by the (possibly) alien linguistic usage of the believer, is the interpretation of constants, including "logical" ones like the negation sign.
It remains open if hyperintensionality also concerns the categorial properties of constants and the semantic rules. At least from one point of view this seems to be not the case, as, on the one hand, operators like believe that are combined only with sentences that are well-formed in the sense of the actual syntax (in this way these operators are different from quotation marks), and, on the other hand, the expressions combined with these operators have to be interpreted in a way that makes clear (also) the actual meaning of the expression (otherwise the belief etc. of the subject of the operator would not be expressed by the sentence).
What, then, is changed by the operators in question? It seems plausible to explain the fact that (7) is not entailed by (6) by saying that at those contexts that are relevant for Fritzchens belief nicht is not interpreted classically and that Fritzchen is, as for that, "logically blind".
A different though similar case is exemplified by (3) - (5); here the linguistic usage governing the contexts relevant for Fritzchens belief does not contain a meaning postulate yielding the equivalence of 2 and 5 minus 3.
These cases can be summed up as follows: the first case arises if the constants occurring in the embedded sentence are interpreted differently in relevant contexts, the second case arises if certain meaning postulates that introduce further constants lack in relevant contexts.
Yet if a sentence of the kind a glaubt daß ϕ (a believes that ϕ) is true, the following has to be the case: at the relevant contexts ϕ gets the same truth value (i.e. true) according to the interpretation the constants in ϕ get (1) by the linguistic usage of the actual context and (2) by the linguistic usage predominant for the individual denoted by a.
ϕ has to be true (1) at just those contexts whose worlds (for this individual) are doxastic alternatives to the world of the actual context; likewise ϕ has (2) to be true at all contexts like them, but for the difference that the linguistic usage there (possibly) is not the actual one.

Probably it would be too weak to postulate that the linguistic usage at the second set of contexts just is different from the actual one; for each world being a doxastic alternative to the actual world, far too many contexts could be made up. Perhaps an ad-hoc device of the following sort will do here: let there be a function f^{des} that gives, for each person u, a designated context; this context can be imagined to be the actual one for that person, or the one that is linguistically ideal, or it somehow represents his theory about the world, or something of this kind.

If we accept this so far we can write down the truth conditions for sentences with glauben daß and sozusagen ($\mathcal{D}_\alpha$ is a function that maps expressions of L into intensions according to α):

(18) $\mathcal{D}_\alpha$(a glaubt daß ϕ)($\langle c_o', c_o \rangle$) = 1 iff for all $c \in C$

such that

a) $f_W(c) \in \{w \mid f_W(c_o)\ R^{Bel}_{\mathcal{D}_\alpha(a)(\langle c_o', c_o \rangle)}\ w\}$

b) $f_X(c) = f_X(c_o)$ if $X \notin \{W, SG\}$

it is the case that

c) (1) if $f_{SG}(c) = f_{SG}(c_o)$, then $\mathcal{D}_\alpha(\phi)(\langle c, c_o \rangle) = 1$

and

(2) $\mathcal{D}_\alpha(\phi)(\langle c', c_o \rangle) = 1$, where c' is like c, with the possible difference that

$f_{SG}(c') = f_{SG}(f^{des}(\mathcal{D}_\alpha(a)(\langle c_o', c_o \rangle)))$

(19) $\mathcal{D}_\alpha$ (sozusagen ϕ)($\langle c_o', c_o \rangle$) = 1 iff

$\mathcal{D}_\alpha(\phi)(\langle c_o'', c_o \rangle) = 1$, where

a) $f_{SG}(c_o'') = f_{SG}(c_o)$

b) $f_X(c_o'') = f_X(c_o')$ if $X \neq SG$

3. The working of (18) can be demonstrated, for the problem of the sentences (6) and (7), by a small interpretation α' of the kind sketched above.

Let propositions (= intensions of sentences) be functions in

$2^{C \times C}$, i.e. they can be represented by sets of pairs of contexts. A sentence ϕ' is a consequence of a sentence ϕ iff the proposition assigned to ϕ by all interpretations is included in the one assigned to ϕ'.

Let α' be such that $U = \{u\}$, $C = \{c_o, c_1, c_2\}$, $W = \{w_o, w_1, w_2\}$, $SG = \{sg_o, sg_1\}$, $f_W = \{<c_o,w_o>, <c_1,w_1>, <c_2,w_2>\}$, $f_{SG} = \{<c_o, sg_o>, <c_1,sg_1>, <c_2,sg_1>\}$, $f^{des}(u) = c_1$, $R_u^{Bel} = \{<w_o,w_1>, <w_o,w_2>\}$.

As the interesting difference between sg_o and sg_1 is the interpretation of <u>nicht</u> let them both be alike except that:

(2o) $F(sg_o)$(<u>nicht</u>) = that function

$g \in (2^{C \times \{c_o\}})^{(2^{C \times \{c_o\}})}$ that, for $p \in 2^{C \times C}$, $<c,c_o> \in C \times \{c_o\}$,

$g(p)(<c,c_o>) = 1$ iff $p(<c,c_o>) = o$.

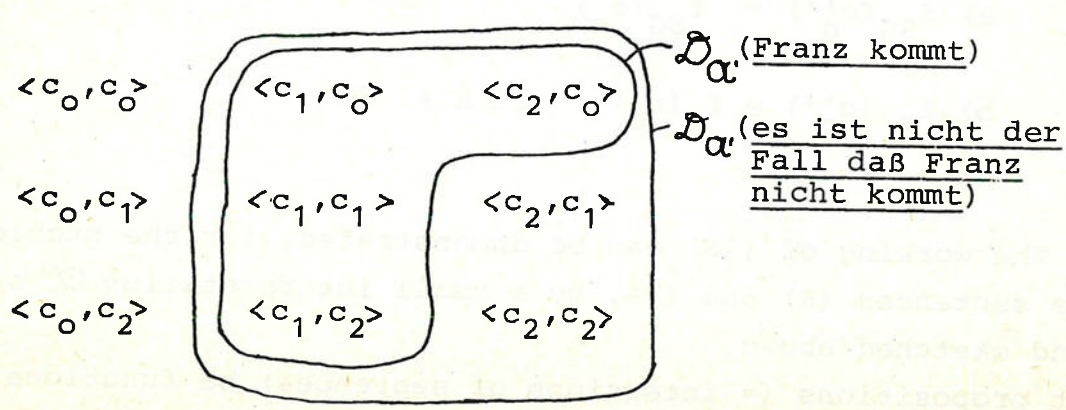
$F(sg_1)$(<u>nicht</u>) = that function

$g \in (2^{C \times \{c_1\}})^{(2^{C \times \{c_1\}})}$ that, for $p \in 2^{C \times C}$, $<c,c_1> \in C \times \{c_1\}$,

if $g(p)(<c,c_1>) = 1$, then not $p(<c,c_1>) = 1$.

Let <u>Franz kommt</u> be an unanalysed sentence, then within the matrix of pairs of contexts we can, as for the distribution of the intension of <u>Franz kommt</u>, have the following situation:

(21)



$\mathfrak{A}'$ is a case where (6) above comes out true, at the contexts whose worlds (for u) are doxastic alternatives to w_o; the sentence es ist nicht der Fall, daß Franz nicht kommt is true there, both from the linguistic point of view of sg_o (= $f_{SG}(c_o)$) and from that of sg_1 (= $f_{SG}(c_1) = f_{SG}(c_2)$).
Franz kommt, however, is not true at all these contexts; $\langle c_2, c_1\rangle$ and $\langle c_2, c_2\rangle$ show that $F(sg_1)$(nicht) allows for a "weaker" meaning of that word.
Therefore, $\mathfrak{A}'$ is a case for (6) being true and (7) being false.

4. The truth conditions for glauben daß as given in (18) are still insufficient in an important respect. No account is taken of the possibility that the "alien" linguistic usage may be such that it does not provide an interpretation for the words in ϕ. (13) is an example of this situation as the linguistic usage that is actual for Caesar probably is some form of Latin, i.e. a linguistic usage with a syntax different from that of (13).
Generally, one could imagine that, to handle these cases, one would need some kind of translation of ϕ; an interpretation then should contain meaning postulates for semantical relations between words of different syntaxes.
This would mean that a structure would be defined on SG which would take account of the fact that

1. there are sets of linguistic usages that do not differ as for their syntax (or: as for relevant parts of their syntax) - one might define languages as such sets -;
2. dependent on time, history, experience, knowledge, theories etc. there will be a variable translatability between constants of different linguistic usages, perhaps even of their syntactic rules.

That the translatability of constants is dependent on theories that define them for a given linguistic usage explains why (22) sounds strange:

(22) Plato believed that the gene pool of the Guardians would probably produce some children with low I.Q.'s.[5)]

(18') is a modification of (18) which tries to capture these reflections; it is sufficient to modify only clausula c)(2).
((18') presupposes that the notion "translation of a sentence"

has been defined.)

(18') a) .... }
b) .... } as in (18)
c) (1).. }

(2) $\mathcal{D}_{\alpha}(\phi')(\langle c', c_o \rangle) = 1$, if there is $\phi' \in \mathcal{A}_L$ and if c' is as c, with the possible difference that $f_{SG}(c') = f_{SG}(f^{des}(\mathcal{D}_{\alpha}(a)(\langle c'_o, c_o \rangle)))$ where

① $\mathcal{A}_L$ is the set of expressions whose interpretation is defined by $f_{SG}(f^{des}(\mathcal{D}_{\alpha}(a)(\langle c'_o, c_o \rangle)))$

② ϕ' is a translation of ϕ. (If there is no such ϕ' $\mathcal{D}_{\alpha}$(<u>a glaubt daß ϕ</u>)$(\langle c'_o, c_o \rangle)$ is not defined)

It is to be expected that different classes of words show different translatability properties.
Probably indexical words vary less within one "language" than others (cf. <u>I</u>, <u>you</u>, <u>tomorrow</u>, etc.), whereas scientific terms probably vary considerably within one "language" (cf. the simple mathematical terms <u>2</u>, <u>5 minus 3</u> discussed above), but are easily translatable across "languages", in any case into words interpreted by a scientific linguistic usage; this would explain why scientific vocabulary can easily be taken over by other languages.
Sentence (22) shows that also time plays a crucial rôle for translatability. Here, translations of scientific terms are concerned, and although the linguistic usage we attribute to Plato is a more or less scientific one, (22) sounds strange (not false). I think it does so because a term like <u>gene pool</u> didn't have an equivalent in Plato's time.
In this, there is a difference between (22) and the situation expressed by (23) in which <u>Fritzchen</u> is to be a contemporary (and which might be true):

(23) Fritzchen believes that the gene pool of Plato's Guardians probably would produce some children with low I.Q.'s.

To conclude with, (24) shows that the operators of the <u>sozusagen</u>

class are able to copy linguistic usages on others that are remote in language and time:

(24) Plato believed that, as it were, the gene pool of the Guardians would probably produce some children with low I.Q.'s.

Notes:

1) The term has been coined, as far as I know, in Cresswell (1974)
2) Cf. Cresswell (1974) p. 2.
3) The differentiation of worlds and ways of fixing the interpretation of (the constants of) a language is due to Thomason (1973)
4) The relations between certain kinds of indices such as contexts and worlds have been reflected and worked out in great detail in Lutzeier (1974).
5) The sentence is taken from Hill (1976), p. 215.

References:

Cresswell, M. (1974) Hyperintensional Logic, to appear in Studia Logica

Hill, Chr.S.(1976) Toward a Theory of Meaning for Belief Sentences, Philosophical Studies 3o, pp. 2o9-226

Lewis, D. (1972) General Semantics, in: D.Davidson and G.Harman (Ed.) Semantics of Natural Language, Dordrecht pp. 169-218

Lutzeier, P.(1974) Der Aspekt "Welt" als Einstieg zu einem nützlichen Kontextbegriff für eine natürliche Sprache, unpubl. dissertation Univ. of Stuttgart

Montague, R.(197o) English as a Formal Language, in:B.Visentini et al.(Ed.) Linguaggi nella Società e nella tecnica, Milan, pp.189 -224

Thomason, R.(1973) Necessity, Quotation, and Truth: An Indexical Theory, unpubl.paper, Univ. of Pittsburgh

Thomason, R.(1974) An Interpretation of First Order Logic with Belief, unpubl.paper, Philosophy Dept. Univ. of Pittsburgh

EPISTEMIC PRAGMATICS: OUTLINES AND TWO APPLICATIONS¶

Jeroen Groenendijk
Martin Stokhof

0. Introduction

The basic assumption from which the contents of this paper start is that an adequate theory of meaning of a language should consist of at least a semantic theory, which gives a recursive characterization of the truth conditions of the sentences of that language, and a pragmatic theory, which gives a recursive characterization of the correctness conditions. One of the grounds for adopting this assumption is that there are aspects of the meaning of certain expressions and constructions which cannot be captured in terms of truth conditions, but which should be described in terms of the conditions under which these expressions and constructions can be used correctly.(1) An important and interesting part of these conditions are those which concern the information of language users. The part of a pragmatic theory which deals with these conditions we call 'epistemic pragmatics'. The main Gricean maxims of quality, quantity and relation can be reformulated in such a theory, not, however, as mere slogans but as correctness conditions which are recursively characterized.
In order to state the correctness conditions concerning the information of language users, an epistemic pragmatic theory should contain a framework which fulfills the following requirements: (a) the information of language users about a compound sentence can be computed from the information they have about its compounds; and (b) the information about an atomic sentence can be computed from the information about the denotation of the expressions occurring in it. Furthermore, the framework should enable one to represent the information of language users about the information of other language users in such a way that conditions (a) and (b) are met.

In this paper we want to illustrate this concept of an epistemic pragmatic theory by developing a framework fulfilling the requirements mentioned above for two very simple languages. After we have done so, we will concentrate not on the formulation of correctness conditions(2), but on the analysis of two problems. The first problem concerns the distinction between the specific and non-specific use of terms, such as a picture in the sentence a picture is missing from the gallery. The second problem concerns the de dicto/de re distinction between objects of belief, as in the sentence John believes that Peter passed the exam.

1. Propositions and connectives

We will now sketch an epistemic pragmatic theory for a simple propositional language containing propositional variables, p_0, p_1, p_2..., and the usual connectives $\neg$, &, v, en $\rightarrow$. The core of such a theory is a definition of the notion of an 'epistemic model'. The purpose of defining epistemic models is to evaluate formulas of the language with respect to (the information of) a language user. Therefore, an epistemic model EM will contain, as one of its components, a non-empty set I of language users. The values which are assigned by the valuation function of a model to pairs consisting of a formula and a language user, represent the possible situations with respect to the information a language user may have about the truth status of a formula.
It is important to notice that the phrase 'information about the truth status of a formula' is used without any factual implications. I.e. the phrase is used in such a way that a language user can be said to have the information that a formula ϕ is the case even in a situation in which ϕ in fact is false.
One can distinguish three possible situations with respect to the information a language user x may have about the truth status of a formula ϕ:

- x has the information that ϕ is true;
- x has the information that ϕ is false;
- x has no more information than that either ϕ is true or ϕ is false: i.e. x has no opnion about ϕ.

These three situations are represented by the epistemic values {1}, {0}, {0,1}, respectively, where 1 and 0 represent the truth values 'true' and 'false'. As remarked above, we should also be able to talk about the information a language user x has about the information a language user y has about a formula ϕ, or, generally, about the information x_1 has about ... the information x_n has about ϕ. Therefore, the valuation function assigns epistemic values to pairs of formulas and n-tuples of language users. Some examples of possible situations and the values representing them, are:

- x has the information that y has the information that ϕ is true: {{1}};

- x has the information that y has no opinion about ϕ: $\{\{0,1\}\}$;
- x has no more information than that either y has the information that ϕ is true, or y has the information that ϕ is false, or y has no opinion about ϕ, i.e. x has no opinion about what information y has about ϕ: $\{\{1\},\{0\},\{0,1\}\}$;
- x has the information that y has the information that it is not the case that z has no opinion about ϕ: $\{\{\{0\},\{1\}\}\}$.

The general picture is as follows: starting from S_0, the set of truth values $\{0,1\}$, we build S_1, the set of epistemic values that can be assigned to an ordered pair consisting of a formula and one language user, as follows:

$S_1 =_{df} POW(S_0)\setminus\{\emptyset\}$, where POW is the powerset operation

From S_1 we build S_2, the set of epistemic values that can be assigned to a pair consisting of a formula and a sequence of two language users, in the same way. In general:

$S_n = POW(S_{n-1})\setminus\{\emptyset\}$, for all $n>0$

We can now define an epistemic model EM as a triple $\langle I,\{0,1\},V\rangle$, in which V is a valuation function taking ordered pairs consisting of a formula and an n-tuple of language users into S_n. The clause for atomic formulas and the ones for negation and conjunction of the recursive definition of V are:

(1) $V(p,i^n)\in S_n$, for every ordered n-tuple i^n of elements of I

(2) $V(\neg\phi,i^n) = NEG[V(\phi,i^n)]$, where NEG is defined as follows:
$NEG[1] = 0$; $NEG[0] = 1$; $NEG[X] = \{NEG[x] \mid x\in X\}$, for $X \neq 1,0$

(3) $V(\phi\&\psi,i^n) = CONJ[V(\phi,i^n),V(\psi,i^n)]$, where CONJ is defined as follows:
$CONJ[1,1] = 1$; $CONJ[0,1] = CONJ[1,0] = CONJ[0,0] = 0$;
$CONJ[X,Y] = \{CONJ[x,y] \mid x\in X \;\&\; y\in Y\}$, for $X,Y \neq 1,0$

The other connectives can be defined in terms of negation and conjunction in the usual way.
This framework meets the first compositionality requirement mentioned in section 0. For example, if $V(p,x) = \{0,1\}$ and $V(q,x) = \{1\}$, then the definitions imply that $V(\neg p\&q,x) = \{0,1\}$.
Note that this framework 'contains' the ordinary semantics of two-valued propositional logic, in this sense that the values assigned to pairs consisting of a formula and the empty sequence of language users, are as in ordinary propositional logic.

The framework can be used to formulate correctness conditions concerning the information of language users. For any formula ϕ, it can be defined whether it is correct for a language user x to utter ϕ adressing a language user y in a situation described by an epistemic model EM. In this way, the main Gricean maxims can be captured (3).

The framework as it is sketched here gives the basic tools for the representation of information of language users, but it needs to be (and has been (4)) enriched in such a way that also information of language users about logical and non-logical dependencies between formulas can be incorporated. Incorporation of information about dependencies between formulas is needed, for example to give a more satisfactory treatment of disjunctions. In the framework sketched here, a language user can only have the information that a disjunction is true if he has of (at least) one of the disjuncts the information that it is true.
We will not incorporate information about dependencies between formulas in this paper, because it is not strictly necessary for the treatment of the problems we want to discuss here.

2. Predicates, constants and quantifiers

In this section we will outline how the framework can be extended to apply to a language with predicates, constants and quantifiers. In doing so, we illustrate how the framework fulfills the second compositionality requirement that the information a language user has about a sentence must be computed from the information he has about the denotation of the expressions occurring in it.

Let us first limit ourselves to a language only containing one-place predicates, P, Q, P_0, P_1, ..., and individual constants, a, b, c, a_0, a_1, ... (and the usual propositional connectives).
The actual denotations of individual constants are, as usual, individuals. And the actual denotations of one-place predicates are sets of individuals. With respect to the information a language user x may have about the denotation of an individual constant α a variety of situations are possible. Some examples of such situations and their representations in the framework are:
- x has the information that the individual constant α denotes the individual $\underline{a}$, represented as $\{\underline{a}\}$;

- x has no more information than that the individual constant α either denotes the individual a, or denotes the individual b, his information does not tell him which of the two it is, but he does know that it is one and only one of them: $\{a,b\}$;
- x has no information at all about the denotation of α, as far as his information goes, the denotation of α could be any individual in the domain A: A.

The information a language user x has about the denotation of an individual constant α is represented as a subset of the domain of individuals A, specified in the epistemic model. The interpretation function F of an epistemic model assigns subsets of the set of individuals A to pairs consisting of an individual constant and a language user.

With respect to the information of a language user about the denotation of one-place predicates the situation is quite similar. Just as in the case of individual constants, the interpretation function F assigns sets of what are the actual denotations of one-place predicates, being sets of sets of individuals, to pairs consisting of a predicate and a language user. Examples of situations with respect to the information a language user x may have about the denotation of a one-place predicate δ and the corresponding representations are:

- x has the information that the predicate δ is true of the individuals a, b and c (and false of all other individuals): $\{\{a,b,c\}\}$;
- x has the information that the predicate δ is true of either a and b, or a, b and c, i.e. x is not sure whether δ is true of c, but he is sure that δ is true of a and b and false of all other individuals: $\{\{a,b\},\{a,b,c\}\}$;
- x has the information that δ is true of all individuals: $\{A\}$;
- x has the information that there is no individual of which δ is true: $\{\emptyset\}$.

As before, we not only want to represent the information of a language user about the denotation of constants and predicates, but also the information of x_1 about ... the information of x_n about the denotation of these expressions. Some examples:

- x has the information that y has the information that α denotes a: $\{\{a\}\}$;
- x has the information that either y has the information that α denotes a, or y has the information that α denotes b: $\{\{a\},\{b\}\}$;
- x has the information that y has the information that either α denotes a or b: $\{\{a,b\}\}$;
- x has the information that y has the information that z has the information that δ is true of two of the three individuals a, b and c: $\{\{\{\{a,b\},\{b,c\},\{a,c\}\}\}\}$

These considerations lead us to the following notion of an epistemic model for this language. An epistemic model EM is a quintuple $\langle A, I, \{0,1\}, F, V\rangle$, in which A is the domain of individuals, of which I, the set of language users, is a subset. $\{0,1\}$ is the set of truth values. F is the interpretation function which is defined as follows:

$F(\alpha, i^n) \in A_n$, for all individual constants α and n-tuples i^n of elements of I;
$F(\delta, i^n) \in A'_{n+1}$, for all predicates δ and n-tuples i^n

The domains A_n, A'_n are constructed from the domain A as follows:

$A_0 =_{df} A$; $A_n =_{df} POW(A_{n-1}) \setminus \{\emptyset\}$, for n>0

$A'_0 =_{df} POW(A)$; $A'_n =_{df} POW(A'_{n-1}) \setminus \{\emptyset\}$, for n>1

To the definition of the valuation function V we add the following clause for the evaluation of atomic formulas:

(0) $V(\delta(\alpha), i^n) = T[F(\delta, i^n), F(\alpha, i^n)]$, where T is defined as follows:
(a) if $x \in A$, $Y \in A'_1$, then $T[x,Y] = 1$ iff $x \in Y$, $= 0$ otherwise;
(b) if x,Y otherwise, then $T[x,Y] = \{T[z,U] \mid z \in x \ \& \ U \in Y\}$

The following examples may serve to illustrate this definition:

- suppose $F(a_1, x) = \{\underline{a}\}$ and $F(P_0, x) = \{\{\underline{a},\underline{b}\}\}$, then $V(P_0(a_1), x) =$ $\{T[z,U] \mid z \in \{\underline{a}\} \ \& \ U \in \{\{\underline{a},\underline{b}\}\}\} = \{T[\underline{a},\{\underline{a},\underline{b}\}]\} = \{1\}$;
- suppose $F(a_3, x) = \{\underline{a},\underline{c}\}$ and $F(P_4, x) = \{\{\underline{a},\underline{b}\},\{\underline{a},\underline{b},\underline{c}\}\}$, then $V(P_4(a_3), x) =$ $\{T[\underline{a},\{\underline{a},\underline{b},\underline{c}\}], T[\underline{a},\{\underline{a},\underline{b}\}], T[\underline{c},\{\underline{a},\underline{b},\underline{c}\}], T[\underline{c},\{\underline{a},\underline{b}\}]\} = \{0,1\}$;
- suppose $F(a_8, xy) = \{\{\underline{a}\}\}$ and $F(P_9, xy) = \{\{\{\underline{a},\underline{b},\underline{c}\}\},\{\{\underline{a},\underline{b}\},\{\underline{a},\underline{b},\underline{c}\}\}\}$, then $V(P_9(a_8), xy) = \{T[\{\underline{a}\},\{\{\underline{a},\underline{b},\underline{c}\}\}], T[\{\underline{a}\},\{\{\underline{a},\underline{b}\},\{\underline{a},\underline{b},\underline{c}\}\}]\} =$ $\{\{T[\underline{a},\{\underline{a},\underline{b},\underline{c}\}]\},\{T[\underline{a},\{\underline{a},\underline{b}\}], T[\underline{a},\{\underline{a},\underline{b},\underline{c}\}]\}\} = \{\{1\},\{1,1\}\} = \{\{1\}\}$.

Let us now extend the language with the existential and universal quantifiers. We start from the observation that in order to compute the information of a language user x about the truth status of quantified sentences such as <u>All men are mortal</u> and <u>A picture is missing</u> it suffices to know what information x has about the denotation of the predicates <u>man</u>, <u>mortal</u>, <u>picture</u>, and <u>missing</u>. In fact, this is a special case of the general point that it is possible to do predicate logic without any variables (cf. Quine 1966). Our reason for doing away with variables in the context of epistemic pragmatics is that it is quite unclear to us what in that context the relation between variable assignments and the information of language users would be.

For the purposes of this paper it suffices to take into consideration only one kind of quantified formulas, viz. those which are constructed according to the following rule: if δ and γ are one-place predicates, then $\exists\delta(\gamma)$ and $\forall\delta(\gamma)$ are wellformed formulas.[5] The valuation of quantified formulas comes to a comparison of the interpretations of the predicates occurring in them. In this comparison the notions of intersection and subset play a role, as one might have expected. The clauses in the definition of the valuation function V for the kind of quantified formulas defined above, are as follows:

(4) $V(\exists\delta(\gamma), i^n) = E[F(\delta, i^n), F(\gamma, i^n)]$, where E is defined as follows:

(a) if $X, Y \in A_1'$, then $E[X,Y] = 1$ iff $X \cap Y \neq \emptyset$, $= 0$ otherwise;

(b) if X,Y otherwise, then $E[X,Y] = \{E[x,y] \mid x \in X \;\&\; y \in Y\}$

(5) $V(\forall\delta(\gamma), i^n) = U[F(\delta, i^n), F(\gamma, i^n)]$, where U is defined as follows:

(a) if $X, Y \in A_1'$, then $U[X,Y] = 1$ iff $X \subseteq Y$, $= 0$ otherwise;

(b) if X,Y otherwise, then $U[X,Y] = \{U[x,y] \mid x \in X \;\&\; y \in Y\}$

Consider the following two very simple examples:

- suppose $F(P_8, x) = \{\{\underline{a},\underline{b}\},\{\underline{a},\underline{b},\underline{c}\}\}$, and $F(P_4, x) = \{\{\underline{a},\underline{b},\underline{c},\underline{d}\}\}$, then $V(\forall P_8(P_4), x) = \{U[\{\underline{a},\underline{b}\},\{\underline{a},\underline{b},\underline{c},\underline{d}\}], U[\{\underline{a},\underline{b},\underline{c}\},\{\underline{a},\underline{b},\underline{c},\underline{d}\}]\} = \{1\}$;
- suppose $F(P_2, x) = \{\{\underline{a},\underline{c}\}\}$ and $F(P_1, x) = \{\{\underline{a},\underline{b}\},\{\underline{b}\}\}$, then $V(\exists P_2(P_1), x) = \{E[\{\underline{a},\underline{c}\},\{\underline{a},\underline{b}\}], E[\{\underline{a},\underline{c}\},\{\underline{b}\}]\} = \{0,1\}$.

To conclude this section, it should be noted that not only has the language defined here a very limited expressive power, but that the epistemic framework developed for this language suffers from the same defect as the framework for propositional languages sketched in section 1: it needs to be extended in order to be able to represent information ot language users about logical and non-logical dependencies between the denotations of expressions. However, for the present purposes these aspects may be left out of consideration.

3. Specific and non-specific use of terms

In this section we will indicate how the framework outlined in the previous sections can be applied to the analysis of some aspects of the distinction between the specific and non-specific use of terms. This distinction, wellknown and often discussed, but not always understood in exactly the same way, can best be clarified by the following example[6]:

(1) A picture is missing from the gallery

There are two different kinds of circumstances in which this sentence can be used.

One can use the term a picture to refer to a specific piece of art, and one can use it to refer to any old picure exhibited in the gallery. Although perhaps the first kind of circumstances may seem to be more likely, the second is certainly possible. To follow Kasher & Gabbay 1976, one can imagine a gallery fitted with an alarm system which gives an acoustic signal in a control room whenever one of the pictures hanging in the gallery is taken away. The guard, alarmed by this signal, could utter sentence (1) without having any idea about which of the pictures is missing. This latter kind of use we call non-specific, the former specific.

As is not too often noted, this distinction seems to apply not only to the use of indefinite terms, but also to that of definite terms, such as proper names and definite descriptions. Consider the following sentence:

(2) Mary talked to dr. Johnson

The proper name dr. Johnson can be used non-specifically, i.e. without having any idea as to who the referent of the name dr. Johnson is, whereas at the same time sentence (2) is used correctly, e.g. if one knows that Mary talked to every veterinarian in town and that one of them happens to be called dr. Johnson. This kind of use of proper names is even a necessary requirement for the correct utterance of sentences such as:

(3) I wonder which of these men is John Smith

(4) I don't know who John Smith is

The same holds, mutatis mutandis, for the use of definite descriptions like the most competent veterinarian in town, and even, as we would argue, for terms like every veterinarian in town.

The main controversy over the specific/non-specific dichotomy is whether it constitutes a semantic or a pragmatic distinction. If one upholds that it is a semantic distinction, then one is committed to the view that a sentence like

(5) John talked to a Swede

is ambiguous. I.e. that (5) has two distinct readings, represented by two different logical forms which have different truth conditions. If one claims that it is a pragmatic distinction, one claims that (5) is not ambiguous, i.e. that there is just one kind of circumstances under which (5) is true, and one explains the distinction in terms of the different kinds of circumstances under which (5) can be used correctly. Kasher & Gabbay have claimed indeed that the specific/non-specific distinction is a semantic one, i.e. that (5) is ambiguous. According to them, (5) is true on the specific reading if and only if John talked to some Swede whom the speaker can 'canonically identify' (i.e. can identify by means of a proper name or a suitable definite description). On the non-specific reading then, (5) would be true if and only if there is some Swede whom John talked to. We agree wholeheartedly with Klein when he says:

> "This line of reasoning is, in my opinion, patently incorrect. However strongly a person who uttered (3.17) [our (5) Gr/St] conveyed that he was capable of identifying a particular Swede to whom John talked, I do not agree that he could be accused of saying something false just because either (a) he could not in fact make a canonically identifying reference to any Swede, or (b) John talked only to Swedes he could not canonically identify."
>
> [Klein 1977 p.17]

One might go even further, suppose John talked to Sven whom the speaker can canonically identify, and suppose further that the speaker thinks he talked to Lars whom he can also identify canonically, would these be circumstances under which (5) would be false? Certainly not.

Besides, Kasher & Gabbay's criterion for canonical identifiability, i.e. identifiability by means of a proper name or suitable definite description, does, in our opinion, not guarantee specificity at all, since proper names and definite descriptions themselves can be used non-specifically, as we remarked above.

Kasher & Gabbay claim that ambiguity tests such as conjunction reduction provide evidence for their thesis that sentence like (5) are ambiguous. In our opinion Klein has convincingly refuted this claim. For reasons of space, we will not go into the respectove arguments here, but refer the reader to Kasher & Gabbay 1976, Klein 1977 and Klein 1978.

Contrary to Kasher & Gabbay then, we want to claim that the distinction between specific and non-specific use of terms, at least in sentences not containing intensional verbs, is of a purely pragmatic nature. We also want to claim that it is not a pure dichotomy, but rather, that specificity comes in degrees. We will say something more about sentences containing intensional verbs later on, but let us first consider some examples intended to show how the specific/non-specific distinction can be handled in the framework of epistemic pragmatics as it is sketched in the previous sections.

Let us represent the sentence

(1) A picture is missing from the gallery

by the formula $\exists P(Q)$, in which P and Q stand for the one-place predicates picture and is missing from the gallery respectively. This formula is a wellformed formula of the language defined in the previous section and it can be interpteted according to the rules given there. With respect to the information of a language user x about the denotations of the predicates occurring in the formula, the following situations, among others, are possible.

(a) F(P,x) = {{a,b,c}} and F(Q,x) = {{a}}

According to x there are three pictures: a, b, and c, and there is exactly one object missing: a. This is a characteristic case of specific use of the term a picture in the context is missing from the gallery. The fact that there are more pictures according to x does not prevent him from using the term a picture specifically. In this case, the fact that according to x the only object missing is a helps to use the term a picture to refer specifically to the object a in the context is missing from the gallery. Notice that the pragmatic character of the specific/non-specific distinction is reflected by the fact that it is formulated in terms of the values assigned by the function F to the predicates P and Q with respect to the language user x. These values are not related to the actual denotations of P and Q, and only the actual denotations play a role in the semantics, in the determination of the actual truth value of ∃P(Q).

(b) F(P,x) = {{a,b,c}} and F(Q,x) = {{a},{b},{c}}

Again, there are three pictures according to x: a, b, and c, and there is exactly one object missing, but this time x has no more information than that the missing object is either a, or b, or c, which one of the three, he doesn't know. This is a typical case of non-specific use of the term a picture in the context is missing from the gallery.

(c) F(P,x) = {{a,b,c},{a,b}} and F(Q,x) = {{a}}

In this case x is not sure whether c is a picture, but he is sure that the only object missing is a and that a is a picture. So, no matter the uncertainty about what the pictures are, x uses the term a picture to refer specifically to the object a in the context is missing from the gallery.

(d) F(P,x) = {{a,b,c}} and F(Q,x) = {{a},{a,g}}

This time, there is no doubt about the pictures, but x is not sure whether the object g is missing. But since x is sure that g is not a picture and that a is the only object which is both a picture and missing, x still uses the term a picture to refer specifically to the object a in the context is missing from the gallery. If g would have been a picture too according to x, i.e. if F(P,x) would have been {{a,b,c,g}}, then the term a picture would not have been used specifically by x.

(e) F(P,x) = {{a,b,c}} and F(Q,x) = {{a,b}}

Since there are, according to x, two objects missing which are both pictures, the term a picture cannot be used to refer to a specific object in this case.

(f) F(P,x) = {{a,b,c,d,e,f}} and F(Q,x) = {{a},{b},{c},{d},{e},{f}}

In this situation x uses the term a picture extremely unspecifically. If F(Q,x) would have been {{a},{b},{c},{d}}, then it would have been used less unspecifically. If F(Q,x) would have been {{a}} the term would have been used to refer specifically to the object a in the context is missing from the gallery. This illustrates that specificity comes in degrees, as we remarked earlier.

(g) F(P,x) = {{a}} and F(Q,x) = {{a,b}}

In this case there is just one picture according to x, viz. a, and this picture is one of the missing objects according to x. This is a clear case of specific use of the term a picture, but this time the specificity does not depend on the context is missing from the gallery. Even if F(Q,x) would have been {{b,c}} the term a picture would still have been used specifically. But this, to be sure is a borderline case. If F(P,x) is such that x has the information that there is only one picture and x has the information which object in the domain is this picture, then, no matter the linguistic context, the term a picture is used specifically to refer to that object.

These examples show that, except for borderline cases such as (g), the specific use of an indefinite term not only depends on the information a language user has about the denotation of the predicate corresponding to the common noun phrase in the term, but also on the information he has about the denotation of the predicate corresponding to the verb phrase. This is captured in the definition below. In this definition we use P and Q as variables over common noun phrases such as picture and verb phrases such as is missing from the gallery respectively, and P and Q are the corresponding one-place predicates of our logical language. Further, if a language user utters a sentence a P Q (represented by a formula ∃P(Q)), we say that he uses the term a P in the context Q.

We can now define the notion SPEC(x,a P,Q,c,EM), according to a language user x the term a P refers specifically to the object c in the context Q, in a situation described by an epistemic model EM, as follows:

SPEC(x,a P,Q,c,EM) iff $F(P,x) = \{\{c\}\}$ or $\forall X \in F(P,x), \forall Y \in F(Q,x): X \cap Y = \{c\}$,
where F is the interpretation function of the model EM

According to the definition, a language user uses the term a picture to refer specifically to an object c in the context is missing from the gallery if either he is sure that c is the only picture, or, he is sure that c is the only thing which is both a picture and missing.
Notice that there is a certain connection between the specific use of the term a P in the context Q and the sincere use of the sentence a P Q. A language user x uses the sentence a P Q sincerily iff $V(\exists P(Q),x) = \{1\}$, i.e. if x has the information that $\exists P(Q)$ is true. If x uses the term a P to refer specifically to an object, say c, and if the specificity depends in part on the information x has about the denotation of Q, (i.e. if $F(P,x) \neq \{\{y\}\}$, for all individuals y), then specificity implies sincerity. If x isn't sure that there is at least one object which is both a picture and missing, and that is required for the sincerity of a picture is missing from the gallery, he can't be sure that c is the only object which is both a picture and missing, which is required for the specific use of a P to refer to c. In case the specific use of a P depends only on the information one has about the denotation of P, it is possible to use the sentence a P Q insincerily. If there is exactly one object which is a picture according to x, this does not guarantee that there is at least one object which is both a picture and missing according to him.
One might wonder whether it is ever correct to use an indefinite term specifically. If one can use an indefinite term specifically, isn't it then always also possible to use a definite description or proper name to refer to the individual in question? And if so, wouldn't this disqualify the use of an indefinite term as being not optimal informative? Not necessarily so, it can very well be the case e.g. that the use of the indefinite term a picture is more informative than that of a proper name, e.g. in case according to the speaker the hearer has no idea whatsoever about the referent of that name, or at least does not know that it is the name of a picture. It might also be the case that the speaker has no proper name or definite description for the object in question at his disposal. So, the use of an indefinite term to refer to a specific object is not necessarily incorrect, and will in many cases be the best one can do.

As we remarked above, in our opinion not only indefinite, but also definite terms can be used both specifically and non-specifically. But as we shall see, there is one imporatant difference. Whereas, except for some borderline cases, the specific use of indefinite terms depends also on information about the denotation of predicates outside the indefinite term itself, this is never the case with definite terms.
Let us consider what would be situations in which proper names and universally quantified terms (being the kinds of definite terms that can be represented in the language developed in the previous section) are used specifically and non-specifically. If we use j as the individual constant representing the proper name *John* and P and Q as one-place predicates representing the expressions *man* and *walks* respectively, then the formula $Q(j)$ represents the sentence *John walks* and the formula $\forall P(Q)$ the sentence *every man walks*. Typical (and in fact the only) cases in which the proper name *John* is used specifically by a language user x are cases in which x has no doubts about the denotation of the proper name, e.g. in case $F(j,x) = \{\underline{a}\}$. Analogously, typical (and again in fact the only) cases in which a universally quantified term *every P* is used specifically by x are cases in which x has no doubts about the denotation of P, e.g. in case $F(P,x) = \{\{\underline{a},\underline{b},\underline{c}\}\}$. Notice that in case of these definite terms, certainty about their denotation guarantees specificity of their use. This is not true of indefinite terms (vide example (b) above in which $F(P,x) = \{\{\underline{a},\underline{b},\underline{c}\}\}$ and $F(Q,x) = \{\{\underline{a}\},\{\underline{b}\},\{\underline{c}\}\}$). In this situation, the term *every P* is used to refer specifically to the set of objects $\{\underline{a},\underline{b},\underline{c}\}$. Notice that with respect to these definite terms, certainty about their denotation guaranteed specific use. This is not true of indefinite terms (vide situation (b) above in which $F(P,x) = \{\{\underline{a},\underline{b},\underline{c}\}\}$ and $F(Q,x) = \{\{a\},\{b\},\{c\}\}$). On the other hand, in case of an indefinite term *a P*, uncertainty about the denotation of P does not necessarily prevent it from being used specifically. If x is to refer specifically to an object *c* in the context *Q* with an indefinite term *a P* without being sure about the denotation of P, then two conditions must be satisfied: (i) x must be sure that *c* belongs to the denotation of P; (ii) x must be sure that *c* is the only thing which both belongs to the denotation of P and to the denotation of Q. However, in case of definite terms, the equivalent of the first condition can never be met. If x does not know for sure to which individual the proper name *John* refers, he can't be sure that a certain individual, say *a*, to which he wants to refer specifically is the denotation of that proper name, and that would be the requirement corresponding to condition (i) in case of proper names. And analogously, if x doesn't know for sure to which set

of individuals the predicate P refers, he can't use the term every P to refer specifically to a certain set of objects, say {a,b,c} because he can't be sure that this set is the denotation of P, and that would be the requirement corresponding to condition (i) in case of universally quantified terms.

The conclusion must be that the only kind of specific use of definite terms is the kind of use illustrated by the examples given above. This leads us to the following definition of the notions SPEC(x,b,c,EM), according to a language x the proper name b refers specifically to the object c in a situation described by an epistemic model EM, and SPEC(x,every P,C,EM), according to x the term every P refers specifically to the set of objects C in a situation described by EM (P and b are variables over common noun phrases and proper names respectively, P and b stand for the corresponding predicate and individual constant of the logical language):(7)

SPEC(x,b,c,EM) iff F(b,x) = {c} , where F is the interpretation function of EM

SPEC(x,every P,C,EM) iff F(P,x) = {C} F is the interpretation function of EM

Let us conclude this section by making some remarks about sentences containing an intensional verb such as to owe, which might support the view that sometimes besides the pragmatic specific/non-specific distinction, a similar semantic notion plays a role. It seems reasonable to claim that a sentence like:

(6) Cleon owes Pericles a horse

is ambiguous. One reading would be the one in which Cleon has paid his debt if he has given Pericles a member of the set of horses, another one would be the one in which he has paid his debt only if he has given him a particular member of that set, say Jojo Buitenzorg. The latter one would be semantically specific. It should be noted that the pragmatic specific/non-specific distinction at least also applies to the semantically specific reading. One can use (6) on its specific reading in two different kinds of circumstances: one can have the information which particular horse it is that Cleon owes to Pericles, e.g. one can have the information that Cleon owes Pericles Jojo Buitenzorg. One can also have the information that it is a particular horse which Cleon owes Pericles, but without knowing exactly which one it is, Jojo Buitenzorg, Kees Verkerk or Speedy Volita

Concerning the application of the pragmatic specific/non-specific distinction to the semantically non-specific reading, things are much more unclear. An answer to the question whether the pragmatic distinction applies here as well must await an epistemic pragmatic treatment of higher order languages, since we believe that only these languages provide a suitable framework for a representation of the semantically non-specific reading of sentences such as (6).

4. An ambiguity concerning objects of belief

In this section we will discuss a second application of the framework of epistemic pragmatics. We will be concerned with the two distinct readings of sentences like:

(7) John believes that Peter passed the exam

Under the first reading (7) is true if and only if John believes of a certain person that he passed the exam, and this person is in fact called Peter. On this reading it does not matter whether or not John believes of this person that he is called Peter. On the second reading, (7) is true if and only John believes that a person he believes to be called peter passed the exam. On this reading it does not matter whether Peter in fact passed the exam. The first reading is often called the de re reading, the second the de dicto reading .

One would expect that the framework of epistemic pragmatics, which is intimately tied to the notion of information of language users can handle this ambiguity in a natural way. We will see that this is indeed the case. For this purpose we have to enrich the framework developed sofar with a doxastic operator, representing belief.

To the language given in section 2, we add a set of indexed operators: B_i, B_j, B_{i_0}, B_{i_1}, ..., and the following syntactic clause: if ϕ is a wellformed formula, and B_x an indexed operator, then $B_x(\phi)$ is a wellformed formula. The valuation of formulas of the form $B_x(\phi)$ will be defined in terms of the information x has about ϕ, $B_x(\phi)$ will be assigned the value 1 if and only if according to the information of x, ϕ is true (remember that the phrase 'has the information that' is used without factive implication), and will be assigned the value 0 otherwise.[8] Further, the valuation function will be defined in such a way that a language user y will have the information that $B_x(\phi)$ is true if and only if y has the information that x has the information that ϕ is true. The clause of the recursive definition of the valuation function V for formulas of the form $B_x(\phi)$ is as follows:[9]

(6) $V(B_x(\phi),i^n) = BEL[V(\phi,i^n{}^\frown x)]$, where BEL is a function from S_{n+1} to S_n such that:

$$BEL[\{1\}] = 1;\ BEL[\{0\}] = BEL[\{0,1\}] = 0$$

$$BEL[Y] = \{BEL[y] \mid y \in Y\}, \text{ for } Y \neq \{1\}, \{0\}, \{0,1\}$$

In this definition $i^n{}^\frown x$ stands for the concatenation of i^n and x.

Notice that according to this definition, the operator B_x creates a truly intensional context. $V(B_x(\phi),i^n)$ is not computed from $V(\phi,i^n)$, but from $V(\phi,i^n{}^\frown x)$.

This can be illustrated by the following example:

Let e be the empty sequence, then $V(B_i(P(a)),e) = 1$ iff $BEL[V(P(a),i)] = 1$. This is the case iff $V(P(a),i) = \{1\}$. And this is the case iff every element of $F(a,i)$ is an element of every element of $F(P,i)$.

So, the actual denotations of the constant a and the predicate P do not enter in the evaluation of the truth value of the formula $B_x(P(a))$.
Notice that if we consider the formula $B_j(P(p))$ to be a representation of sentence Peter respectively, and the predicate P the verbphrase passed the exam, we only obtain the *de dicto* of (7). In order to obtain the *de re* reading as well, we have to enrich the syntax of the logical language in such a way that sentence (7) can be represented by another formula as well. This can be done in the following way. A general feature of a language without variables is that the various sentential operators, such as the connectives, tense operators, and in this case the belief operator, do not only operate on sentences, but on predicates too. Another way of saying this is that these expressions operate on predicates with an arbitrary number of places, where sentences are considered to be zero-place predicates. The need for this device can be illustrated by a simple example. In order to construct in a language without variables a formula which corresponds to the following formula of a language with variables: $\exists x(P(x) \& \neg Q(x))$, we need to apply negation as an operation on the predicate Q. In that way we can obtain the formula $\exists P(\neg Q)$. Without this device one could only construct the formula $\neg(\exists P(Q))$, which is the non-variable equivalent of $\neg\exists x(F(x) \& G(x))$. The result of application of negation to a (one-place) predicate δ is straightforward:

$F(\neg\delta,i^n) = COMPL[F(\delta,i^n)]$, where COMPL is a function from A'_{n+1} to A'_{n+1} such that

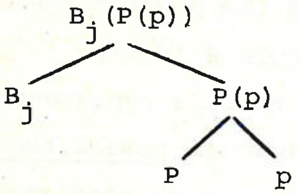
(a) if $X \in A'_i$, then $COMPL[X] = A - X$

(b) if X other wise, then $COMPL[X] = \{COMPL[x] \mid x \in X\}$

Given this definition the formula $\exists P(\neg Q)$ means what $\exists x(F(x) \& \neg G(x))$ usually means.
For a definition of the full apparatus needed for the construction of a language without variables which is equivalent with the language of standard predicate logic, see Quine 1966 and Groenendijk & Stokhof 1978b. In the latter, the epistemic pragmatics of this language enriched with epistemic and modal operators is given in detail.
If we allow the belief operator not only to be applied to formulas, but also

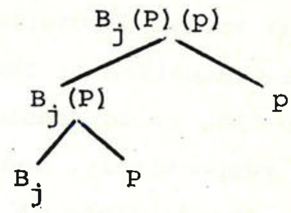
to one-place predicates, by adding a syntactic clause like: if δ is a 1-place predicate and B_x an indexed operator, then $B_x(\delta)$ is a 1-place predicate, then the two readings of sentence (7) can be represented by two different formulas which can be constructed as follows:

(7')



'de dicto'

(7")



'de re'

The interpretation of the syntactic application of an indexed operator B_x to a predicate δ, resulting in the predicate $B_x(\delta)$ is as follows:

$F(B_x(\delta),i^n)$ = INTER[$F(\delta,i^n x)$], where INTER is a function from A'_{n+2} to A'_{n+1} such that

(a) if $X \in A_2$, then INTER[X] = $\bigcap_{x\in X}$

(b) if X otherwise, then INTER[X] = {INTER[x] | $x\in X$}

It will be clear that the operator B_x applied to a predicate δ creates an intensional comtext, since $F(B_x(\delta),i^n)$ is defined not in terms of $F(\delta,i^n)$, but in terms of $F(P,i^n{}^\frown x)$.

What remains to be shown is that the formulas $B_j(P(p))$ and $B_j(P)(p)$ are assigned different truth conditions within this framework:

- $V(B_j(P(p)),e) = 1$ iff
  BEL[$V(P(p),j)$] = 1 iff
  $V(P(p),j) = \{1\}$ iff
  $\{T[x,Y] \mid x\in F(p,j)\ \&\ Y\in F(P,j)\} = \{1\}$, i.e. iff
  every element of $F(p,j)$ is an element of every element of $F(P,j)$, i.e. iff
  every object that according to j could be the denotation of p is an element of every set of objects that according to j could be the denotation of P

- $V(B_j(P)(p),e) = 1$ iff
  $T[F(p,e),F(B_j(P),e)] = 1$ iff
  $F(p,e) \in F(B_j(P),e)$ iff
  $F(p,e) \in$ INTER[$F(P,j)$] iff
  $F(p,e) \in {}_{X\in F(P,j)}$ i.e. iff
  the actual denotation of the proper name p is an element of every set of objects that according to j could be the denotation of P

It will be clear that the truth conditions of these two formulas differ. The formula representing the de dicto reading of sentence (7) does not imply the formula representing the de re reading: if $F(p,e) = \underline{a}$, $F(p,j) = \{\underline{b}\}$, and $F(P,j) = \{\{\underline{b},\underline{c},\underline{d}\}\}$, then the formula $B_j(P(j))$ is true and $B_j(P)(p)$ is false. Neither does the formula representing the de re reading imply the formula representing the de dicto reading: if again $F(p,e) = \underline{a}$ and $F(p,j) = \{\underline{b}\}$, but now $F(P,j) = \{\{\underline{a},\underline{c},\underline{d}\}\}$, then $B_j(P)(p)$ is true and $B_j(P(p))$ is false.

Let us conclude this section, and thereby this paper, by pointing out that a language user can significantly be said to use the proper name Peter specifically (or non-specifically) in uttering sentence (7) only if he uses it on its de re reading. In case he uses it on its de dicto reading the question simply does not arise. In terms of the framework this is expressed by the fact that evaluation of the formula $B_j(P)(p)$ with respect to the information of x, i.e. the computation of $V(B_j(P)(p),x)$, involves considerations concerning the information of x with respect to the denotation of p, i.e. $F(p,x)$, but the evaluation of the formula $B_j(P(p))$ with respect to the information of x does not involve considerations concerning the information of x about the denotation of the constant p representing the proper name Peter.

NOTES

¶ The preparation of this paper was part of the research project 'Performatives and Pragmatics', supported by the Netherlands Organization for the advancement of Pure Research. We would like to than Renate Bartsch, Johan van Benthem and Theo Janssen for their comments on part of the material presented here.

1. See Groenendijk & Stokhof 1978a for empirical and theoretical arguments which justify this assumption.
2. For a formulation of the Gricean maxims in terms of recursively characterized correctness conditions, see Groenendijk & Stokhof 1978b.
3. For definitions see Groenendijk & Stokhof 1978b.
4. See Groenendijk & Stokhof 1978b.
5. For a full treatment of a language with the same expressive power as standard predicate logic, see Quine 1966 and Groenendijk & Stokhof 1978b.
6. See the references in Kasher & Gabbay 1976.
7. A quite similar definition might be given for the specific use of definite descriptions of the form <u>the P</u>.
8. This is, in fact, a simplification. For an adequate semantics of belief, in particular in connection with a semantics for knowledge, it is necessary to distinguish information of language users according to its degree of reliability. Since it plays no essential role in the problems discussed here, we have omitted this complication, but see Groenendijk & Stokhof 1978b, which also contains a critique of the Hintikka-type semantics for the notions of knowledge and belief.
9. A more complex definition which allows also for a 'subjective interpretation' of the index of the operator is:

$$V(B_x(\phi), i^n) = \bigcap_{y \in F(x, i^n)} BEL[\, V(\phi, i^n \frown y)]$$

The reason for omitting this aspect here is that eventually we analyze <u>believe</u> as a two-place predicate, in which case a subjective interpretation of the subject term is automatically allowed for. For details, see Groenendijk & Stokhof 1978b.

REFERENCES

GROENENDIJK, Jeroen & Martin STOKHOF, 1978a, 'Semantics, Pragmatics and the Theory of Meaning', in: Journal of Pragmatics, Vol.II, 1

GROENENDIJK, Jeroen & Martin STOKHOF, 1978b, Epistemic pragmatics. A Logical Analysis of the Role of Information in Language Use, Amsterdam (in preparation)

KASHER, Asa & Dov GABBAY, 1976, 'On the Semantics and Pragmatics of Specific Non-Specific Indefinite Expressions', in: Theoretical Linguistics, Vol.III,1/2

KLEIN, Ewan, 1977, 'Referential Opacity', ms., Amsterdam

KLEIN, Ewan, 1978, On Sentences which report Beliefs, Desires, and other Mental Attitudes, dissertation, University of Cambridge

QUINE, W.V.O., 1966, 'Variables Explained Away', in: QUINE, Selected Logic Papers, New York, Random House

SURFACE COMPOSITIONALITY AND THE SEMANTICS OF MOOD

Roland R. Hausser

0. Introductory remarks

The goal of the present paper is to extend the principles of truth-conditional semantics to non-declarative sentence moods.[1)] My basic hypothesis is that the different syntactic moods should be characterized semantically in terms of their characteristic kind of possible denotation. In the same way as the declarative mood is characterized semantically by the fact that declaratives denote propositions (i.e. functions from points of reference into truth values), I want to characterize the imperative and the interrogative mood by assigning suitable and natural kinds of possible denotations, which are a strictly compositional result of the characteristic syntax defining each mood.

In order to motivate my hypothesis I will outline a syntactico-semantic analysis of the English declarative, imperative, and various kinds of interrogatives in the style of a canonical extension of the grammar defined in Montague's PTQ (Montague 1974, chapter 8). This mode of presentation allows to formulate the specifics of my analysis in brief, yet formally precise terms.

I will rely in particular on the intensional logic of PTQ, which is a type system with λ-calculus, and on the categorial surface syntax of PTQ, which characterizes the meaning of English surface expressions indirectly by systematic translation into formulas of intensional logic. The semantic representations are interpretable in explicit model-theoretic terms according to the truth-conditions of intensional logic as stated in PTQ. In order to handle certain phenomena of context-dependency arising with the imperative and with the interpretation of 'non-redundant answers', I will also refer to the system of a context-dependent intensional logic based on context-variables presented in Hausser (1978).

---

1) The term 'mood' can be used either in connection with <u>verbal mood</u>, such as indicative, subjunctive, optative, etc., or in connection with <u>syntactic mood</u>, such as declarative, imperative, or interrogative. The specifics of a certain mood may vary from language to language, and different languages may have different moods, as a comparison of the verbal moods of Greek vs. English demonstrates. The present paper will be concerned with syntactic mood only.

Unfortunately, in PTQ a declarative sentence is defined as denoting a truth-value rather than a proposition (i.e. a function from points of reference to truth-values). I agree with Tichy (1971) and Maderna (1974),who have critized such 'relicts of extensionalism' in Montague's semantics.Rather than redefining PTQ in this respect (which would be space consuming, but technically routine), I will gratuitously intensionalize translation-formulas where it is semantically relevant.

1. Syntactic mood versus speech acts

When we talk about syntactic moods we are talking about formal properties of linguistic surface expressions. These are to be kept clearly distinct from (the properties of) the speech acts in which a linguistic expression may function.
For example, (1)
1) Could you pass the salt?
is by virtue of its form an expression of the interrogative mood. If we use (1) at the dinner table we are normally performing a speech act of a request. But we may use (1) also as a question, for example when we use (1) to ask a disabled person about his physical abilities. The difference between these two types of speech acts may be characterized in terms of different appropriate responses: if (1) is used as a request, the appropriate response would be passing the salt; if (1) is used as a question about the hearers state of recovery, on the other hand, a simple 'yes' or 'no' would be an appropriate response. The crucial point is, however, that the literal meaning of (1) is exactly the same in the two different speech acts indicated.

In the same sense as a linguistic expression like (1) may be used in different types of speech-acts, different linguistic expressions may be used to perform the same type of speech act. Consider for example (2) and (3):
2) You know where the can opener is. (declarative)
3) Get the can opener! (imperative)
I take it that the locutionary acts indicated in (4) and (5)
4) John says to Mary at 10.45: 'You know where the can opener is.'
5) John says to Mary at 10.45: 'Get the can opener!'
may constitute the same illocutionary acts, namely (6):
(6) John requests from Mary at 10.45 to give him the can opener.
In other words, (2) and (3) may be used in locutionary acts which are equivalent on the illocutionary level, despite the fact that (2) and (3) are syntactically of different mood and denote semantically different things: while (2) denotes a proposition, (3) does not.

There is a certain feeling that declarative sentences are used primarily or predominantly as statements, while imperatives and interrogative sentences are used mostly as requests or questions, respectively. The above examples demonstrate clearly, however, that the syntactic mood does not determine the speech act. Rather, syntactic mood participates with all the other linguistic properties of a given surface expression ϕ in delimiting the set of use-conditions of ϕ. Since there is no one to one relation between syntactic moods and speech acts, it would be a mistake to implement speech act properties in the semantic characterization of syntactic mood.

The conclusion stands in contrast to the practice of Austin (1959), Lewis (1972) and many others, who proceed in their analyses on the assumption that examples like (7) and (8) are semantically equivalent.

7) I order you to leave. (declarative)

8) Leave! (imperative)

Before we scrutinize this assumption, let us note that (7) and (8) are of quite different syntactic structures. Comparison of (7) and (9), furthermore,

9) You ordered me to leave.

shows clearly that (7) and (9) are in the same syntactic paradigm.

It is curious that Austin and Lewis draw quite contrary conclusions from the presumed equivalence. Lewis recognizes that (7) is a declarative sentence and therefore wants to treat it semantically as a proposition. The presumption of semantic equivalence between (7) and (8), however, leads him to the counterintuitive consequence that (8) likewise denotes a proposition. Austin, on the other hand, notes correctly that (8) does not denote a function from points of reference into truth values (proposition). The presumption of semantic equivalence between (7) and (8), however, leads Austin to the counterintuitive conclusion (and contrary to Lewis) that (7) does not denote a proposition, which means breaking the linguistic paradigm (according to Austin, (9) denotes a proposition while (7) does not).

Austin and Lewis agree that (7) and (8) are semantically equivalent and disagree about whether (7) (the explicit performative sentence) should denote a truth value or not. I take the position that (7) is a normal declarative sentence and denotes a proposition, while (8) is an imperative and does not denote a proposition. Consequently, (7) and (8) cannot be semantically equivalent in the sense that they have the same denotation at any given point of reference: if (7) and (8) have different types of possible denotations they cannot be semantically equivalent.

The assumption of semantic equivalence between, e.g., (7) and (8) must go. Instead the relation between (7) and (8) is one of overlapping use conditions, just as in the case of (2) and (3) above. Furthermore, it is easy to show that an explicit performative sentence like (7) denotes a proposition (pace Austin). Imagine a run down Hobo who walks by the docked Q.E. II, throws an empty champagne bottle at her hulk and says:

10) I hereby christen this ship Q.E. II.

If we were to report this incident by saying (11),

11) That hobo christened this ship the Q.E. II.

(11) would be false under the indicated circumstances. Thus the so-called happiness conditions for an act of christening must be part of the truth conditions of the performative verb christen. (11) is false because the happiness conditions for christening are not satisfied in the 'ceremony' referred to. And (10) would be as false as (11). This truth value assignment reflects that in the indicated circumstances no christening has taken place in the moment of pronouncing (10).

The alleged semantic equivalence between, e.g. (7) and (8)

7) I order you to leave.

8) Leave!

has served as the corner stone of the so-called performative analysis of non-declaratives. Lewis, for example, treats (12b), (13b), and (14b) as paraphrases of (12a), (13a) and (14a), respectively, and proposes to derive the latter from the former via meaning-preserving transformations.

12a) I state that you are late.	12b) You are late.
13a) I command that you are late.	13b) Be late!
14a) I ask whether you are late.	14b) Are you late?

Lewis says:

> "I propose that these non-declaratives (i.e. (13b) and (14b), R.H.) ought to be treated as paraphrases of the corresponding performatives, having the same base structure, meaning, intension, and truth value at an index or an occasion. And I propose that there is no difference in kind between the meanings of these performatives and non-declaratives and the meanings of the ordinary declarative sentences considered previously."
> Lewis 1972, p. 208

It seems to me, however, that by deriving (12b), (13b) and (14b) from the indicated declaratives, Lewis relies not only on the unsubstantiated semantic equivalence discussed above, but also fails to respect the distinction between syntactic mood and speech acts. The problems resulting are easy to see: if the surface syntactic form of declaratives, imperatives, and interrogatives is systematically related to the 'underlying' performative clauses 'I state', 'I command' and 'I ask', respectively, then also examples like (15a,b) or (16a,b) are rendered as paraphrases:

15a) I command that you feel yourself at home.
15b) Feel yourself at home!
16a) I hereby ask you whether you could pass the salt.
16b) Could you pass the salt?

The desire to reduce different syntactic moods to the same kind of meaning is also apparent in the analysis of Stenius (1967), according to which examples like (12b), (13b) and (14b)
12b) You are late.
13b) Be late!
14b) Are you late?
share the same sentence radical (propositional content) and differ only in their underlying mood operators. However, as Lewis points out correctly, 'it is hard to see how it (i.e. the method of sentence radicals, R.H.) could be applied to other sorts of questions ...' (op.cit., p. 207). Compare for example (17) and (18):
17) Who came?
18) John came.
There is no complete propositional content which could serve as the basis of (17).

The only reason for the popularity of the performative approach and the sentence radical method I can see, is that it derives different moods from the same kind of structure. This way one can continue to practice logic in the ways one has long been accustomed to and evades the task to provide the syntactic structure of each mood with its own characteristic semantics. While I believe that the principles of referential semantics should be rigorously maintained, there is no reason why we should not use these principles creatively. The performative approach and the sentence radical approach proceed as if <u>propositions</u> were the only interpretable expressions of logic, which is a completely unwarranted assumption.

Consider for example (19) and (20):
19) Stop to kiss Mary.
19') $\hat{}$stop'(kiss'$_*$(m))
20) John stops to kiss Mary.
20') $\hat{}$stop($\hat{}$j, $\hat{y}$ kiss'(y, $\hat{P}$P{$\hat{}$m}))

The denotation of a complex IV-phrase like (19) can be interpreted in exactly the same compositional manner (relative to a model and a point of reference) as a complete declarative sentence like (20). The only difference is that (19) denotes a function from points of reference into functions from individuals into truth values, such as indicated in (a):

a) $I \times J \rightarrow (A \rightarrow \{0,1\})$
while (20) denotes a function from points of reference into truth values, such as indicated in (b):
b) $I \times J \rightarrow \{0,1\}$

The meaning of (19)/(19') is characterized in terms of the same kind of truth-conditions as the meaning of (20)/(20'). (It would perhaps be more appropriate to speak of denotation-conditions rather than truth-conditions, since the rules in question characterize the meaning of expressions of any semantic type in general model theoretic terms.)

Propositions are like other possible denotations in that they occur as the denotation of subordinate clauses. So why should not other types of possible denotation be like propositions in that they may serve as the denotation of complete linguistic expressions? I will present a treatment of interrogatives (section 5), according to which (21) translates into (21'):

(21) Who talks?

(21') $^\wedge \underline{P}\,[\,\underline{P}\{^\wedge talks'\}\,] \in ME_{<s,f(t/T)>}$
(where $\underline{P}$ is a variable of type $<s,f(T)>$

(21') is treated as a complete expression of intensional logic, denoting a function from points of reference into sets of term-denotations (John', Bill', the man you saw yesterday', etc.) According to my view, this type of possible denotation is as characteristic of the type of interrogative involved as propositions are the characteristic type of possible denotations for declaratives. As far as I can see, the indicated semantic characterization of the meaning of (21') does in no way exceed the basic principles of intensional logic.

The point I want to make here is that it is logic-wise conceivable that possible denotations other than propositions may serve as denotation of complete linguistic expressions, such as, for example, interrogatives, or imperatives. Note in this connection that a type system (such as the one employed the intensional logic defined in PTQ) provides infinitely many different types of possible denotations (compare the recursive definition of types, p. 256, and the recursive definition of possible denotations, p. 258, in PTQ). Later (section 4, 5) I will discuss the further question of which types of possible denotations should be defined as characteristic of which kinds of syntactic mood. This decision will be motivated in linguistic terms, since we proceed on the hypothesis that different syntactic moods are different kinds of syntactico-semantic composition occurring in natural language.

2. Linguistic standards of analysis (surface compositionality)

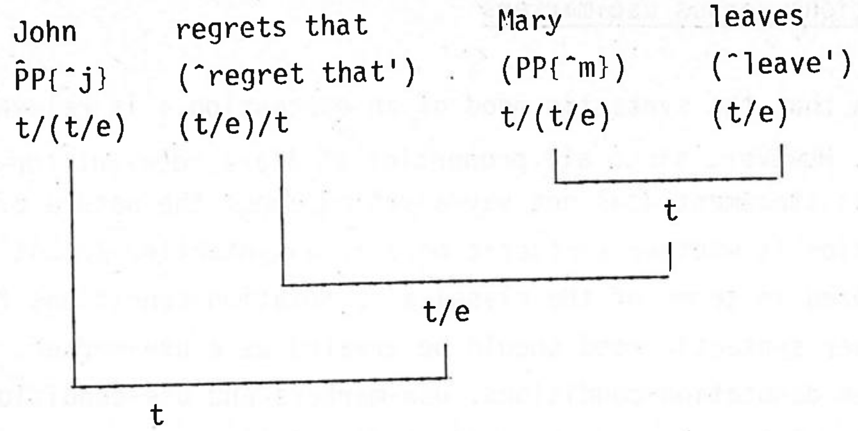
One reason why I have chosen to present my analysis of syntactic mood in form of an extension of PTQ is that PTQ is a complete grammar in the sense that the generation and interpretation of a fragment of English is coordinated in a rigorously formal generative system. Lewis (1972), on the other hand, is incomplete because the transformations which supposedly make an explicit performative categorial structure into a (non-)declarative surface structure are not spelled out. I take it, however, that it would be a rather nasty job to explicitly define these transformations and give a 'reasonable' linguistic motivation for them (though I have no doubts that it could be done).

As long as we do not coordinate the surface syntactic analysis with the semantic characterization of an expression in terms of a complete, formal generative grammar generating a fragment if a natural language, we can postulate any kind of logic for the characterization of syntactic mood.
The confusion of semantic properties and speech act features shows, furthermore, that in addition to the methodological standard of completeness we need some kind of linguistic standard to guide our use of mathematical power to linguistically well-motivated analyses. But which standard of linguistic analysis should we adhere to?

The assumption that the semantics of natural languages works like the semantics of formal languages in that the meaning of complex expressions is the systematic result of the meaning of the basic parts (and the mode of syntactic combination) suggests a principle which I would like to call the principle of surface compositionality. According to this principle, the semantic representation of a linguistic expression should contain nothing that does not have concrete surface motivation. Furthermore, a surface compositional analysis must characterize explicitly how the meaning of a complex surface expression is composed from the meaning of its basic surface constituents. For example, an analysis which derives passive via transformation from the corresponding active violates the principle of surface compositionality though it might still be complete in the sense that the linguistic expressions investigated are systematically characterized in model-theoretic terms by means of formal translation into intensional logic (compare Partee (1975), Cooper (1975)).

Thanks to the presence of λ-calculus, the principle of surface compositionality can be easily accomodated in a PTQ-style grammar. Consider for example (23):

23)



The semantic representation in (23) assigns the standard logical translations to the surface constituents - apart from the unresolved analysis of <u>regret that</u>, which would require further attention. Since the surface syntax and the logical syntax are completely parallel in their respective function-argument structure, the composition of the meaning in the surface expressions is explicitly characterized. Yet our surface compositional analysis remains strictly within the realm of referential semantics. The formula in (23) reduces 'automatically' via λ-conversion into the more standard looking but equivalent formula (24):

24) ^regret that'(^j, ^leave'(^m))

Neither Lewis (1972) nor Hintikka (1976) follow the principle of surface compositionality because their semantic representations contain a lot more than can be surface-syntactically motivated, and because the structure of the surface syntax is completely disregarded in their respective semantic analyses. The movement, deletion, and insertion transformations of transformational grammar in general run counter to the principle.

The linguistic literature of recent years has been quite concerned with 'evaluation measures', designed to help finding the 'right' grammar for a natural language (c.f. Chomsky (1965), Partee (1978)). It seems to me that the combination of

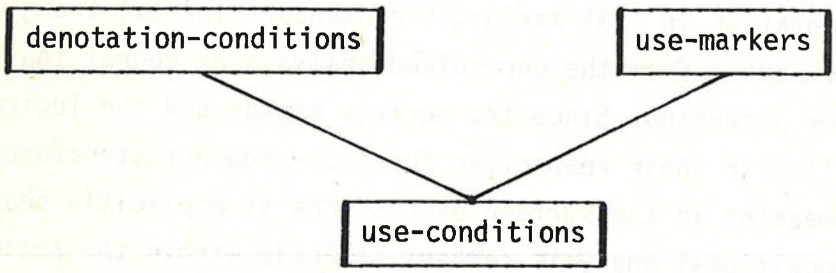
a) the principle of formal completeness (regarding the coordination of syntax and semantics), and
b) the principle of surface compositionality

would result in a restriction on linguistic analysis of natural language which would induce the following desirable properties: an approach conforming to (a) and (b) would have to be such that (i) different 'local' analyses in the system must really be shown to be <u>compatible</u> with each other, and (ii) all aspects of linguistic analysis would have to be motivated over concrete surface properties (rather than putative speech act and/or paraphrase properties).

3. Denotation conditions versus use-markers

There is no question that the syntactic mood of an expression ϕ is relevant for the use-conditions of ϕ. However, since all properties of ϕ are relevant for the use-conditions of ϕ, this statement does not say anything about the nature of syntactic mood. The real question is whether syntactic mood is a syntactico-semantic phenomenon to be characterized in terms of the classical denotation-conditions (truth conditions) or whether syntactic mood should be treated as a use-marker. The interrelation between denotation-conditions, use-markers and use-conditions of an expression of natural language may be indicated as follows:

25)



The position that mood is something like a use-marker is quite clearly expressed in the following quotation from Lewis (1972):

> "Fundamentally, however, the entire apparatus of referential semantics (whether done on a categorial base, as I propose, or otherwise) pertains to sentence radicals and constituents thereof. The semantics of mood is something entirely different. It consists of rules of language use... In abstract semantics, as distinct from the theory of language use, a meaning for a sentence should simply be a pair of mood and an S-meaning (mood being identified with some arbitrarily chosen entities)."
>
> Lewis (1972), p. 207

Yet, there is no one to one relation between mood and use/speech-act (compare section 1). Furhtermore, to identify mood with arbitrarily chosen entities violates the principle of surface compositionality (c.f. section 2). Finally, why should referential semantics be limited to sentence radicals? As Lewis points out himself, sentence radicals are a theoretical construct and do not normally appear on the linguistic surface.

4. A compositional analysis of imperatives

The crucial question for a surface compositional analysis of non-declarative sentence moods is: What do complete but non-declarative expressions denote? I take it that an imperative denotes a property (roughly that property which the speaker wants the hearer to acquire). This concept is formally captured in the semantic representation (33') of (33):

33) Leave! $\in P_{IV}$

33') $\hat{}\lambda x\,[\Gamma_2\{x\} \wedge \text{leave}'(x)] \in ME_{<s,f(IV)>}$

(33') denotes the property of being the hearer (Γ_2) and to be leaving. Γ_2 is a context-variable representing the property of being the hearer. For detailed definitions of the presumed treatment of context-dependency see Hausser (1978), where Γ_2 appears in the translation of second person pronouns such as you ($\in B_T$), your ($\in B_{T/CN}$), yours ($\in B_T$), our ($\in B_{T/CN}$), etc. Γ_2 is of type $<s,f(CN)>$.

The indicated treatment of imperatives may be implemented into the PTQ-system by adding the following two rules:

S20. If $\alpha \in P_{IV}$, then $F_{20}(\alpha) \in P_{IV}$ and $F_{20}(\alpha)$ is α!

T20. If $\alpha \in P_{IV}$ and α translates into α', then $F_{20}(\alpha)$ translates into $\hat{x}\,[\Gamma_2\{x\} \wedge \Gamma'(x)]$.

Thus, (33) and (33') are derived in our extension as follows:

34) leave! $\in P_{IV}$
  | S20
leave $\in B_{IV}$

35) $\hat{x}\,[\Gamma_2\{x\} \wedge \text{leave}'(x)]$
  | T20
leave'

Our analysis may be motivated linguistically as follows:

It seems to be a universal linguistic fact that imperatives do not have subjects. Therefore, it seemed best to avoid the postulation of an 'underlying' subject. Since the syntactic structure of English imperatives is identical to the structure of un-inflected IV-phrases, it suggests itself to take such IV-phrases as the point of departure for the derivation of imperatives. The presence of the context-variable Γ_2 in the semantic representation of imperatives is motivated over the universal fact that second person personal T-pronouns serving as the direct or indirect object of the main verb in an imperative must be reflexive pronouns. With regard to this important detail, our treatment of imperatives plugs in directly to the surface compositional treatment of pronouns presented in Hausser (1978) (for the analysis of reflexives see section 8). The context-variable Γ_2 is crucial, furthermore, for distinguishing imperatives semantically from the corresponding IV-phrases (which are not considered to be complete linguistic expressions). The 'propositional content' of imperatives, finally is captured in the semantic representation without the counterintuitive assumption that imperatives denote propositions and without invoking putative speech-act properties.

The fulfillment-conditions for imperatives used as requests may be roughly characterized as in (36):

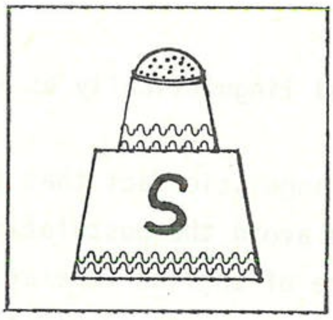
36) If the speaker s utters $f_{20}(\alpha)$ (where $\alpha \in P_{IV}$) towards a hearer h at time j in order to make a request, then this utterance is a fulfilled request if and only if there is a time j', j' > j, such that $[h=x \wedge \alpha(x)]$ is true at j'.

Whether a certain imperative expression is used as a request or an order, etc. depends on pragmatic criteria concerning the status of the speaker (authority), his wishes (sincerity), etc. But assuming that an utterance of an imperative constitutes a request, (36) specifies the conditions under which this request is fulfilled.

Our semantic representation of the imperative (33) may seem rather parsimonious. Is (33') sufficient as the semantic representation of the literal meaning of (33)? And what is the literal meaning? In order to clarify my position regarding the relation of meaning and use, semantics and pragmatics, it might be helpful to consider a metaphor.

Imagine A and B sitting at the dinner table, B looks quietly at A and flashes the little card indicated in (37).

37)



Would not A quite normally understand the flashing of this card under the indicated circumstances as a request to pass the salt? No illocutionary force indicator is necessary to use the card in (37) for making a request. Furthermore, we could use the same card also in a descriptive function if the circumstances are of an appropriate nature. For example, (37) could serve as a label on a box containing little bags with salt (in distinction to another box containing little bags with sugar).

Now, does (37) have two or more meanings depending on how many different uses we can find for it? Or is (37) one sign with one meaning? Though (37) has an extremely rudimentary 'surface-syntax' and there is no special syntactic mood to speak of we may extend the performative approach to this example by postulating 'underlying' meanings.

> If (37) is used as a request then it means 'I request from you that you pass the salt.'. But if (37) is used as a label then it means 'This container contains salt.'.

I would like to argue, on the other hand, that (37) has only one simple meaning, as represented in the picture of a salt-shaker. We may say that this picture denotes in the same sense as the English word salt-shaker.

It seems to me that the study of speech-acts makes sense only if we operate with concrete linguistic signs in their literal meaning. Only if we stick to the meaning concretely given in the sign can we study the use-conditions of this sign. By making explicit assumptions about the believes, intentions, emotions, and preferences of the speaker and other relevant features of the utterance situation, we can systematically deduce the intended use of the sign. The theory of pragmatics should be able to account for different uses of cards like (37) in the same way as it should account for the use of natural language signs.

In order to further motivate my surface compositional approach (with its consequences for syntax, semantics, and pragmatics), I will present below an analysis of the interrogative mood which illustrates how 'literal' the translation of linguistic surface structures into the standard notation of intensional logic can be (though much could be improved in a more thorough and detailed analysis).

5. The syntax and semantics of interrogatives[1)]

While the imperative mood in English is characterized by one basic type of surface syntactic construction, the interrogative mood exhibits a systematic variety which creates a whole spectrum of different surface syntactic constructions. For reasons which will become apparent shortly, we may distinguish between the following two basic types of interrogatives: non-restricted interrogatives and restricted interrogatives.

---

1) The discussion of interrogatives in the present section summarizes the main points of Hausser (1977, chapter 5).

Examples of non-restricted interrogatives are (38-42):

38) Who came? (John.)

39) Why did John come? (Because he admires Mary.)

40) When did John arrive? (Early.)

41) How often did John kiss Mary? (A 103 times.)

Examples of restricted interrogatives, on the other hand, are (43) and (44):

43) Did John kiss Mary? (Yes./No.)

44) Does Mary love John or Bill? (Yes./No./John./Bill.)

The expressions in brackets in the above examples are compatible answer expressions. In the same sense as we have to distinguish between an interrogative X and a speech act of asking which contains X as a token, we have to distinguish between 'answer-expressions' and the speech act of answering. Which kind of expressions may serve as answers to which kind of interrogatives? Consider the following example:

45) John asks Bill: "Who came?"

Bill answers:

a) "Peter." (non-redundant answer)

b) "Peter did." (pseudo-redundant answer)

c) "Peter came." (redundant answer)

d) "Peter, on Mary's bicycle." (explicatory non-redundant answer)

e) "Peter did, on Mary's bicycle." (explicatory pseudo-redundant answer)

f) "Peter came, on Mary's bicycle." (explicatory redundant answer)

The non-redundant and pseudo-redundant answers (i.e. (45a), (45b), (45d) and (45e) are special surface expressions which (i) are not a complete declarative expression, (ii) denote a truth-value (proposition) if interpreted relative to a suitable interrogative, and (iii) exhibit highly specific structural properties which delimit the class of interrogatives they may function to answer. The redundant answers (i.e. (45c) and (45f), on the other hand, are complete and regular declarative expressions. Whether these declaratives are used as answers or not is solely a matter of the speech-act they are used in.

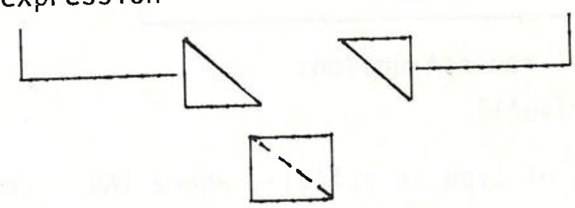
The answers indicated in (38-44) are all non-redundant answers. They demonstrate that a non-redundant answer-expression may be related only to two kinds of interrogative because it is structurally incompatible with all other kinds. The expression "John.",for example, is a possible answer only to (38) and (44), on the 'disjunctive reading' of (44). The declarative sentence (46),

46) John came early because he admires Mary.

on the other hand, may serve as an explicatory redundant answer to (38), (39), and (40). Which interrogative (46) is used to answer is indicated only in the intonation. Note that this specific intonation is a use-marker: it has no influence on the denotation of the sentence.

How can the structural properties of interrogatives, answer-expressions, and the important property of compatibility observed in the above examples be formally characterized in a complete and surface compositional analysis? In the analysis presented in Hausser & Zaefferer (1977) and Hausser (1977, chapter 5), an interrogative is analyzed syntactico-semantically in such a way that it fits together with a compatible non-redundant answer and vice versa. Their combination results in a proposition which is semantically equivalent to the translation of the corresponding redundant answer:

48) interrogative expression — non-redundant answer



redundant answer

Note that the interrogative elements <u>who</u>, <u>why</u>, <u>when</u> and <u>how</u> in the non-restricted interrogatives (38-41) have the <u>same</u> syntactic category as the <u>constituents</u> of their respective non-redundant answers. This fact is taken into account in the analysis of the following example:

49) Who came? ($\varepsilon P_{t/T}$) John. (εP_t)

$\hat{}\lambda \underline{P}_0 [\underline{P}_0\{\hat{}come'\}]$ $\Gamma^{t/T}(\hat{}\lambda P [P\{\hat{}j\}])$

where $\underline{P}_0$ is a variable of type <s,f(T)> ranging over term-denotations, while $\Gamma^{t/T}$ is a context-variable of type <s,f(t/T)> ranging over translations of t/T-interrogatives provided by the context.

The non-redundant answer (49b) is interpreted with respect to the interrogative (49a) by replacing the context-variable by the translation of (49a). The result is (50a), which reduces via λ-conversion to (50b).

50a) $\hat{}\lambda \underline{P}_0 [\underline{P}_0\{\hat{}come'\}](\hat{}\lambda P [P\{\hat{}j\}])$

50b) $\hat{}come'(\hat{}j)$

(50b) is the translation of the corresponding redundant answer "John came", as desired. For the details of <u>contextual reconstruction</u> (i.e. replacement of a context-variable as indicated in (49)) see Hausser (1978).

According to the indicated treatment, an interrogative denotes a function from points of reference into sets of corresponding non-redundant answer constituent denotations. The category variation (type-variation) of different interrogatives (i.e. interrogatives with interrogative elements of different category) is crucial for the characterization of compatibility between an interrogative and its non-redundant answers. As an example of a t/IAV-interrogative consider (51):

51) When did John arrive? ($\varepsilon\ P_{t/IAV}$) Early. ($\varepsilon\ P_t$)

$\hat{}\lambda a_n^t\ [\hat{P}P\{\hat{}j\}(\hat{}a_n^t\{\hat{}arrive'\})\ ]$ $\Gamma^{t/IAV}(\hat{}early')$

result of contextual reconstruction:

$\hat{P}P\{\hat{}j\}(\hat{}early'(\hat{}arrive'))$

(the variable a_n^t is of type <s,f(IAV)>, where IAV = IV/IV, ranging over the denotation of time-adverbs. In order to differentiate between different kinds of adverbs, I assume that the variables representing IAV-interrogative elements are sorted with respect to time, place, manner, etc.)

Comparison of (50) and (51) illustrates that the compatibility between an interrogative and a non-redundant answer expression depends largely on the category correspondence between the interrogative and the context-variable in the answer (where the type of the context-variable must always be such that the answer denotes a proposition). Further restrictions are induced by sub-categories such as different kinds of adverbs:

52) * Where did John go? Yesterday.

Where corresponds to an IAV-variable sorted with respect to places. Though the constituent in the answer-expression in (52) is of category IAV, it is not of the right subcategory.

Our semantic treatment covers also the multiple interrogatives (which are likewise non-restricted interrogatives). Consider (53):

53) Who kissed whom where? ($\varepsilon\ P_{(((t/T)/T)/IAV)}$)

$\lambda a_2^p \lambda \underline{P}_1 \lambda \underline{P}_0\ \ [\underline{P}_0(\hat{}a_1^p(\hat{}see'(\underline{P}_1)))\ ]$

John Mary in the kitchen. ($\varepsilon\ P_t$)

$\Gamma^{t/T/T/IAV}(\hat{}in\ the\ kitchen')(\hat{P}P\{\hat{}m\})(\hat{P}P\{\hat{}j\})$

$= \Gamma^{t/T/T/IAV}(\hat{P}P\{\hat{}j\},\ \hat{P}P\{\hat{}m\}\ \hat{}in\ the\ kitchen')$

where a_2^p is a variable of type <s,f(IAV)> sorted with respect to places. $\underline{P}_0$ and $\underline{P}_1$ are variables of type <s,f(T)>.

Contextual reconstruction renders again a formula which is equivalent to the translation of the corresponding redundant answer:

John kissed Mary in the kitchen.

$\hat{P}P\{{}^\wedge j\}({}^\wedge in\ the\ kitchen'({}^\wedge kiss(\hat{P}P\{{}^\wedge m\})$

(53) demonstrates that the category of interrogative elements as well as their number and relative order determine the category of the whole interrogative expression. Furthermore, the number of interrogative elements in a non-restricted interrogative is theoretically indefinitely large and restricted only by limitations of perception (a parallel case is the stacking of relative clauses). This means that there are indefinitely many kinds of non-restricted interrogatives, having different types of possible denotations and different types of compatible non-redundant and pseudo-redundant answers. Semantically, we may motivate the postulated syntactico-semantic structure of interrogatives by saying that interrogatives denote <u>sets of</u> their non-redundant <u>answer constituent-denotations</u>.

We have captured the intuition underlying the hypothesis in Hamblin (1973) that interrogatives denote <u>sets of possible answers</u>, though in a manner quite different from Hamblin's. Hamblin proceeds on the assumption that an answer must be a declarative sentence, which leads him to define interrogatives as uniformely denoting <u>sets of propositions</u>. He motivates this move by proposing to treat declaratives as denoting unit-sets of propositions. Thus imperatives and declaratives denote the same kind of function - though imperatives are apparently not included in this 'generalization'. Needless to say, the question of compatibility between interrogatives and suitable answer-expressions is not raised in this paper.

The basic principles of my approach to non-restricted interrogatives are suited also for a logically standard treatment of restricted interrogatives. Consider (54):

54) Did John leave?

$${}^\wedge\lambda r_n\,[\,r_n({}^\wedge leave'({}^\wedge j)) \wedge [\,(r_n = \hat{p}\,[\,{}^\vee p\,]\,) \vee (r_n = \hat{p}\,[\,{}^\vee\!\sim p\,]\,)]]$$

No.

$$\Gamma^{t/(t//t)}(\hat{p}\,[\,{}^\vee\!\sim p\,]$$

(r_n is a variable of type $ME_{f(t//t)}$. I assume that <u>yes</u> and <u>no</u> ($\epsilon\ B_{t//t}$) translate into $\hat{p}\,[\,{}^\vee p\,]$ and $\hat{p}\,[\,{}^\vee\!\sim p\,]$, respectively. Contextual reconstruction renders an expression which is equivalent to the translation of the corresponding redundant answer 'John did not leave', i.e. <u>${}^\wedge\!\sim leave'({}^\wedge j)$</u>.)

Yes/no-interrogatives like (54) are morphologically and semantically related to the so-called alternative interrogatives, such as (55),

55) Did John kiss Mary or Sally?

which is ambiguous between the following two meanings:

55a) $\hat{r}_n\,[\,r_n(\text{kiss}'({}^\wedge j,\ \hat{P}\,[\,P\{{}^\wedge m\} \vee P\{{}^\wedge s\}\,])\,)\,]$ (yes/no-reading)

55b) $\hat{\underline{P}}_n\,[\,\text{kiss}'({}^\wedge j,\ \underline{P}_n) \wedge [\,(\underline{P}_n = \hat{P}P\{{}^\wedge m\}) \vee (\underline{P}_n = \hat{p}p\{{}^\wedge s\})\,]\,]$ (term-reading)

Expressions like (54) or (55) are called restricted interrogatives because the possible values of the variables in their translation is restricted via explicit enumeration. Note that there are again indefinitely many kinds of restricted interrogatives, as shown by example (56):

56) Did John kiss Mary or Bill Suzy? ($\varepsilon\ P_{t/T/T}$) John Mary.

The different readings (55a) and (55b) are marked on the surface in terms of different intonation. These intonation patterns constitute an instance of 'syntactic intonation' (are not a use-marker). The rules of implementing the indicated syntactico-semantic treatment of restricted and non-restricted interrogatives, redundant and non-redundant answer-expressions as well as the so-called 'indirect questions' into a complete and canonical extension of PTQ are given in Hausser (1977, chapter 5), and need not be repeated here.

6. Conclusions

Let us return now to our basic hypothesis according to which different moods should be characterized semantically in terms of their characteristic kind of possible denotation (induced by their characteristic syntax). We have seen that different moods arise in form of very clearly defined different kinds of surface structures. We argued that

a) declaratives denote propositions, i.e. they are expressions of type $<s,t>$,

b) imperatives denote porperties of individual concepts (restricted in the indicated way), i.e. they are expressions of type $<s,f(t/e)>$,

c) interrogatives denote functions from s into sets of non-redundant answer constituent denotations, i.e. they are expressions of type $<s,<<s,\alpha>,t>>$, where α is the type of the constituent of the respective compatible non-redundant answer expressions.

Thus there is no type-overlap between different syntactic moods. It seems to me that the kinds of possible denotations assigned to different moods are fairly natural. To say that an imperative denotes properties of individual concepts (restricted to the hearer) is no less motivated than saying that declaratives denote propositions.

While the presented treatment of non-declaratives is logic-wise completely standard, it is linguistically close to the surface and maintains the original categories of the constituents. In line with the principle of surface compositionality, our semantic representations of different moods contain nothing that is not concretely reflected in the linguistic surface.

Our analysis of mood has shown that the principle of surface-compositionality leads not only to a highly restricted syntactico-semantic analysis of expressions of natural language, but is also of greatest consequence for the analysis of the use-conditions of these expressions. The systematic study of speech-acts cannot be successful as long as we do not proceed on the basis of concrete linguistic signs in their literal meaning.

REFERENCES

Austin, J.L. (1962), How to do things with words, Cambridge, Mass., Harvard University Press.

Chomsky, N. (1965), Aspects of a theory of syntax, Cambridge, Mass., MIT-press.

Cooper, R. (1975), Montague's semantic theory and transformational syntax, unpublished Ph.D. dissertation, University of Massachusetts.

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

Hausser, R. (1974), Quantification in an extended Montague grammar, unpublished Ph.D. dissertation, University of Texas at Austin.

Hausser, R. (1976a), 'Presuppositions in Montague grammar', in Theoretical Linguistics, Vol. 3.3.

Hausser, R. (1976b), 'Scope ambiguity and scope restriction in Montague grammar', in Amsterdam Papers in formal grammar, Vol. 1 (eds. R. Bartsch, J. Groenendijk, M. Stokhof, Department of Philosophy, University of Amsterdam).

Hausser, R. (1977), Linguistic cross-connections, ms.

Hausser, R. (1978), 'How do pronouns denote?', in Syntax and semantics, vol. 10, eds. F. Heny and H. Schnelle, Academic Press, New York.

Hausser, R. and D. Zaefferer (1977), 'Questions and answers in a context-dependent Montague grammar', to appear in Formal semantics and pragmatics of natural language, eds. F. Guenthner and S.J. Schmidt, Synthese Language Library, Reidel, Dordrecht.

Hintikka (1976), The semantics of questions and the questions of semantics, Acta Philosophica Fennica, vol. 28, no. 4, North-Holland Publishing Company, Amsterdam.

Lewis, D.K. (1972), 'General semantics', in Semantics of natural language, eds. Davidson and Harman, Reidel, Dordrecht.

Maderna, P. (1974), 'Some remarks on Montague's language L_0 in comparison with Tichy's language L_μ', in Linguistische Berichte 34/74.

Montague, R. (1974), Formal philosophy, The papers of Richard Montague, edited and with an introduction by Richmond Thomason, Yale University Press, New Haven.

Partee, B. (1975), 'Montague grammar and transformational grammar', Linguistic Inquiry, IV:2.

Partee, B. (1978), 'Montague grammar and the well-formedness constraint', in Syntax and Semantics Vol. 10, eds. F. Heny and H. Schnelle, Academic Press, New York.

Stenius, E. (1967), 'Mood and language-game', Synthese 17, p. 254-274.

Tichy, P. (1971), 'An approach to intensional analysis', Nous V/3, 1971, 273-297.

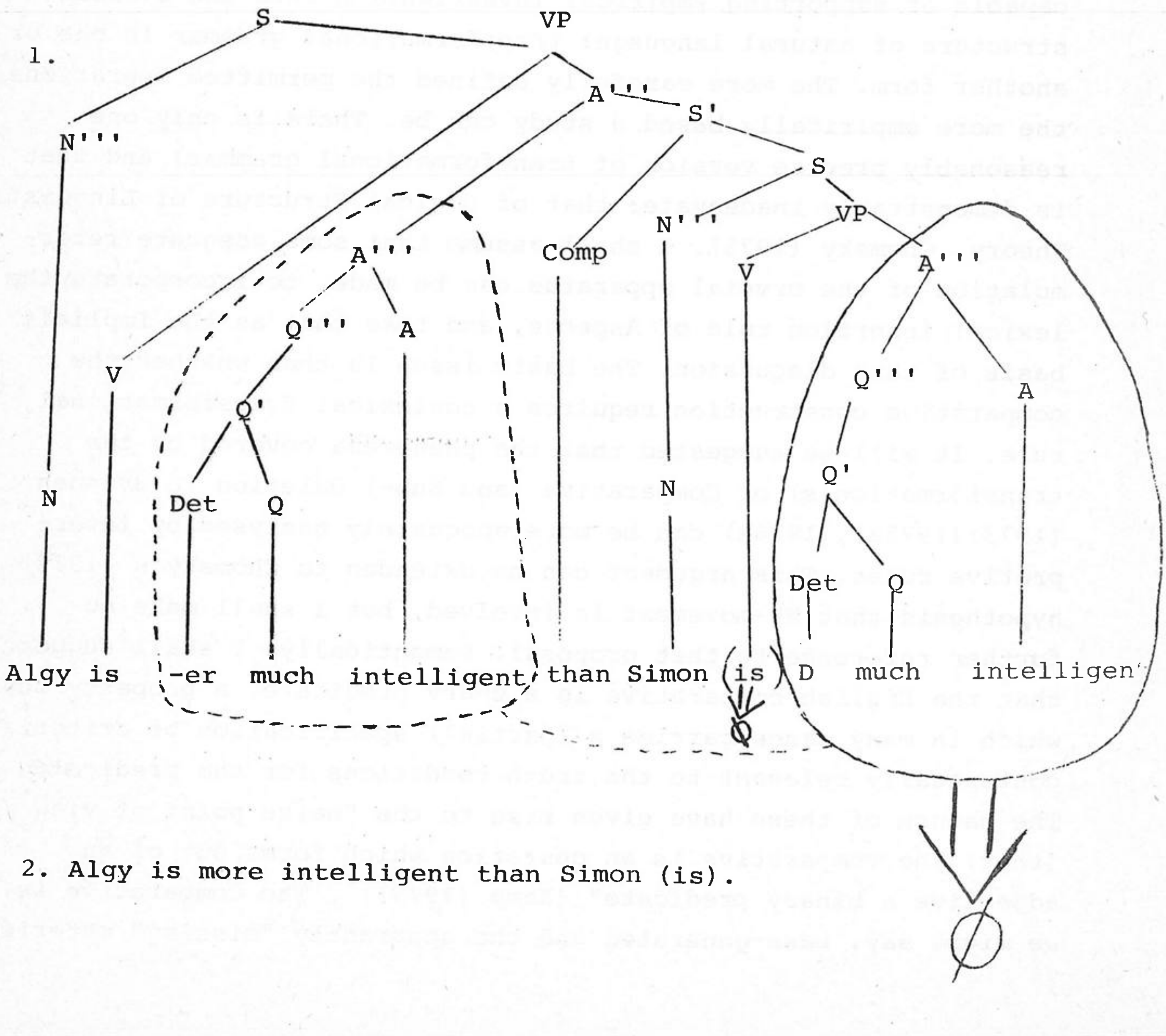
COMPARING ADJECTIVES -- AND GRAMMARS[1]

Frank Heny

An adequate treatment of the English comparative would presuppose a grammatical theory capable of handling virtually the whole of natural language. So I shall not attempt it. The construction involves practically every syntactic and semantic problem that I am aware of. Is deletion, or movement an operation that is indispensable in linguistic theory; will lexical rules (or some kind of categorial grammar of a constrained type) suffice to give a natural account? Problems of scope and opacity interact with the more basic semantic questions: are adjectives properties of individuals, functions from properties to properties -- or in some way normative, binary predicates? These issues underlie the ensuing discussion but will seldom emerge overtly since space dictates economy. There is only one theory of linguistic structure capable of supporting empirical investigation into the overall structure of natural language: transformational grammar in one or another form. The more carefully defined the permitted operations the more empirically based a study can be. There is only one reasonably precise version of transformational grammar; and that is demonstrably inadequate: that of Logical Structure of Linguistic Theory, Chomsky (1975). I shall assume that some adequate reformulation of the crucial apparatus can be made, to incorporate the lexical insertion rule of Aspects, and take that as the implicit basis of this discussion. The basic issue is then whether the comparative construction requires a nonlexical transformational rule. It will be suggested that the phenomena covered by the transformation(s) of Comparative (and Sub-) Deletion in Bresnan (1973)(1975a),(1976a) can be more adequately analysed by interpretive rules. This argument can be extended to Chomsky's (1977) hypothesis that WH-movement is involved, but I shall make no further reference to that proposal. Semantically, I shall suggest that the English comparative is a unary predicate, a property thus, which in many cases carries a (partial) specification of criteria contextually relevant to the truth conditions for the predicate. The nature of these have given rise to the "naive point of view (that) the comparative is an operation which forms out of an adjective a binary predicate" (Kamp (1975)) . The Comparative is, we might say, base-generated and the apparently "missing" material is to be handled interpretively. But neither by a delta-filling rule, as suggested by Williams (1977), nor by a rule of anaphora, as suggested by Jackendoff (1977). Nor does it help much to translate into even so versatile a logical language as Cresswell's lambda categorial language (Cresswell(1973)(1976)). Some indication will be given of what seems required if the crucial properties of the construction are to be given due weight: a mechanism partially specifying truth conditions.

The following remarks are based upon and owe everything to Bresnan's cited analyses of the construction; they are in no sense an attack on her analysis but point out cases where because of the theoretical grounding of that analysis it was in principle unable to provide a satisfactory account. The foundation was a deletion transformation, Comparative Deletion, which operated on structures of the form (1) to yield sentences like (2) :

1.



2. Algy is more intelligent than Simon (is).

As indicated, the A''' in the subordinate sentence is removed by a deletion rule, under the condition that it not be distinct from the A''' in the main clause: where the latter has -er, the former is empty, represented D in (1). The -er then cliticizes onto much to form more; for adjectives in the class of tall, more+Adj forms Adjer, thus, taller. An obligatory rule of much-deletion applies after -er, if present, has cliticized onto much; if at that stage, much is immediately to the left of an adjective it deletes. Hence, *(as) much intelligent becomes (as) intelligent. So brief a summary does not do justice to the original but is adequate for the criticisms that follow.

The fact that * too much tall is ungrammatical is to be explained by much ⇒ Ø / ____Adj. (Bresnan(1973) p. 278.). The word much is freely generated before nouns (He hasn't much sense.) and adjectives in the quantifier slot. If -er or some other item intervenes, the rule of Much deletion is blocked, but otherwise it applies. But analogous to the just cited "He hasn't much sense", we have "He has little intelligence": Little and much appear to be more or less parallel forms. We do not get *(too) little tall, any more than * (too) much tall. So it might be reasonable to attribute this to a little -deletion rule. This would also explain *John is as little tall as Bill and * John is little tall, ungrammatical forms precisely counterbalancing similar forms with much, and would provide a source for John is less tall than I thought he was, -er little becoming less. But of course that is nonsense. Every form putatively derived from a deleted much will also come from a deleted little. If there is any relationship at all between the postulated sources and meaning, such a hypothesis must be excluded.[2]

Trouble with little is not the only trouble with much. It has plenty of its own. Consider the following:

3. a. *John has much water in his glass
   b. John hasn't much water in his glass
   c. Has John much water in his glass

Some speakers of English might not totally reject a positive declarative like (3a); they would all agree that it is odd, and that the others are not. So it seems that much is not quite so freely distributed before nouns as we perhaps assumed.
Either there will have to be a rule deleting much before nouns but only in positive declaratives (deriving John has water in his glass from (3a)), or we shall have to constrain the domain of

lexical insertion of *much*. Presumably the choice of treatment could affect the account of *much* in adjective (and other quantified) phrases, and will in turn be influenced by such example as:

4.a. *Mary is very asleep
   b. *Mary is much asleep
   c. Mary is very much asleep
   b. Mary is more asleep than you think she is

5.a. Mary is very alive
   b. *Mary is much alive
   c. Mary is very much alive
   d. Mary is more alive at this party than she was at the last.

(In (5) it is obviously a special sense of *alive* that is intended; one not obtained with a plain positive adjective.) Grammaticality varies considerably in forms with and without both *much* and *very*. If a deletion rule for much is involved in whatever permits the (d) forms to be generated, this little transformation is going to have to be prevented from operating before *asleep* if *very* precedes it, and yet retain the obligatory force it has for most adjectives if there is no *very* present. Before *alive* in this special sense, it will be optional after *very* (or obligatory with optional base generation) and obligatory as usual if there is no *very*. Now this kind of idiosyncratic behavior governed by lexical items is just exactly what has been used at other times to motivate a lexical rule rather than a transformation. While it is true that we are not yet directly concerned with the rule of Comparative Deletion itself, it seems clear that the distributional facts about *much*, far from supporting an account of comparatives involving deletion of a lower phrase containing *much*, non-distinct from a higher phrase containing the same item and later yielding either *more* or the comparative adjective, are actually themselves a trifle unmanggeable. It begins to look as though *much* might not be the source of *more*, and hence be unavailable for deletion in the lower sentence.

In fact there is some slight evidence from a different source that *more* is a distinct item:

6.a. Did Sandy eat {many / *much} apples?
   b. Did Sandy eat {*many / much} ice cream?
   c. Did Sandy eat as {*many / *much} apples as he did ice cream?
   d. Did Sandy eat more apples than ice cream?

Only when the word *more* is involved can an acceptable comparison of mass and count be made. (And not always then, but that does not seem relevant.) It has been supposed that both many and much yield more when combined with -er. There may be other explanations of the above facts, but a very simple one is that the item more, occuring separately in the lexicon, is not sensitive, itself, to the relevant distinction.

Furthermore, there is reason to doubt the derivation of a comparative adjective form from a more supposedly cliticized onto the corresponding positive form -- the putative source, as we have seen, for the comparative. There are a number of adjectives which appear to undergo this "process" optionally. Perhaps sure (of oneself) is such an adjective:

7.a. John is a little more sure of himself these days.

b. John is a little surer of himself than he used to be.

And there is a modifier, unbelievably which patterns with positive adjectives, never with compared adjectives, but fairly freely with periphrastic forms based on more.

8.a. Jane is unbelievably tall.

b. *Jane is unbelievably taller than I expected.

c. This problem is unbelievably more difficult than I expected.

Even those that seem to "optionally" undergo "cliticization":

9.a. John is unbelievably more sure of himself than I expected.

b. *John is unbelievably surer of himself than I expected.

Even if we seldom find absolute differences in grammaticality in such pairs, it will require some fancy footwork to explain whatever difference there is (and that seems considerable in (9)) if a non-lexical transformation is responsible for converting forms like (9a) into (9b).

It is time to turn from the minor rules to the putative transformation of Comparative Deletion itself. A general methodological remark is called for. I am arguing against a transformational analysis. It has been more or less universally assumed by those working within the general framework developed by Chomsky in recent years that one is fairly well justified in assuming standardly accepted transformations unless good arguments are provided against them. Thus, it is quite normal to argue for lexical analysis. This is misconceived. Only if there are very good arguments can one adopt a non-lexical rule as part of an analysis. I have argued this point at some length elsewhere and will not pursue it here. (Heny (forthcoming, a). The relevance of this observation to what follows is just this: I shall try to show here that for reasons much like those provided by Chomsky for a lexical analysis of certain nominals, one would be led to adopt such an analysis for constructions putatively derived by Comparative Deletion; but it is not really incumbent upon me, or anyone else, to show that comparatives cannot be derived by transformation. On the contrary, if anyone proposes such a derivation he must show that it has some very distinct advantages over a lexical account.

Comparative Deletion has been left quite crucially underdefined. In Bresnan (1976,a,b.) there is actually no mention of comparative morphemes at all; the rule is to apply crosscategorially to X'', subject to the relativized A-over-A constraint and recoverability conditions. In Bresnan (1977) the morphemes than and as are required to be in the complementizer of the subordinate clause. That is obviously the intention, in all the published versions. But is it really an adequate reflection of the range of forms which appear to be relevantly close in form and meaning to those constructions that contain these two items?

Consider the following:

(10) John is the same height as Bill (is).

This is extremely similar in both form and meaning to

(11) John is as tall as Bill (is)

There are many such cases:

(12) a. Jane took the same number of pens to school as John did.

b. Jane took as many pens to school as John did.

It is worth pointing out that these additional as constructions behave just like standard comparatives in regard to their boundedness, to constraints such as the noun phrase constraint, contraction etc. This is shown in the following examples, which I will not discuss in detail. Reference to Bresnan's papers will make them clear to those who are not able to interpret them.

(13) a. Jerome took the same number of pens as Bill thought Algernon was accusing Max of having taken.

b. *Gillian took the same number of pens as Manfred knew an old schoolmate who had taken.

c. I will am taking as many pens to Chicago as Pete is to Salzburg.

d. *I will take as many pens to Chicago as Pete's to Salzburg.

Of course this should in principle be no trouble, for the rule is already formulated, at least in Bresnan (1977), with as in the complementizer position. But the problems arise with the supposedly "non-distinct" lower constituent subject to obligatory deletion. Returning to the simplest example, (10), it is quite difficult to suggest anything remotely suitable. "John is the same height as Bill is x much height" is obviously a non-starter.

Once such deviant forms are noticed, with even an as appropriately contained in the complementizer, it is a relatively simple matter to extend the problem yet further, For the word after is not difficult to analyse as a comparative in form and meaning on some level, and interestingly enough it behaves at least in certain respects like one.

(14) a. John left after Bill (left).

b. John left after Bill decided that his mother did.

c.*John left after Bill knew someone who had.

And of course before acts in very much the same way. In both cases, the contraction facts seem, in so far as they support a deletion analysis, just as relevant here as in the case of "real" comparatives.

(15) a. John is going to Chicago before his brother is to L.A.

b.*John is going to Chicago before his brother's to L.A.

The floodgates are open. Or at least there is one more form that seems closely akin to the comparative which nevertheless is hard to analyse by a deletion, like, as in

(16) John runs now like his father used to say I ran.

Clearly the meaning of the lower sentence is not that of "I ran" but rather "I ran thus" or something of that sort. How this could figure in the definition of a transformation deleting under non-distinctness is hard to see. In these cases we find no sign at all of an as or a than in the complementizer; if the transformation is to be extended, it seems that it will have to be governed by an odd assortment of lexical itcms. Perhaps it is as well to establish quickly that it will need to be governed. Unless we require that one of the appropriate morphemes appear properly embedded, it seems that Bresnan's rule would generate Jake and Jeff are tall(er) men that have gold, from Jake and Jeff are (-er) much tall men that have D much gold. The rule will actually have to be rather severely constrained to prevent this kind of thing.

Turning away from the problem of constraining the rule to operate on just the right forms, while including, as it seems certain would be necessary, some of these additional constructions, I want to point to a number of cases which are obviously less closely related to the comparative, but which quite clearly are, in some way, related. Syntactic rules like transformations do not in principe seem adequate to the task of capturing partial relatedness, degrees of relatedness and so on between these forms. Lexical rules obviously are. If some instances are covered by a transformation, others (because of their irregularity, for example) by lexical rule, it is doubly hard to envisage a notion of interrelatedness emerging from the description.

One, perhaps not quite serious, example of these "extended" comparatives comes readily to mind. What is the source of:

(17) I would rather be playing a mandolin than banging a typewriter.

We might propose:

(19) I would er much RATH be playing...than I would D much RATH be....

but that is quite absurd as a serious proposal within the extended standard theory of syntax, though interestingly reminiscent of generative semantics. There is in fact one part of the analysis forced upon Bresnan by the constraints within which she was working, in particular the assumption of a transformational rule deriving all (standard) comparatives, that is perhaps even more typical of that framework I mean the assumption of an underlying copula of abstract form, symbolised =, which is never realized at the surface and in that way is able to account for the grammaticality of John is taller than six feet, but not of *...than six feet is.

An alternative solution is available for such forms. Hankamer (1973) has in fact argued that there are two words than in English, one of which takes sentences, one noun phrases and prepositional phrases. Assuming that six feet is a noun phrase of some sort, then the above example would not be generated by Comparative Deletion at all, but would presumably be base generated as it is, and the ungrammatical *John is taller than six feet is might be excluded if John is tall and Six feet is tall, both of which seem grammatical, could be appropriately distinguished so that Comparative Deletion could not be applied to delete a part of some extension of the latter on the basis of what seems the equivalent part of the former.

Be that as it may, the proposal that certain comparatives are generated in the base, and not by deletion at all raises very serious problems in principle for the deletion account of the remaining forms, and leads naturally to a consideration of the semantics of the construction, to which we must turn directly. The problem is this. If some comparatives can be given appropriate syntax and semantics on the basis of apparently degenerate forms possessing the overt appearance of just one simplex sentence, then could not all comparatives be assigned appropriate status in the grammar on the basis of surface forms? This will depend in part on how extensive the "noun phrase" comparatives are. They seem to be quite central to the construction. If some kind of interpretive rule must deal with all comparatives that have superficially nothing but a noun phrase or PP after than , then the interpretive component will need to be quite powerful. While evidence adduced by Hankamer does not seem overwhelmingly convincing for English, there is no doubt that it would simplify our account of such forms to exclude them from those derived by deletion -- still assuming that there is a rule of comparative deletion in the grammar of English. Considerations relating to scope play some role in this. For while (19a) is ambiguous (19b) is not.

(19) a. John believes Sue is taller than she is.

b. John believes Sue is taller than herself.

Different derivations would obviously simplify an account of this fact, though how far such an account could be based on a distinction between a deletion vs. non-deletion account remains an open question worth careful investigation. In addition to scuh considerations, however, there are some interesting vaguenesses associated with the noun phrase construction which could rather easily be attributed to a derivation from a form close to the surface. A typical example is

(20) I'm sure you can find a nicer girl friend than Mary

This is open to several interpretations. It could be said if Mary had been the girl friend of the addressee but was no longer. Then the natural continuation in a comparable sentential comparative would be ...was. Said when Mary remained his girl friend at the time of utterance, it would be naturally completed ...is. If she is proving elusive and the speaker is seeking to dissuade the hearer from further pursuit, it could be continued ...might have been. And said in an effort to dissuade the speaker from an adventure that is (certain) to take place in the future, it might end ....will be. A deletion account of such examples is obviously, on the basis of assumptions about, and attempts to constrain such rules, on the part of those who continue to believe in them, highly unsatisfactory. It would appear that the language is permitting a considerable degree of vagueness and indeterminacy, providing rather minimal constraints on the conditions for truth.

One might of course seek to set these cases off completely from those derived from full sentences. In principle there could be nothing against such a proposal. However, there are reasons for thinking that the boundary between the two constructions simply will not settle in the right place. In particular, there are "sentential" and hence "deletion" comparatives where there simply is no suitable source for a deletion which will yield the right input to the semantics. Some are indeed very similar to the noun phrase cases. They have been called, by Williams "null complement anaphora", or at least are related to cases which he has discussed under that heading:

(21) John drove faster than I had hoped.

(* John drove faster than I had hoped he drove.)

(22) Jane drove faster than I like

(* Jane drove faster than I like Jane drove (fast).)

(23) Mike drove faster than I am used to.

And so on. There is once again a range of interpretations possible, just as in the case of the noun phrase comparatives; yet these have at least the structure of partial sentences.

Finally, there are cases with the structure of full sentences, where the interpretation is more or less fully determined, in fact, but where it does not correspond to any putative deep structure appropriate for derivation by a rule of deletion under non-distinctness. For the meaning that seems to be borne by the lower sentence actually contains additional items or what appear to be semantically distinct elements. Some simple cases, and probably the least troublesome for a deletion rule, involve negative items like little,the negative counterpart of much which on a very superficial level was hard to reconcile with the analysis. A sentence like (24) seems most likely to be derived under the standard account from a source containing few, as indicated:

(24) Sam has fewer books than his brother has

Sam has er few books than his brother has D few books.

It seems at the very least questionable whether we would want to use the putative source if Sam's brother had many thousands of books ...it is in fact he that has more than Sam. Comparable positive instances have in fact been discussed by Bresnan. She has suggested that in a sentence such as John drove more carefully than I did, there need be no implication that I drove carefully at all, though there is an implication that I drove; and this is consistent with a semantics for comparatives somewhat along the lines of the commonly assumed binary predicate over degrees. Merely to assign some degree on the scale of carefulness to my driving would be, positively, careful. The greater discomfort that one experiences in saying of a person who is specifically claimed to have more books, perhaps a very great number in fact, that he has a certain degree of fewness of books (to twist the language a little), is merely a reflection of the marked value of negative scalar items like few, young, small etc.

Be that as it may, and I leave this question open, there are nevertheless clear cases where these forms, and their positive counterparts too interact with certain modal notions to yield unquestionably inappropriate input to the semantics on a simple-minded deletion analysis.

The following two sentences provide quite clear examples, one positive, one negative.

(25) John shot fewer ducks than he was allowed to shoot.

(26) John shot more ducks than he ought to have shot.

The first would presumably be derived by non-distinct deletion from a source something like

(25') John shot er few ducks than he was allowed to shoot D few ducks

But clearly the point of the sentence is, in its standard interpretation, that he was allowed to shoot a certain number of ducks, and did not exceed his quota...i.e. he was allowed to shoot "D many" ducks. Someone allowed to shoot few ducks would be a reluctant hunter, forced to go out into the field but not forced to shoot many ducks. This is not part of the meaning of the sentence at all. The second example has an equally odd source:

(26') John shot er many ducks than he ought to have shot D many ducks.

But there was no number of ducks that John ought to have shot at all. He ought not to have shot more than a certain number.

Rather similar effects can turn up in quantified sentences of certain sorts. Thus, the following must not be understood to embody the claim that there was some distance such that several of the boys ran that far but John ran further. Rather, the point is that there is a certain distance that John ran and the boys in question did not run that far.

(27) John ran further than several of the boys.

Given the existence of the various irregularities of different sorts outlined here, it seems highly dubious that a transformation of comparative deletion is an appropriate way of analysing comparatives; but of course it remains to be shown how an interpretive account might work. It will obviously depend in part (and to a much greater degree than a syntactic rule need do) on an adequate account of the semantics of comparatives.

And that having been said the question obviously arises what in the context in which the notion "interpretive rule" is used here, an interpretive rule is. Williams (1977) has already suggested using one, so has Hellan(1977); and so Jackendoff (1977) has argued, for very different reasons from those given above, that one must have an interpretive acccount of the comparative. The first two are assuming rules that will copy material appearing elsewhere in a syntactic stree structure into nodes left empty (but generated) in an Aspects-type base structure. Jackendoff suggests that what is needed is an anaphoric pronoun, and in his examples clearly shows a single element at the point where the former comparative deletion rule would have operated to delete structure.

Neither proposal seems adequate. Pronouns do not obey the constraints typical of movements (and deletion) rules in standard transformational analysis. Thus a sentence such as John spoke to the man who had hit him is fine, with the pronoun inside the complex NP anaphoric to John. We have seen that comparatives obey the constraint. The proposal to use a rule filling in empty nodes appears to suffer from all the disadvantages, or nearly all the disadvantages of a transformational account, for the irregularities that we have pointed out cannot be taken care of any better by a rule which is essentially the converse of a transformation. What about an essentially interpretive rule, in the sense of a rule which assigns meanings to lexical items and then seeks to combine these to yield an appropriately projected semantics? Using a rough approximation to Cresswell's lambda categorial language as a vehicle for representing the structure of the result (but without any commitment to the appropriateness of this in the long run), let us see what might be done to interpret the simplest case, the plain noun phrase comparatives with adjectives, in particular a rule that will assign the sentence John is taller than Bill a suitable truth conditional semantics. The following is a rough approximation:

$\boxed{\text{-er}} \overset{tr}{\Longrightarrow} \lambda x_0^{<0,1,deg>} \langle \lambda x_0^{<<0,deg>,<0,1,deg>>} \langle \lambda x_1^{<0,1,deg>} \langle \lambda x_0 \langle \lambda \delta_0 \langle x_1^{<0,1,deg>}, x_0, \delta_0 \rangle \triangleright$

$\langle x_0^{<<0,deg>,<0,1,deg>>}, x_1^{<0,1,deg>} \rangle\rangle, x_0^{<0,1,deg>} \rangle\rangle\rangle$

$\underset{NP}{\boxed{\text{than}}} \overset{tr}{\Longrightarrow} \lambda x_1 \langle \lambda x_2^{<0,1,deg>} \langle \lambda \delta_1 \langle x_2^{<0,1deg>}, x_1, \delta_1 \rangle\rangle\rangle$

SO:

1. than Bill $\overset{tr}{\Longrightarrow} \lambda x_2^{<0,1,deg>} \langle \lambda \delta_1 \langle x_2^{<0,1,deg>}, \underline{Bill}, \delta_1 \rangle\rangle$

2. taller $\overset{tr}{\Longrightarrow} \lambda x_0^{<<0,deg>,<0,1,deg>>} \langle \lambda x_1^{<0,1,deg>} \langle \lambda x_0 \langle \lambda \delta_0 \langle x_1^{<0,1,deg>}, x_0, \delta_0 \rangle \triangleright$

$\langle x_0^{<<0,deg>,<0,1,deg>>}, x_1^{<0,1,deg>} \rangle\rangle \underline{tall} \rangle\rangle$

3. taller than Bill $\overset{tr}{\Longrightarrow} \lambda x_1^{<0,1,deg>} \langle \lambda x_0 \langle \lambda \delta_0 \langle x_1^{<0,1,deg>}, x_0, \delta_0 \rangle \triangleright$

$\langle \lambda x_2^{<0,1,deg>} \langle \lambda \delta_1 \langle x_2^{<0,1,deg>}, \underline{Bill}, \delta_1 \rangle\rangle x_1^{<0,1,deg>} \rangle\rangle \underline{tall} \rangle$

which becomes: $\lambda x_0 \langle \lambda \delta_0 \langle \underline{tall}, x_0, \delta_0 \rangle \triangleright \langle \lambda x_2^{<0,1,deg>} \langle \lambda \delta_1 \langle x_2^{<0,1,deg>}, \underline{Bill}, \delta_1 \rangle\rangle \underline{tall} \rangle\rangle$

which becomes: $\lambda x_0 \langle \lambda \delta_0 \langle \underline{tall}, x_0, \delta_0 \rangle \triangleright \langle \lambda \delta_1 \langle \underline{tall}, \underline{Bill}, \delta_1 \rangle\rangle$

apply to John: $\lambda \delta_0 \langle \underline{tall}, \underline{John}, \delta_0 \rangle \triangleright \lambda \delta_1 \langle \underline{tall}, \underline{Bill}, \delta_1 \rangle$

One point may be worth making briefly in connection with the semantics of ⊳ above. It is of course a binary predicate. The arguments are degrees. I have chosen to represent the comparative thus not purely arbitrarily, but because it seems to be the best approximation to an appropriate account that we have. In particular, it seems that the discussion in Kamp (1975) is in one crucial aspect unconvincing. The complexity of many adjectives relative to various criteria is obviously important. But these criteria appear to be held constant in any given context, and this permits the analysis to be made in terms of a simple (partial) ordering. An example will illustrate the crucial aspects of what is too long an argument to enter into properly here. Under Kamp's analysis of very and rather in terms of degrees of truth, it is impossible to conclude from John is very clever and Bill is (only) rather clever that John is cleverer than Bill. This is clearly wrong. The point is that in evaluating a single sentence, a single set of criteria are consistently applied, and comparatives are necessarily contained within a single sentence. Hence it seems perfectly permissible to adopt the reduction to a binary predicate.

Now, how usefull will a set of lexical meanings of the sort sketched above, be? We have 'read back' material into a suitable semantic structure, certainly without making any use of overtly repeated elements of the natural language syntax. Presumably such entries could be extended (with considerable complication) to do the same for other cases, though even for a simple extension to handle Bill has a bigger book than Sam we would have to build into the interpretation of -er and/or than variables to handle the predicate have, the property book, another individual variable, and so on. But the project is really pointless. For this more "seriously" interpretive framework is doing just what the mapping into empty nodes did; in particular, it is going to have to be separately complicated, as far as one can tell, to handle each of the little complications that have beset the other analyses.

Let us try again. We did not attempt to motivate the assumption that the comparative is a binary predicate. It has often been pointed out, e.g. by Bartsch and Vennemann (1972), that comparatives and ordinary positive adjectives have a good deal in common. Some analyses have been proposed of the positive which set that up as a binary predicate. The details do not matter. For I am going to suggest that the structure of English(the way in which comparatives are used, in other words, is such that we must analyse them as properties, with, in many contexts but not always, the requirement that certain rather specific aspects of the attendant truth conditions be overtly specified. Perhaps the mere fact that the comparative occurs in a higher sentence than the clause that goes with it is enough to suggest that this may have some value. But there are specific comparatives, (often overtly excluded from attempts at analysis) which suggest that this may well be the right direction to look:

(28) a. I don't have any more apples to give you
b. I have just received several more of those books
c. Mary grew taller and taller
d. Can you reach any higher?
e. Give me the larger piece of pie.
f. He is growing angrier every minute.

The suggestion is that in forms which have "missing" material this is at least in some cases actually not being specified. Just as in *Horses eat grass* and *Horses graze in that field* natural language is crucially under-specifying and we misconstrue the situation if we map material "back" where it never belonged. But what of the "deletion site" for comparatives? Surely that cannot be left unspecified. The suggestion (and on this occasion it must remain only that) is that there is again no need to specify....all that is necessary is that there be an indication of the relevant *place* in structure. For the comparative is, I have claimed a property, not a binary predicate, and it is obviously totally irrelevant to add, in the often highly degenerate sketch of criteria that constitute the comparative "clause" or phrase, information which exists overtly where it is needed, in the higher sentence.

FOOTNOTES

1. A number of the arguments that are used here have been around for quite a long time. I have not tried to attribute them to specific individuals. It is however important that I not give the impression that Bresnan is unaware of them; in many cases they appear in some form in footnotes to her own work. One worries about counter-examples when they suggest an alternative.

   Many people have helped to shape this paper. Ewan Klein has contributed in many ways expecially in connection with the abstraction.

2. Jackendoff (1977) suggests an interesting but still far from adequate account of little.

   For further discussion see Heny (forthcoming, b.).

References :

Bartsch, R. and T. Venneman (1972) Semantic Structures, Athenaum Verlag, Frankfurt/Main.

Bresnan, J. (1973 "Syntax o the Comparative Clause Construction in English," Linguistic Inquiry 4, 275-343.

Bresnan, J. (1975a) "Comparative deletion and constraints on transformations," Linguistic Analysis 1, 25-74.

Bresnan, J. (1976a) "On the form and functioning of transformations," Linguistic Inquiry 7, 3-40.

Bresnan, J. (1977), "Variables in the theory of transformations," Formal Syntax, Academic Press, Inc. New York. 157-197.

Chomsky, N. (1975), The Logical Structure of Linguistic Theory, Plenum Press, New York.

Chomsky, N. (1977), "On wh-movement," Formal Syntax, ed. by P.W. Culicover, Th. Wasow, A. Akmajian. Academic Press, Inc. New York. 71-133.

Cresswell, M.J. (1973), Logics and Languages, Methuen & Co. Ltd., London.

Cresswell, M.J. (1976), "The Semantics of Degree," Montague Grammar, ed. B. Partee. Academic Press, Inc. New York.

Hankamer, J. (1973) "Why There are Two than's in English," Papers from the Ninth Regional Meeting, Chicago Linguistics Society, 179-191.

Hellan, L. (1977) "Some syntax and/or semantics of ((at least) Norwegian) comparatives."

Heny, F.W. (forthcoming a.) Review of LSLT to appear in Synthese.

Heny, F.W. (forthcoming b.) "The Grammar as a Lexical Transformation".

Jackendoff, R. (1977), "Constraints on phrase structure rules," Formal Syntax, Academic Press Inc. New York. 249-285.

Jackendoff, R. (1978), $\bar{X}$-Syntax, MIT Press, Cambridge, Mass.

Kamp, J.A.W. (1975), "Two theories about adjectives". Formal Semantics of Natural Language, ed. E.L. Keenan, Cambridge University Press, London. 123-156.

Williams, E.S. (1977), "Discourse and Logical Form," Linguistic Inquiry 8-1, 101-139.

Compositionality and the form of the rules in Montague grammar

by

T.M.V. Janssen

ABSTRACT

Two fundamental principles concerning Montague grammar are proposed and a formalization of these principles is given. It is investigated which consequences these principles have for the possible forms of the rules in a Montague grammar.

KEY WORDS & PHRASES: *Montague grammar, Frege's principle*

1. THE PRINCIPLES

In the last years several extensions and variants of PTQ (MONTAGUE 1973) have been published. They deal with phenomena which are not treated in PTQ or are treated in a defective way. In order to do so, often new kinds of rules and new technical tools are incorporated in the system; a most interesting example is the attempt to make a synthesis of Montague grammar and transformational grammar (PARTEE 1973,1975). This growth of the kinds of rules is not restricted by PTQ since PTQ does not provide a description of what a possible rule is like. On the contrary, its syntactic rules suggest that any description of the desired effect by means of an English sentence is acceptable. This way of description has been used in semantics as well. Since this tool has a great expressive power, it is not surprising that the new proposed rules differ considerably.

Such an uncontrolled growth of the types of rules is for several reasons undesirable. The most important are probably the considerations of methodological nature. A standard definition provides, for instance, for a languange in which one can formulate observations, relations and generalizations and it provides a good starting point for formulating extensions, restrictions or deviations of the framework. Some objections of practical nature are as follows. If each "extension" uses a deviant framework, then one has to start each time anew in obtaining intuitions about the properties of the system and to check whether old knowledge still holds. If one uses a computer program then it means that one has to rewrite the program completely, probably using new data-types. As long as there is no standard, it is impossible to avoid ad-hoc programming. Objections of a linguistic nature are presented by PARTEE 1978. She wishes to restrict the concept of Montague grammar in order to come to a characterization of the possible natural languages. Finally, an important justification for this research can be found in the interesting results obtained by it.

Our investigations concerning the possible kinds of rules in a Montague grammar will be based on two principles. These principles have a simple and natural formulation, they are intuitively very appealing and I expect that hardly anybody will disagree with them. The priciples are:

SEMANTIC COMPOSITIONALITY PRINCIPLE:

The meaning of a compound expression is built up from the meanings of its constituent parts.

SYNTACTIC COMPOSITIONALITY PRINCIPLE:

Each syntactic rule operates on the well-formed expressions of specified categories in order to produce a well-formed expression of a specified category.

The semantic compositionality principle has a long history. It goes back to Tarski and Frege and therefore it is sometimes called the Fregean principle. This principle gave rise to the development of model-theoretic semantics for logic. The principle is fundamental for the Montague style of dealing with semantics and it is widely known and generally accepted among people working in Montague grammar. The syntactic compositionality expresses the way in which PTQ works and the way in which the syntax of much formal grammars is organized. The formulation of the syntactic compositionality principle is almost the same as the formulation of the well-formedness constraint of PARTEE 1978. We have chosen another name for the principle because we will give it a different interpretation from what Partee does; the name well-formedness constraint will be used for an interpretation which is more according to her ideas.

As a direct application of the well-formedness constraint Partee considers a rule which constructs adnominal adjectives from relative clauses. Its syntactic function F_i has the effect that:

F_i*(immigrant who is recent) = recent immigrant.*

The input for this rule is an ill-formed expression *(immigrant who is recent)* and she judges that therefore such a rule is interdicted by the well-formedness constraint. This argmentation is, in my opinion, based upon a confusion between a general definition of a possible grammatical rule, actual examples of such a rule and the notion adequacy of a grammar.

On the one side one may consider an abstract definition of the kind of grammars one wishes to use. This definition contains a definition of what the possible rules of a grammar are. A simple example would be the definition of a context-free rewriting rule. It is a natural requirement that on the basis of this definition one is able to decide whether a certain text describes a rule which satisfies the definition. If there is no effective way to obtain a definite answer on this question, such a definition is useless. The definition of the grammar also contains a description of how the rules should be used to produce strings over some alphabet. These strings are called the *well-formed expressions* over this alphabet generated by the grammar. A subset of the well-formed expressions is called the *generated language*. The generated language of the PTQ grammar for a fragment of English consists of the generated expressions of the category t. The language of intensional logic consist of all expressions generated by its grammar. According to the PTQ grammar for English the expressions *love* him_1 is well-formed whereas *love her* is not well-formed since this expression is not generated by the grammar.

On the other side one may consider some actual grammar satisfying the definition and some -in an other way defined- language (for instance English). Then one may ask whether this language is the same as the language generated by the grammar; In other words one may ask whether the grammar is adequate for that language. One should be aware that adequacy of a grammar is completely different notion from the notion definition of a grammar. Grammars may exist without being adequate for some natural language: in the definition of a grammar adequacy plays no role.

Clearly Partee understands by a well-formed expression an expression occuring as subexpression in some English sentence. Her well-formedness constraint states that all expressions produced by the grammar must be well-formed expression in the above sense. This is a mixing up of definition and adequacy and makes her constraint to an unusable one. Suppose that one is presented a list of rules and one is asked whether this list constitutes according to her constraint a list or rules of a grammar for English. In order to answer this question one may start to produce some strings and ask each time whether it is an well-formed expression English. Suppose they are well-formed, then one cannot conclude that the list obeys the constraint since not all possible outcomes of the rules are considered. One has to try and try again, but never the definite answer "yes" can be given (related questions in formal language theory are known to be recursively unsolvable). It is of course possible for some special lists to obtain a definite answer (e.g. if the list consists of one non-recursive rule), but a constraint one can check only for some special lists and not for others is not a usuable constraint on the system. So the well-formedness constraint cannot be accepted as a formal constraint on the possible rules in a Montague grammar. It must be considered as expressing just some hope or desire: namely that an adequate grammar of English will be such that it contains no rule which produces *recent immigrant* from *immigrant who is recent*. In section 4 we will try to guarantee some of this desire by a formal constraint.

We have to say ourselves what we understand by the phrase "well-formed expression" in the syntactic compositionality principle. It is the other interpretation: the well-formed expressions are the expressions generated by the rules of the grammar. Any output of a syntactic rule is produced by that syntactic rule, so it is a wellformed expression. The phrase in the principle stating that the rules produce well-formed expressions is a pleonasm. The same holds for the inputs: the only possible expressions of specified categories are the expressions generated by the grammar. The word well-formed gives no restriction on the expressions, it does not contribute anything to the meaning of the syntactic compositionality principle and it may be ommited. In section 4 we will consider what is left of the principle.

Thus we observe that the interpretation of the syntactic compositionality principle is not uniquely determined. There occur so much vague, undefined words in the principles that they hardly say anything; everybody can find its own interpretation in them. Maybe this explains why these principles are so attractive and acceptable. In order to give substance to the principles, we must make the interpretation of several such words explicit. The guiding policy will be to give the principles as much content as possible, they should express sufficient and necessary conditions. It will appear that the compositionality principles together with the additional definitions give rise to interesting claims about the possible rules in a Montague grammar: several rules proposed in the literature are interdicted by the principles. If someone is not willing to accept some of the consequences of the principles, he may reject the principles, or he may give another interpretation for the vague words. We will consider just <u>a</u> formalization, although I consider it as a rather straightforward one.

2. SEMANTIC COMPOSITIONALITY

In this section we will make explicit what we understand by building up the meaning of a "compound expression" from the meanings of its constituent parts. By a compound expression we understand an expression that is produced by some syntactic rule from some input expression. It does not matter how complex or simple the expression is. By speaking about "constituent parts" we do not mean that the compound expression is divided into parts and that we combine the meanings of these parts. We allow that certain parts have no meaning in isolation; such expressions are called syncategorematic expressions (examples from PTQ: *every, such that*). Since they do not contribute a meaning from which we can build a new one, there is no use in considering them as constituted parts. We also allow that words from the constituent parts are disappeared in the compound expression (in PTQ: term-substitution). By the parts of a compound expression we understand the expressions which served as inputs for the syntactic rule according to which the compound expression is constructed. But knowing these inputs is not enough to determine in which way they are constituents (*John* and *love* him_1 can be combined in at least two ways). We have to know in which sequel they served as inputs and of which syntactic rule. Therefore, by "the meanings of its constituent parts" we understand a list of meanings of syntactic expressions together with the information in which sequel they served as inputs and of which syntactic rule. This is the information which is available for constructing the meaning of the compound expression. At the one hand we may use all this information. For each syntactic rule we may combine the meanings of the constituent parts in another way. Therefore we will have for each syntactic rule a separate semantic rule. On the other hand we may use only this information. If the principle would be understood as stating that you may use this information, but also other information if you wish so, then the principle would not state the whole truth and it would become a hollow phrase.

What does "meaning" mean? Let us consider the case of declarative sentences. According to PTQ, such a sentence is translated into an expression of intensional logic (henceforth IL). The interpretation of this expression with respect to some point of reference and some variable assignment yields a truth-value. This truth-value does not constitute the meaning of the sentence; meaning is more. The meaning of a sentence determines the circumstances under which the sentence is true or false. So knowing the meaning is (for all variable assignments) knowing for what kind of indices the interpretation yields true and for which false. Generally stated, it is knowing (for each variable assignment) a function from indices to the possible denotations of a certain type: it is knowing the intension.(The reader might consult LEWIS 1972 for a more extensive argumentation to this approach to semantics). So meaning is some element in the domains of the model. This element can be represented by some expression from IL. The interpretation of this formula with respect to a certain index gives us the value of the intension function for that index. By an expression representing a certain meaning we understand such an IL-formula. Notice that the translation function from PTQ yields such a meaning representating expression. One should be aware that such an expression is not identical with the meaning it represents. One and the same meaning can be represented by several expressions. Each of them is equally good in this respect.

The formalization of the concept meaning we considered above, it not the last and final answer. The intuitive concept meaning is rather vague, one may relate it with several phenomena, which are not covered by the above approach. In our formalization all tautologies have the same meaning, if one is interested to discriminate among them, another formalization is needed. (See LEWIS 1972). No extension of PTQ actually uses this formalization of meaning, in essence all have the same formalization: meaning is an intension.

What is allowed for building up new meaning? Meanings are functions, and we can do with them everything that can be done with functions. I see no argument to restrict our tools here. For building a new meaning from old meanings we allow every method which can be used to define a new function from old ones. We consider some examples.

1. If the meanings of the expressions ϕ and ψ are functions yielding truth-values, then we may define a new function which yields true for a certain index if and only if both the meaning of ϕ and of ψ yield true. Then the new meaning can be represented by $\phi \wedge \psi$.

2. Suppose that the interpretation of α may operate on the meaning of β, then we may define a new meaning as the intension yielding at each point of reference the result of this operation. This new meaning is represented by $\alpha({}^\wedge\beta)$.

3. Suppose that the interpretation of η yields a truth value and η contains the free variable z then one may define as new meaning that function which yields true for an index if there is a variable assignment for the variable z such that the interpretation of η yields true for that assignment. This new meaning can be represented by $\exists z\eta$.

One observes that in all the cases above, the interpretation of the new meaning consists of a single IL-formula containing the representations of the old meanings. These occur unchanged and identifiable in the representation of the new meaning: the representation of the first constituent part occurs unchanged at a specified position, the second at another position and so on. If we would allow that the representations of the old meanings become changed, we would define an operation on representations. This does not always define an operation on the meanings they represent (see also section 3). We make the semantic compositionality principle operational by requiring that *we may only define operations on meanings by a providing a single formula from IL (or an appropriate extension thereof) which contains at specified positions the unchanged representations of the input meanings*. One might at first glance be tempted to think that this operational version is more restrictive than the original formalization we give of the principle. This is, however, not completely true; it is a restriction on the format in which operations on meanings can be represented. If someone considers IL as being too restrictive for his purposes, he may extend IL by new operators etc. The interpretation of such operators must recursively be defined as is done in PTQ for the usual ones ($\Box$,H,$\exists$,$\wedge$). The main advantage of requiring that the representation of an operation on meanings is presented by an IL expression is that this way of presenting guarantees that the compositionality principle is obeyed automatically.

We have explained rather extensively how we came to our formalization and operationalization of the semantic compositionality principle. I hope that no one will be surprised by it, we followed the obvious way from intuition towards a precise formalization. Some of the implications of this formalization have been stated by others, although not with the argumentation and coherence as we did. PARTEE 1973 states that the "the translation rule must be such that the translation of the input expression must occur intact in the translation of the output". PARTEE 1978 mentions as a "semantical constraint" that the meaning of a compound should be given by means of an IL formula. The more surprising is it, that several authors present extensions or variations of PTQ which do not obey the so fundamental principle of semantic compositionality. In most cases it is not the complexity of the problems which makes it difficult to obey the principle; in contrary, often it is not so difficult to provide a proposal which is in accordance with the principle. Probably one is not aware of the implications of the principle one accepts; therefore it seemed useful to present such an extensive argumentation. In section 3 we will consider some proposals which do violate the principle.

3. EXAMPLES CONCERNING SEMANTIC COMPOSITIONALITY

1. Shake John awake

DOWTY 1976 treats, among others, the semantics of factive constructions such as *shake John awake*. In order to do so, he extends the language of intensional logic with two operators: CAUSE and BECOME. Interesting for our discussion is his treatment of CAUSE. In order to define its in interpretation Dowty adds "to the semantic apparatus of PTQ a selection function f that assigns to each wff ϕ and each $i \in I$ a member $f(\phi,i)$ of I. [Intuitively $f(\phi,i)$ is to be that i' most like i with the (possible) exception that ϕ is the case [..]]". Then the interpretation of CAUSE reads:
"If $\phi,\psi \in$ ME then $(\phi \text{ CAUSE } \psi)^{A,i,j,g}$ is 1 if and only if $[\phi\wedge\psi]^{A,i,j,g}$ is 1 and $[\neg\psi]^{A,f(\neg\phi,i),j,g}$ is 1.

The function f is defined on IL-expressions and not on the interpretations of these expressions. As a consequence CAUSE is an operator on IL-expressions and not on the meanings they represent. This is illustrated as follows. The definition of f allows that for some ϕ,η,i holds that $f(\phi\wedge\eta,i) \neq f(\eta\wedge\phi,i)$. So it may be the case that $\neg(\phi\wedge\eta) \text{ CAUSE } \psi^{A,i,j,g}$ yields 1 whereas $\neg(\eta\wedge\phi) \text{ CAUSE } \psi^{A,i,j,g}$ yields 0. So the compositionality principle is not obeyed and, moreover, the implications of this CAUSE are incorrect. A correction is possible by taking as domain for f the intensions of formulas: f assigns to each $d \in D_{<s,t>}$ and $i \in I$ a member $f(d,i) \in I$. Then a situation as described above is automatically excluded. The interpretation of CAUSE now becomes

"If $\phi,\psi \in ME_t$ then ϕ CAUSE $\psi^{A,i,j,g}$ is 1 if and only if $[\phi\wedge\psi]^{A,i,j,g}$ is 1 and $[\neg\psi]^{A,\underline{i},j,g}$ where $\underline{i} = f((^\wedge\neg\psi)^{A,i,j,g},i)$.

2. Horse Cannonero

DELACRUZ (1976) considers expressions like *the horse Cannonero*. Such expressions belong to a category $\bar{T}$ and they are generated by the following rule:

S3.1 If $\alpha \in B_T$ and $\zeta \in B_{CN}$ then $F_{21}(\zeta,\alpha)$ $P_{\bar{T}}$, provided that whenever ζ is of the form he_n, $F_{21}(\zeta,\alpha) = \alpha$; otherwise $F_{21}(\zeta,\alpha)$ = the ζ α.

Examples:

F_{21}*(horse,Cannonero) = the horse Cannonero*

F_{21}*(horse,*he_1*)* = he_1

Translation rule:

T3.1 If $\alpha \in B_t$, $\zeta \in B_{CN}$ and α,ζ translate into α',ζ' respectively, then $F_{21}(\zeta,\alpha)$ translates into α' if α is of the form he_n; otherwise $F_{21}(\zeta,\alpha)$ translates into

(1) $\lambda P\exists y[\forall x[[\zeta'(x) \wedge \lambda P\lambda zP\{^\wedge\lambda x[^\vee z={}^\vee x]\}(^\wedge\alpha')(x)] \leftrightarrow x=y] \wedge P\{y\}]$

Translation rule T3.1 refers to the syntactic form of the input expressions of the syntactic rule. This means that in order to obtain the representation of the compound expression we need more than only the meanings of the inputs for the syntactic rule and the information which rule is used. So T3.1 violates the formalization of the semantic compositionality principle. The correction of this rule can be provided for in the syntax. The phenomenon considered by Delacruz provides evidence that among the terms we should distinguish syntactically Proper names and indexed pronouns and ask in S3.1 just for a proper name as input. Notice that the formula Delacruz presets is not the simplest one. I would prefer

(2) $\lambda P\exists y\forall x[\zeta'(x) \wedge \alpha'(^\wedge\lambda z[^\vee x={}^\vee z]) \leftrightarrow x=y] \wedge P\{y\}]$.

3. Easy to please

This example concerns a rule which is so close to a correct formulation that I would not like to call it a violation of the principle; is rather a (illustrative) slip of the pen. The main reason for mentioning it, is that we will use it in the discussion of the syntactic compositionality principle. We consider the following rule from PARTEE 1973.

Derived verb phrase rule:

If $\phi \in P_t$ and ϕ has the form ${}_t[{}_T[he_i]{}_{IV}[\alpha]]$, then $F_{104}(\phi) \in P_{IV}$, where $F_{104}(\phi) = \alpha'$, and α' comes from α by replacing each occurence of *he*$_i$, *him*$_i$, *him*$_i$*self* by *he**, *him** *him***self* respectively.

Examples:

F_{104}(*he*$_1$ *sees him*$_1$*self*) = *see him***self*

F_{104}(*he*$_7$ *is easy to please*) = *be easy to please*.

Translation rule

If $\phi \in P_t$ and ϕ translates into ϕ', then $F_{104}(\phi)$ translates into $\lambda x_i \phi'$.

From the formulation of the translation rule it is not as evident as in the previous example that the translation rule uses syntactic information. In order to decide what the actual translation is ($\lambda x_1 \phi$ or $\lambda x_2 \phi$ or ...) one needs to know the index of the first word of ϕ. The correction of this rule rather simple, in analogy of term-substitution in PTQ we give the syntactic operation an index as parameter: so F_{104} is replaced by $F_{104,i}$. In a later paper (PARTEE 1977) she corrected the rule in this way.

4. John who runs

BARTSCH 1976 and BARTSCH 1978 considers term phrases containing non-restrictive relative clauses. Such expressions are produced from a term and a sentence by the following rule (BARTSCH 1978)

S4. If α is a term and β a relative sentence, then $\beta(\alpha)$ is a term. [...]

The corresponding translation rule reads

T4. If α' is the translation of the term α and $RELT(\lambda x\ \beta'(x))$ is the translation of the relative clause β from S4, then $(RELT(\lambda x\ \beta'(x)))(\alpha')$ is the translation of $\beta(\alpha)$, and for all terms α with $\alpha' = \lambda P(..P(\nu)...)$ we have: $(RELT(\lambda x\ \beta'(x)))(\lambda P(...P(\nu)...)) = \lambda P(...\beta'(\nu)\ \&\ P(\nu)...)$.

Take as the representation for the meaning of the term *every man*

(3) $\lambda P \forall v[\text{man}'(v) \rightarrow P(v)]$.

If we combine (3) with the translation of some relative clause, the effect of T4 is well-defined. We might also consider another representation for the meaning of *every man*. Let Q be a variable of the same type as P and in (1) and Let $\mathcal{R}$ be a variable of the same type as the translation of term. Now the effect of T4 is not defined for:

(4) $\lambda Q \forall v[\lambda \mathcal{R}[\mathcal{R}(\text{man}') \rightarrow \mathcal{R}(Q)]\ (\lambda P\ P(v))]$.

A reaction on these objections against rules like T4 might be that one adds to the rule a clause stating that if the input formula is not in the required format, it must be reduced to that format. This a is a very simplified formulation of complex way of defining function between meanings. In order to define functions, one has to fulfill the following three requirements.

1. One has to describe exactly for which representations one will define the function.
2. One has to define for all expressions in the subset what the effect of the function is.
3. I consider two alternatives.

3a. One has to prove that each meaning for which we wish to define the function has precisely one representation in the subset defined in 1. (It lies at hand to prove this by providing an algorithm which transforms a given expression into one in the subset. This can, however, not be done since it would bring us in conflict with the undecidability of IL).
3b. One has to prove that each meaning for which we wish to define the function, has at least one representation in the subset and moreover that the result of applying the function to two different representations of the same meaning yields the same result.

Rule T4 does not fulfill the requirements. In general it is a complicated and extensive task to define a function between meaning by defining a function for specially selected representations. It can probably only be done in practice, if one considers a situation with a special structure in which all the proofs become drastically simplified. But if the situation is such a special one, one may expect that the same effect can be obtained in a more direct way as is demonstrated below.

The idea behind our reformulation is that the effect of replacing $P(v)$ by $\beta'(v)$ & $P(v)$ can be obtained by giving $\lambda z[\beta'(z) \;\&\; P(z)]$ as argument of $\lambda P[\ldots P(v)\ldots]$. We must take care of the binding of the variable and thus we come to a version of T4 which is in accordance with our formalization and operationalization of the semantic compositionality principle: T4'. Let α' be the translation of the term α and γ' the translation of the relative clause γ. Then the translation of the compound expression $\gamma(\alpha)$ is:

(5) $\lambda Q(\alpha'(\lambda z[\gamma'(z) \;\&\; Q(z)]))$

One observes that it is not needed to define the intended function from T4 along the laboured way of defining a mapping on special selected representations. The formulation of T4' is more exact and more simple than the formulation in T4.

One might take instead of (5) a more complex representation in which the translation of the term is operand rather than operator. This gives

(6) $\lambda R[\lambda Q(R(\lambda z[\gamma'(z) \;\&\; Q(z)]))](\alpha')$.

It is interesting to compare (8) with the expression RODMAN 1976 gives for the nonrestrictive relative clause:

(7) $\lambda P[\lambda Q\; P\{{}^\wedge\lambda x_n[\gamma \;\&\; Q\{x_n\}]\}]\;({}^\wedge\alpha')$.

It turns out that their approaches are semantically in essence the same. This illustrates the use of writing semantic functions in the same format, from the formulation in T4 the close relationship could not be observed.

5. a[p][q] =: y

JANSSEN & van EMDE BOAS 1977 present a Montague grammar for the syntax and semantics of the assignment statement in the programming language ALGOL 68. They treat the semantics of an assignment to arrays of dimension n by reducing it to the case of dimension n-1. Unfortunately this approach gives rise to problems concerning semantic compositionality principle. Adopting Lewis formalization of meaning (section 2) would make it possible to save the principle. This is, however, a rather sneaky escape since in the context of programming languages all tautologies do have the same meaning. Although the compositionality principle is not standard among computer scientist, the authors prefer to obey it. So the authors have to give up their claim that they can treat the semantics of assignments to arrays without the need of a separate rule for each dimension. Giving up this claim is not too hard: there remain enough reasons for preferring their Montague-still approach. It is the only known semantic treatment of pointers and it is a proposal which (now) obeys the semantic compositionality principle whereas several other proposals do not.

4. SYNTACTIC COMPOSITIONALITY

We will interprete the syntactic compositionality principle (just as we did for the semantic principle), as giving a necessary and sufficient condition. It is not so surprising that a rule may operate on expressions of specified categories. The interesting aspect of the principle is that it states that this is also sufficient. Once the input expressions of specified categories are available, the rule can be applied. One does not need to know in which larger expression context the expression will be used. If several rules can be applied, then they are equally possible: there is no prescribed order among them. The derivational history has no influence on the question whether the rule applies or not. Even the actual form of the input expression is not of importance: if expressions of the required categories are available, the rule always is applicable. *The syntactic rules must be total rules!*

There are several arguments for interpreting the syntactic compositionality principle in this way. The most important one is that the requirement of total rules gives rise, in combination with the semantic compositionality principle, to several important and attractive consequences concerning the form of the rules. One consequence will be discussed below others will be discussed in the next sections. Another argument is that our interpretation of the syntactic compositionality principle expresses the way in which much formal grammars work, for instance PTQ has total rules. It is moreover an attractive and elegant principle because of its analogy with (our formalization of) the semantic compositionality principle. Finally, there is a practical motivation for total rules. Total rules are easier to understand and can easier be handled by a computer program.

The first consequence of having total concerns the "un-well-formed" expressions. Suppose the grammar contains a rule S_i of which the syntactic operation F_i has the following effect:

F_i(*immigrant who is recent*) = *recent immigrant*.

So the rule operates on a common noun-phrase which, according to rule S3.1, must be constructed from the common noun *immigrant* and the sentence *he*$_1$ *is recent*. This sentence must come from the IV-phrase *be recent*. Since we require that the rules are total, we may also combine this IV-phrase with other term-phrases. So also the sentence *John is recent* is generated by this grammar, which is not a correct sentence of English. This example suggests that an adequate grammar for English cannot contain a rule which generates *recent immigrant* in the way rule S_i does.

The above reasoning is not a mathematically proof that it is absolutely impossible that an adequate grammar for English contains the rule S_i. In fact it can be made possible by changing the PTQ rules and applying the following trick. We split each category in two new ones: one that may contain expressions such as *is recent* and one that may not contain them. So the grammar becomes rather complicated while a simple solution lies at hand. It is unlikely that someone will ever write down a grammar as sketched above. If some rule in the grammar introduces an "un-well-formed" expression, then it is due to the totalness of the rules, very difficult to get rid of that expression. So in a certain sense Partees well-formedness constraint is saved. The syntactic compositionality principle, with the interpretation of requiring total rules guides us towards a grammar which fulfills her desire.

5. TRANSFORMATIONS

In this section we will investigate the consequences of the two compositionality principles for the incorporation in Montague grammar of transformations as they are used in transformational grammars. Therefore we consider some fundamental aspects of such transformation. These are:

1. Transformations define mappings from trees to trees rather then from strings to strings.
2. If several transformations can be applied, then the order of application may be prescribed.
3. A transformation applies on one input tree.
4. A transformation imposes certain structural conditions on the analysis of the input tree.

In order to take care of the first point, it is required that a Montague grammar does not operate on plain strings but on trees or equivalently on labelled bracketings. Let us assume that the Montagues framework can be adapted in this way. This change of the system makes all the rules to rules which operate on trees, so in a certain sense all the rules in Montague grammar become transformations. In order to avoid confusion in terminology, we will use the name C-transformation ("Chomskyan") for transformations as they are used in transformational grammars. Once they are incorporated in Montague grammar they are called P-transformations ("phrase-structure"). The second point gives rise to problems concerning syntactic compositionality. It asks for more than only the category of the input expressions: it asks for information about the derivational history of the expressions. In a grammar which obeys the syntactic compositionality principle, there cannot be a prescribed order on the order of applications of the rules: only an implicit ording is possible. Notice that PARTEE 1978 comes, based upon the well-formedness constraint, to the same conclusion concerning this point. The third point is rather peculiar in the context of a Montague grammar. A syntactic rule in a Montague grammar may have any number of inputs. It seems rather artificial to incorporate transformation-like rules in Montague grammar, while at the same time restricting such rules to the case of rules with one input expression. The most problematic is the last point since it implies that C-transformations are partial rules. It is a very important aspect of C-transformations that they have conditions on the structural analysis of their inputs; this aspect makes them very attractive for practical use. It makes it possible to indicate what the relevant trees are without bothering about all details that are considered as being irrelevant. Evidently the aspect of structural conditions where for PARTEE 1978 the main reason for allowing partial rules.

In spite of the above considerations, we will continue to require total rules. We will incorporate C-transformations in Montague grammar by means of a slight reformulation that makes them total. The reader might be surprised by this reformulation and consider it at first glance as a sneaky trick used in order to obey the letter of the principle. This is not completely true: the reformulation expresses a different view on transformations and this has, in combination with the semantic compositionality principle, important consequences.

Let us demonstrate how the reformulation works. Suppose a C-transformation is given in the following format.

> If the input expression is of the category C_1 and it satisfies structural condition SC, then we may apply transformation T in order to obtain an expression of the category C_2 and else we may not apply T.

Its reformulation as a total rule has the following format:

> If the input expression is of the category C_1 then we may apply the transformation T'. Transformation T' reads as follows. If the input expression satisfies the structural condition SC, then apply transformation T and otherwise apply a specified do-nothing transformation.

By a do nothing transformation we understand a rule which gives one of the inputs unchanged as output. Notice that in this reformulation the do-nothing transformation reflects the essential difference with Montague rule descending from C-transformations. So the essential aspect of a P-transformation is the occurrence of a do-nothing transformation.

The reformulation expresses the conception that the P-transformation always applies if an expression of the required category is available and that this application always yields an output. For rules with one input the output may, under certain conditions, be that input itself. Since the input category and the output category are specified in the syntactic rule, this means that for a P-transformation with one input, the output has always the same category as the input. It is a nice coincidence that this is always the case for C-transformations. For transformations with more than input, the use of a do-nothing transformation implies that (at least) one of the inputs must have the same category as the input. This consequence disallows us to impose conditions on the inputs of a rule like S4 (which combines a term-phrase and an IV-phrase to a sentence).

The other consequence of having total rules is that the corresponding semantic rule always applies. We already observed that for every P-transformation it may under certain conditions be the case that one of the inputs is used as output. So in this case the meaning of the output expression is the same as the meaning of that input. We observed in section 3 that it is a consequence of the semantic compositionality principle that the semantic rule must be the same for all inputs: it may not depend on the structural analysis of the input expressions. So in all cases the meaning of the output of P-transformation must be the same as the meaning of the input. The translation rule corresponding with a P-transformation must be the identity mapping on one of its inputs. This means, for instance, that a variant of rule S14 from PTQ cannot impose conditions on its input expressions since the corresponding translation rule is not the identity on one of its inputs. If we restrict our attention to P-transformations with one input, we come to the following consequence. From the two compositionality principles it follows that *the transformations in a Montague grammar must be meaning preserving!* For transformational grammar this is a well-known requirement (e.g. PARTEE 1971). Some examples in the literature of transformations which are not meaning-preserving, will be considered in the next section.

6. EXAMPLES CONCERNING SYNTACTIC COMPOSITIONALITY

1. He_1 is loved

PARTEE 1973 considers the C-transformation Passive Agent Deletion. An example is:

$F_{102}(he_1 \text{ is loved by } him_3) = he_1 \text{ is loved}$

Translation: If $\phi \in P_t$ and ϕ translates into ϕ', then $F_{102}(\phi)$ translates into $\exists x_j \phi'$.

At the one hand this transformation applies only to input trees of a special form. At the other hand the translation rule is not the identity mapping. This means that we cannot reformulate this transformation as a total rule. So the traditional way of dealing with agentless passive is disallowed by the compositionality principles. THOMASON 1976 provides rules for generating agentless passive by means of categorical rules and DOWTY 1978 does so for the case of double passives. So there is an alternative way available for dealing with agentless passive. When we follow this way, it is rather strange to generate passives which do contain an agent by means of transformations; they also can be generated according to the rules of Thomason and Dowty.

2. Mary shakes John awake again

In section 3 example 1 we considered some semantic aspects of the proposals of DOWTY 1976 concerning the treatment of factives. Now we will consider some syntactic aspects. Of course, in doing this we cannot neglect the semantic aspects of the rules. Dowty produces the factive sentence *Mary shakes John awake* from the term *Mary* and the IV-phrase *shake John awake*. This IV-phrase in its turn is obtained from the TV-phrase *shake awake*. The first rule he presents for generating this TV-phrase is as follows.

S30 If $\alpha \in P_{IV}$ and $\phi \in P_t$ and ϕ has the form *he$_n$* is γ then $F_{30,n}(\alpha,\phi) \in P_{TV}$ where $F_{30,n}(\alpha,\phi) = \alpha\phi$.

An example is:

$F_{30,1}$(*shake, he$_1$ is awake*) = *shake awake*.

The corresponding translation rule reads: If α translates into α' and ϕ translates into ϕ' then $F_{30,n}(\alpha,\beta)$ translates into

$$\lambda\mathcal{P}\lambda x\mathcal{P}\{{}^\wedge\lambda y\,[\alpha'(x,{}^\wedge\lambda P[P\{y\}])\ \textit{CAUSE}\ [\textit{BECOME}\ [\phi']]]\}.$$

Again this rule is a partial rule which is not meaning preserving. So we have to find another approach; can the above result be obtained by means of a total rule? For generating expressions like *shake awake* one only needs an adjective and a TV-phrase. So it lies at hand to try the following rule:

S_{930} If $\alpha \in P_{TV}$ and $\beta \in P_{adj}$ then $F_{930}(\alpha,\beta) \in P_{TV}$ where $F_{930}(\alpha,\beta) = \alpha\beta$.

The corresponding translation rule would be

T_{930} If α translates into α' and β translates into β', then $F_{930}(\alpha,\beta)$ translates into

$\lambda P \lambda x P\{{}^\wedge\lambda y[\alpha'(x, {}^\wedge\lambda P\, P\{y\})])$ *CAUSE* [*BECOME* $[\beta'(y)]]]\}$.

Mary shakes John awake again is ambiguous. On the one reading Mary has done it before, on the other John has been awake before. Dowty treats *again* as a sentence modifier and he needs two different sentences in the derivation in order to deal with the ambiguity. He probably starts his investigations along this line for historical reasons: it is the way in which such constructions are treated in generative semantics. By rule S_{930} we are guided to another approach to this ambiguity. The one reading can be obtained by combining *again* with *Mary shakes John awake*, the other by combining it with *shake awake*. We do not go into details of this approach for the following reason. After considering several phenomena concerning factives, Dowty observes that his first approach is not completely adequate. He discusses extensively several alternatives and escapes. Finally he concludes "there would be no reason why we should not then take the step of simplifying rules S30-S32 drastically by omitting the intermediate stage in which a sentence is produced". He presents a rule which is identical with S_{930} as the rule which he considers as the best one. So the syntactic compositionality principle has led us immediately to the solution which is the simplest and the best one. This example suggest that we might derive from the syntactic compositionality principle the advice "to ask for what you need and not for more".

3. Easy to see himself

In section 3 example 3 we considered the Derived Verb Phrase rule from PARTEE 1973. An example concerning this rule was

F_{104}(*he*$_1$ *sees him*$_1$*self*) = *see him***self*.

At the one hand the derived verb phrase rule is a partial rule, at the other hand its output belongs to a different category than its input. Therefore we cannot reformulate this rule as a total one using a do-nothing transformation. So we have to find another treatment for the cases where Partee uses this rule. Let us, in accordance with the syntactic compositionality principle, just ask for what we actually need and not for more. In the above example we only need a TV-phrase. So the syntactic rule becomes

S_{902} If $\alpha \in P_{TV}$ then $F_{902}(\alpha) \in P_{IV}$ where $F_{902}(\alpha) = \alpha$ *him***self*

The corresponding translation rule reads:

T_{902} If α translates into α', then $F_{902}(\alpha)$ translates into $\lambda x\ \alpha'(x,{}^\wedge\lambda P\ P\{x\})$.

PARTEE 1975 provides as an explicit argument for use of the Derived Verb Phrase rule the treatment of the sentence

(8) *John tries to see himself.*

If the derived verb phrase rule is used for (1) then the translation becomes:

(9) $try\ to'({}^\wedge j,{}^\wedge\lambda x_3\ see'(x_3,{}^\wedge\lambda P\ P\{x_3\}))$.

Sentence (1) can also be generated according to the lines of PTQ. Then the translation is

(10) $try\ to'({}^\wedge j,{}^\wedge see'(\lambda P\ P\{{}^\wedge j\}))$.

Partee gives some considerations why she might prefer (9) above (10). Rule S_{902} is compatible with that opinion, it gives rise to a translation which is equivalent with (9). So we have found a simple, total rule.

Partee also presents the example

F_{104}(*he*$_7$ *is easy to please*) = *be easy to please*.

This is not treated by S_{902}. We search for a total rule and thus ask just for what we need for generating *easy to please*. We need an expression like *easy* and some TV-phrase. Let us, as Partee does, assume that we have a special category $\overline{ADJ}$ which contains easy, though etc. The resulting expression easy to please will be of the category ADJ'. Then we are guided to the following rule

S_{903} If $\alpha \in P_{\overline{ADJ}}$ and $\beta \in P_{TV}$ then $F_{903}(\alpha,\beta) \in P_{ADJ}$ where $F_{903}(\alpha,\beta) = \alpha$ to β.

The translation of (this) *easy* must be such that it may be combined with a TV-translation in order to obtain an expression of the type of translations of adjectives. Then the translation rule reads

T_{903} If α translates as α' and β as β' then $F_{903}(\alpha,\beta)$ translates into $\lambda x\ \alpha'(\lambda y \beta'(x,\lambda P\ P\{y\}))$.

Rule S_{903} makes it possible to generate the expressions containing easy to please which we mentioned above. Unfortunately, Partee does not provide an explicit semantics for the source of her construction (*to please him*$_7$ *is easy*), so it is difficult to compare the semantic consequences of S_{903}. I expect that she will finally end up with something close to the result of T_{903}. Concerning the syntax, it is demonstrated that our principles guide us to a much simpler treatment

4. Comparatives

There are more situations where the traditional approach uses rules which ask for more input than actually is needed. We will make some remarks concerning the case of comparatives. An example is:

(11) *Fewer of the women came to the party than of the men.*

This sentence is derived from the (illformed) sentence:

(12) *Fewer of the women came to the party than of the men came to the party.*

This is in its turn derived from a source sentence like

(13) *Fewer of the women came to the party than x many of the men came to the party.*

PARTEE 1978 presents (11) as a difficult case for the well-formedness constraint since it uses an illformed source (12). For the syntactic compositionality principle the intermediate stage of (12) is undesirable since one can hardly get rid of this sentence. Moreover the rules relating comparatives with sources as (12) or (13) are rather complicated. Why following this approach? Maybe one judges that a source like (12) or (13) expresses the semantic content of the comparatives more completely than comparatives. Or one wishes to explain the semantic relations between comparatives by generating them from the same kind of source. In transformational grammar this might be valid arguments, no other formal tools than transformations are available. In Montague grammar there is semantic component in which such semantic relations can be formally expressed. So if we do not need such a source for syntactic reasons we may try another approach. The syntactic compositionality guides us to ask for just what we need. In order to make a comparison between two individuals concerning the amount of some property they have, we need two terms and a property. So we introduce a three place rule which has the effect that

F_{910}*(John, Bill, see women) = John sees more women than Bill.*

The semantic component has to express what is compared; the syntax needs no to do so.

Another rule might compare of two sets in which amount they are involved in a certain property.

F_{911}*(man, boy, come to the party) = fewer of the men come to the party than of the boys.*

One may also compare two individuals for two properties.

F_{912}*(John, Bill, see men, meet women) = John sees more men than Bill meets women.*

These examples do not provide for a treatment of the comparative. They just illustrate the kind of solution we would search for in accordance with the compositionality principles. The other examples in this section make us expect that a solution along these lines will be simpler than the traditional treatment.

REFERENCES

BARTSCH, R., 1976, *Syntax and Semantics of relative clauses,* in GROENENDIJK & STOKHOF 1976.

————, 1978, *The syntax and semantics of subordinate clause constructions and pronominal reference,* in HENY & SCHNELLE 1978.

DOWTY, R.R., 1976, *Montague grammar and the lexical decomposition of Causative verbs,* in PARTEE 1976.

———— , 1978, *Dative movement and Thomasons extensions of Montague grammar,* to appear.

DELACRUZ, E.B., 1976, *Factives and propositional level constructions in Montague grammar,* in PARTEE 1976.

GROENENDIJK, J. & M. STOKHOF (eds.), 1976, Amsterdam papers in formal grammar, vol. 1: *proceedings the Amsterdam Colloquium on Montage grammar and related topics,* Centrale Interfaculteit, University of Amsterdam.

HENY, F. & H. SCHNELLE, 1978, *Selections from the third Groningen Round Table Conference.* (to appear).

JANSSEN, T.M.V. & P. VAN EMDE BOAS, 1977, *On the proper treatment of referencing, dereferencing and assignment,* in Lecture notes in Computer Science 52, *Automata, Languages and programming,* A. SALOMAA & M. STEINBY (eds.), Springer.

LEWIS, D., 1972, *General Semantics,* in DAVIDSON & HARMAN (eds.), *Semantics of natural language,* Reidel. Reprinted in PARTEE 1976.

MONTAGUE, R., 1973, *The proper treatment of Qualification in ordinary English,* in J. HINTIKKA, J. MORAVCSIK & P. SUPPES (eds.), *Approaches to natural language,* Reidel. Reprinted in THOMASON 1974.

PARTEE, B., 1971, *On the requirement that transformations preserve meaning,* in C. FILLMORE & D.T. LANGENDOEN, *Studies in Linguistic Semantics,* Holt, Rinehart & Winston.

__________, 1973, *Some transformational extensions of Montague grammar,* Journal of Philosophical Logic 2, 509-534, Reprinted in PARTEE 1976.

__________, 1975, *Montague grammar and transformational grammar,* Linguistic Inquiry 6, 203-300.

__________, (ed.), 1976, *Montague grammar,* Academic Press.

__________, 1977, *John is easy to please,* in ZAMPOLLI (ed.), *Linguistic Structures Processing,* North Holland.

__________, 1978, *Montague grammar and the well-formedness constraint.* (to appear)

RODMAN, R., 1976, *Scope phenomena, "movement transformations", and relative clauses,* in PARTEE 1976.

THOMASON, R.H. (ed.), 1974, *Formal Philosophy,* Selected papers of Richard Montague, Yale University Press.

__________, *Some extensions of Montague grammar,* in PARTEE 1976.

SEMANTIC FIELDS AND OTHER PROBLEMS IN WORD SEMANTICS RECONSIDERED *

Peter Rolf Lutzeier

When I spoke last time at this conference in Amsterdam I already saw a strong point in making detailed studies about word semantics within Montague grammar. Since then this point has become more and more obvious. What may generally not have become so obvious is the fact that any attempt to give a semantic treatment of some given expressions in English or German leads us back at the same time to very general questions about the utility of the model theoretic approach for a description of a natural language.

We all know that in model theory we assign a semantic type T to each syntactic category c. The corresponding semantic type to a given syntactic category determines the domain and range of the denotation function of every expression belonging to this syntactic category. However, it is normally unspecified which denotation function we should pick out of the whole class of functions from the given domain into the given range for a given expression. As far as I know no one has ever advocated the view that none of the given expressions should have a fixed denotation function. That is not surprising as one could claim that the difference between a language and a practice of making arbitrary noises is that the former embodies a concept of negation. In other words, the existence of conventions of negation is one

---

* This paper is a rough starting point of more longer lasting work in lexical semantics. I am especially grateful to my wife Elizabeth C. Lutzeier for checking my English of the version which was read at the conference.

minimum requirement for a language.[1] So, if we want to take English or German as a language seriously, we have to make sure that the 'negation' particles not and nicht are interpreted as negation particles.[2] Of course, traditionally all the key words for a language in the framework of predicate logic are given a fixed denotation function. All the successful steps towards pragmatic languages have allowed further contributions concerning, for instance, modal, local and temporal expressions. The designation of these key words, sometimes called logical words, is a necessary condition for arriving at all the logical truths of the language. No doubt, this is an important part of a language, but as linguists we cannot be interested only in knowing whether the truth of a certain sentence is inferred by its logical form. We also want to know the contribution of all meaningful parts of the sentence towards its meaning as a whole, instead of having only the contribution of a few designated words. This even holds for all sentences with the logical form of a logical truth, because there are senseless ones, e.g.

(1) No unmarried nail is married.

So, what the cleavage between logical and non-logical particles boils down to is that you give a list of words and prefixes, call them logical particles, and assume that these items deserve special semantic treatment.

---

1) Compare A.Quinton, The a priori and the analytic, Proceedings of the Aristotelian Society 64(1963-4), 31-54. Reprinted in: Philosophical logic (ed. P.F. Strawson), Oxford University Press 1967, 107-128; p.116-17.

2) Obviously, there may be other functions for not. Think of phrases like: It's beautiful here, isn't it? .

Surely, in doing this one is part of an old tradition and ones excuse for doing so might be that it has always been so useful for determining logical truths.

Still, as we have seen, that cannot be enough for linguistics. Should we not postulate for any interpretation the extreme requirement that each basic[3] expression should have its denotation function fixed? Cresswell explicitly favours this view in saying that every word in a natural language is, in some sense, a constant.[4] Hence, he certainly reminds us of the fact that for a natural language we cannot establish interpretations out of the blue. We have to aim at interpretations which conform to our intuitions about the language, especially to our intuitions about the semantic part. In other words, we should aim at so-called 'intended interpretations'. Obviously, more and more investigations about the semantics of single words can help us to get closer to the intuitions of a native speaker and should therefore be encouraged. But do our intuitions about language turn out to be very clear, when we are forced to write them down? I am not so sure about that, so to me it seems to be questionable whether one should state Cresswell's view as an absolute requirement for any interpretation at the moment. Besides its feasibility, merely adding up what is at most loosely connected information about the meaning of single words would hardly be enlightening about the vocabulary of a language; as, for instance, adding up unrelated information about the meaning of individual sentences could not tell us anything interesting about their relations to each other. If we take seriously the idea that the vocabulary of a natural

---

3) At the moment by using 'basic' I just refer to the status of the expressions within grammar. As Werner Abraham rightly points out one would like to have some outside criteria for that status.

4) M.J.Cresswell, Hyperintensional logic (unpublished) Wellington 1974, p.15-16.

language like English or German is suitable for investigations about relations between its members - which I do -, then concentrating on that idea may be a fruitful alternative to the specific investigation of single words.

One traditional method fitting in with this idea is the method of subclassification, because suclassifications within one syntactic category are semantic groupings within that category which reflect different semantic behaviour.[5] I think it will be a promising step forward towards 'intended interpretations', if we aim at subclassifications within all syntactic categories.

Which criteria do we use in order to establish subclassifications within syntactic categories? If we look into Montague grammar, then we realize that a given class of expressions of a certain syntactic category gets divided into several subclasses according to their different behaviour with regard to the semantic relation 'logical consequence'. Let us look at the class of adjectives and the one of transitive verbs. Differences between the well-known sentence pairs

(2) This tulip is a red flower

(3) This tulip is a flower

(4) John is an alleged murderer

(5) John is a murderer

to the effect that (3) is a logical consequence of (2), whereas (5) is not a logical consequence of (4), show that we have a reason for establishing at least two subclasses for the class of adjectives. The same holds for the class of transitive verbs. Some of them allow

---

5) We may need to qualify the idea that different syntactic behaviour is best dealt by assuming different syntactic categories.

the rule of existential instantiation with regard to their direct objects; others do not. Compare Montague's pair find and seek in the following sentence pairs

(6) Rolf finds a proof for Goldbach's assumption

(7) There is something which Rolf finds

(8) Rolf seeks a proof for Goldbach's assumption

(9) There is something which Rolf seeks

(7) is perfectly alright as a logical consequence of (6), whereas in one reading of seek (9) is not a logical consequence of (8).

An important question arises now: Are there other criteria in addition to logical consequence? Syntactic criteria are excluded, because we started on a given syntactic classification.[6] In other words, we are looking for further semantic criteria or possibly the pragmatic criteria which have to be taken into account in semantics too, for instance, criteria involving speech acts. But may it not be that all semantic relations are reducible to logical consequence? Lyons, for instance, describes all semantic relations which he treats in terms of implication.[7] This seems to hold also for the semantic understanding of the notion 'presupposition'.[8]

A good candidate for another criterion for subclassification might be the intuitive notion 'similarity'. Certainly, this notion still needs a lot of study, but I doubt if one can cover it by means of the logical consequence. I expect that a close look at semantic fields will

---

6) See footnote 5.

7) J.Lyons, Introduction to theoretical linguistics Cambridge University Press 1968, chapter 10,p.443-481 or in his new book: Semantics I, Cambridge University Press 1977, chapter 9,p.270-335.

8) Compare E.Keenan, Presupposition in natural logic, The Monist 57(1973),p.344-370.

give us useful insights into 'similarity', because this notion must have played a major role on the intuitive level for the constitution of semantic fields. So, for instance, its members are able to occur in the same verbal context and have to be regarded as being similar to each other at least with respect to this. The study of semantic fields may lead to a better understanding of similarity and therefore be helpful for the problem of subclassification, which, as we have seen, has to be tackled in word semantics. This prospect should be reason enough to deal in more detail with semantic fields. A few other reasons may become clear during my lecture. In addition, a precise idea about semantic fields would make available the vast material about word semantics of traditional linguistics and I would be pleased to support a convergent trend between more logical interests within linguistics and more traditional ones.

I am not interested in an exegesis of the original understanding of semantic fields. What I am interested in is a notion of semantic fields which will give us the opportunity of incorporating into our theory well-established traditional facts in word semantics in a more precise and satisfying way as well as the opportunity of finding new results within word semantics.

Intuitively we know that semantic fields are special paradigms deriving from verbal contexts. The members of a semantic field represent choices expressive of intentions which are possible in the given verbal context. In other words, a semantic field is the repertoire of choices in the language with respect to a given verbal context. There will be intentions which are possible in the given verbal context for which we cannot find any suitable member of the semantic field or for which we cannot find a unique member of the semantic field. Uniqueness, of course, is only guaranteed up to synonymy. So, quite often, several, non-synonymous members can express one intention, especially if it is not a very precise one. These are reasons why there will not normally be for a semantic field in language A such a semantic field in language B that it is a 1-1 correspondent to the one in language A, quite independent of whether we find synonymous verbal contexts in the two languages.

The intention to express something special within a certain verbal context must not be confused with the actual motive of speaking. Whether, for instance, the speaker really wants to deceive the hearer by uttering something specific is not relevant for the question of which choice the speaker should make in order to express in the best way his intention to utter something specific. In other words, only the adequacy on the expressive level with regard to his intention counts not what he really wants to achieve with his utterance. That is why semantic fields are definitely part of semantics. Still it is only the intention of the speaker that counts, therefore the question of the truth value of the possible utterance does not come in either. So much for intuitive remarks. I shall look now at the notion 'semantic field' in more detail. The examples will be given in German, because I take them from German.

First, we must be clear what sort of paradigm we are aiming at. It is the notion 'paradigm' which is determined by a verbal context and it should be suitable for our notion 'semantic field'. So, what sort of verbal context do we allow for our notion 'paradigm'? Let us examine the following examples:

(1) ...ot / r,t,...

(2) sag... / e,st,t,en,t,en

(3) ...schlafen / aus, ein, weiter,...

(4) ... macht mich krank / Schlafen, das Wetter,...

(5) die ... macht mich krank / Krise, Sorge, Frau,...

(6) ein ... Mann / armer, guter, kranker,...

(7) ein Mann, der ... / schläft, lügt,...

For semantic fields we are not interested in phonological paradigms or ones which consist of inflections. Therefore, we want to rule out context (1) and context (2). The rest of the examples are alright, because we do not require that the substitution of an element of the paradigm yields a sentence. This would be an unnecessary restriction as example (6) shows. However, all syntactic concatenations for the verbal context should be done with the aim to get a sentence, which the ones in our examples obviously are. Therefore, we are interested in the following notion 'verbal context':

Definition: Let us add to the basic words one designated element X (X stands for the empty place where the substitutions can be made).
α be a concatenation of basic words and X.
α is a verbal context iff there is a sentence s such that there is an α' which derives from α by substituting elements of sentence s for X and which occurs in s.

Given a sentence, α' is normally not uniquely determined.

Example to our definition: α := ein Mann, der X

(Compare (7))

We want to show that α is a verbal context.

s:= ein Mann, der morgens aus dem Bett fällt, fängt den Tag schlecht an is a sentence.

α':= ein Mann, der morgens aus derives from α by substituting elements of s for X. α' occurs in s, therefore α is a verbal context.

Roughly speaking, it is clear what we can substitute for X in order to check whether a concatenation of basic words and X is a verbal context. It is not clear yet what paradigms consist of. In other words, the question is which items can replace X in a given verbal context in order to arrive at a paradigm? We understand semantic fields as special paradigms, so the answer to this question will be influenced by our conception of semantic fields. We only want to have syntactic entities as members of paradigms. That is not surprising if we remember the following: semantic fields are supposed to be a help in word semantics by revealing relations within the vocabulary of a language, so semantic entities like concepts must not be members of semantic fields.

Basic words are obvious candidates for being members of our paradigms. I would like to point out that the syntactic category of the words to be substituted for X in a given context will not always be uniquely determined by this verbal context. Let us illustrate this fact by the following example:

Hans läuft X is a verbal context. With regard to the simplest distribution of syntactic categories we have two possibilities for the syntactic category of words to

	e	t/e	
1.	t		t/t
2.		t/e	t/e/t/e

be substituted for X: t/t (one place sentence operator) and t/e/t/e (predicate modifier). We get as paradigm: schnell; möglicherweise; wahrscheinlich; gut; langsam;...

Should we allow anything else besides basic words? What about phrases, in other words, concatenations of basic words according to syntactic rules? Compare the following examples:

(1) das Bonbon klebt an der X / Schachtel; Innenseite der Schachtel; frisch gestrichenen Türe;...

(2) das Bonbon klebt X / außen; inwendig; an der Bettdecke; an dem Punkt, an dem Otto erschossen wurde;...

(3) Hans läuft X / schnell; sehr schnell; zu schnell; gemäßigt schnell;...

(4) das Buch ist X / vor zwei Monaten erschienen; soeben auf den Markt gekommen;... .

Again, the decision about having phrases or not depends on our idea of what semantic fields should serve for. If we understand semantic fields as something which could throw some light on the lexical level of a language, then we should not accept bigger phrases in a paradigm. Not surprisingly, I agree with this view, because I see the role of semantic fields in word semantics, for instance, as a guide for subclassifications. Nevertheless, we must not only allow basic words for paradigms; we also have to allow the use of morphological rules and the use of a few syntactic rules, like the use of the rule for combining a noun with the definite or indefinite article. Compare our example (4), where we would like to have phrases like: ein Plagiat, ein Skandal, der Knüller, die Masche. If we did not allow such rules and only asked for basic words, then, for instance, we could not get nouns with different gender in one paradigm relative to a verbal context like (4). The few rules can easily be listed as soon as we have written down the morphological and syntactic rules for a given fragment of German. As example (3) shows, members of the paradigm itself create new verbal contexts; in our example:

(3') Hans läuft X schnell.

This, however, is only of interest for further, new paradigms but not for the original one, because having these bigger phrases like: zu schnell as members in a paradigm would not contribute anything to our questions on the lexical level.

So far, we have got for our notion of paradigm syntactic criteria, first, because paradigms are determined by verbal contexts and verbal contexts were practically based on the criterion of well-formedness, and secondly, because as members we allow basic words and any phrase built by rules from a specified list. However, given a verbal context something still has to govern the substitution process. Compare our example:

(5) Unser Schimmel ist X / ein Klassepferd; ein Ass; elegant; stattlich; schwarz; ein Mensch; vierbeinig; gefleckt;... .

Presumably, schwarz should be excluded from the paradigm for a standard version of German, because it would lead to a contradictory sentence in that version. We already realized earlier on that truth cannot be a criterion, even in an intuitive sense. Still, if we do not want to fall back on the traditional criterion of actual occurrence in a text, which may be important for the possibility of finding new results and for realizing potentials of the language, we have to allow an intuitive notion of possibility for the substitution as a criterion.

No doubt, that is a point which will need further thought and I certainly would like to invite comments on that part as I do for any other part of my lecture.

Because of some vagueness at this point, mainly with regard to the notion 'possibility', I can only give a provisional version of a definition for the notion 'paradigm'.

Provisional definition: α be a class of basic words or phrases built by rules from a specified list.

α is a paradigm iff there is a verbal context c such that for all members m of α there is a possible sentence s such that c occurs in s if we substitute m for X in c.

'Possible sentence', of course, is meant semantically and not syntactically. By the way, we need this extra clause about the existence of a sentence s because we allow 'incomplete' verbal contexts like ein X Mann. One can easily see from the provisional definition that a given verbal context can create different paradigms. Some of them may be ordered by inclusion and others will not. The union and the intersection of two paradigms created by the same verbal context will also yield a paradigm created by this verbal context. It is clear as well that one paradigm can be created by different verbal contexts. A trivial example would be to replace each word in the verbal context by a synonymous word or phrase:

(a) Stefan sucht nach einem X

(b) Stefan ist auf der Suche nach einem X.

I turn now to the notion 'semantic field'. The way we defined the notion 'paradigm' was in order for it to be suitable for a sensible notion of semantic field. Nevertheless, I have not yet solved quite a few of the problems concerning semantic fields, so at the moment I could not come up with a satisfactory definition. I shall need an extended empirical phase, especially with paradigms which are considered to be semantic fields, before I feel competent in giving a tenable definition of semantic fields. So, I prefer to give some hints for a definition and to state some of the problems concerning semantic fields.

We certainly expect from members of a semantic field that they have something semantically in common. We also know that in a paradigm phrases of different syntactic categories can occur. Phrases of different syntactic categories can certainly have something semantically in common, especially when they occur in the same verbal context as the ones of a paradigm. However, members of different syntactic categories would not help in establishing inner semantic relations between one syntactic category. And we really need that in order to get help in subclassification. In other words, in my conception a semantic field only consists of basic words and phrases of one syntactic category.

This is already quite a step forward towards our intuitive idea about semantic fields, but it does not suffice yet. The restriction to one syntactic category will not keep us from results we would not like to have. I suppose for a semantic field we are not just interested in choices expressive of intentions which are possible in the given verbal context, but we are also interested in choices expressive of intentions which are possible in the given verbal context and which are relative to a certain aspect. Let us illustrate this for the syntactic category 'adjective':

(1) das Buch ist X / alt; interessant; neu; originell; umfangreich; dick; langweilig; faszinierend; erschienen; erhältlich; verkauft; vergriffen; veraltet; ... .

'Relative to a certain aspect' could mean for adjectives what has been called 'relative to a certain dimension'.[9] So, for instance, we would have the aspect: 'Possibility of getting it', which would lead to a semantic field 1:

S 1: alt, neu, erschienen, erhältlich, verkauft, vergriffen

and the aspect: 'value of the content', which would lead to a semantic field 2:

S 2: interessant, originell, langweilig, faszinierend, veraltet.

On the other hand, there are possible aspects which must not restrict our choices expressive of intentions, because actual truth must not play any role for semantic fields. For instance, specific points of reference belong to this class; like Tuesday, 10th January 1978 for the verbal context: Das Wetter ist X, whereas an abstract aspect like 'Time', of course, will not. If we took seriously specific points of reference we would not get interesting semantic fields, because, for instance, antonyms would not appear. By the way, the truth issue also forces us to ignore somehow indexical elements, if they occur in a verbal context. Compare our example:

(2) das Wetter ist heute X / schön, regnerisch, windig, windstill, unliebsam, unwirtlich, angenehm,... .

I shall look now at another problem for the definition of semantic fields. At the same time it will lead to some application of the notion 'semantic field'. I shall consider ambiguous basic words or phrases and ask how they enter semantic fields? In paradigms we are used to having ambiguous phrases neither disambiguated by

---

9) Compare R.Bartsch/T.Vennemann, Semantic Structures Athenäum Verlag Frankfurt 1972,p.67.

the rest of the paradigm nor by verbal contexts creating the paradigm. Compare the basic phrase Pferd in the following paradigm:

(3) Rolf griff nach dem X / Pferd, Schlitten, Buch, Rind, Schaf, ...

Here Pferd can still have all its possible interpretations, such as an interpretation as an animal, as a toy, as a chess figure and as a gymnastic apparatus. Does that hold for semantic fields as well? Yes, it does, as we can realize by looking at the following semantic field 3:

S 3: Pferd, Stute, Rappen, Schimmel, Fohlen, Hengst, Mähre, Gaul.

A basic word like Pferd is not disambiguated by such a semantic field, because besides the obvious interpretation as an animal the semantic field 3 does not exclude an interpretation as a toy. Why is that? Well, all the other members of the semantic field 3 allow this interpretation too and an interpretation for a member of a certain semantic field must be according to the inner structure of the whole semantic field. That is the reason why interpretations as chess figure or gymnastic apparatus for Pferd are out, because other members of the semantic field 3 do not have these interpretations. This line of argument takes into account that items entering a semantic field must have something semantically in common.

I mentioned at the beginning of my lecture the role of similarity between expressions for the constitution of semantic fields. Now we found out that semantic fields can serve for the explication of a further notion of similarity, namely the notion of similarity between interpretations of ambiguous phrases. In other words, a semantic field determines the meaning of its members only up to similar interpretations. In our example we get for Pferd the strong similarity between the interpretations as an animal and as a toy. By the way, the similarity between interpretations is independent of similarities between corresponding objects in reality. For instance, we could have a wooden figure in the shape of a horse as a gymnastic apparatus, nevertheless, the interpretation for Pferd as a gymnastic apparatus would still be excluded by the semantic field 3, simply because the other members do not have such an interpretation in a standard version of German. The other interpretations for Pferd, such as the one as a gymnastic apparatus or as a chess figure, arise from other semantic fields. As examples:

S 4: Pferd, Reck, Barren, Ringe

for the interpretation as gymnastic apparatus and

S 5: König, Bauer, Läufer, Turm, Pferd, Dame

for the interpretation as chess figure. These semantic fields show that there is no such similarity between these two interpretations for Pferd, the same holds with regard to the interpretation as an animal or toy. The line of argument would be analogous to the one earlier on. This notion of similarity may help to clarify the traditionally awkward difference between polysemic and homonymic expressions. Intuitively, for polysemic expressions one sees a connection between its interpretations whereas for homonymic ones one does not see a connection. Semantic fields could perhaps be used at least as a necessary test for homonymic phrases in the sense that if an[10] ambiguous phrase is homonymic then there is no semantic field with the questioned phrase as member which

---

10) 'an ambiguous phrase' is to be understood as follows: If it is homonymic then it is a homograph as well and cannot be differentiated on the graphematic level. Of course we would speak of different phrases in the end.

allows all its possible interpretations. I am reluctant to accept this test as a sufficient one, because I am not sure whether we would like to call Pferd a homonymic word. And Pferd really falls under our criterion: I would claim that in a standard version of German there has to occur another term for gymnastic apparatuses in the semantic field in order to get the 'gymnastic apparatus'-interpretation for Pferd at all. However, such an occurrence excludes at least the 'animal'-interpretation, so there is no semantic field which allows all the interpretations for Pferd. What about kosten which with its interpretations as 'tasting' and 'having a certain price' is traditionally regarded as homonymic word? Typical semantic fields for each interpretation are:

S 6: kosten, abschmecken, probieren, versuchen,...
with der Koch X die Speisen as creating verbal context;
S 7: kosten, verlangen, bezahlen, auszeichnen, feilschen, kalkulieren,...
with das Mädchen X zuviel as creating verbal context.

Although S 7 allows an interpretation with regard to meals or beverages, only the aspect 'value' comes in but not the aspect 'tasting'. In general, any semantic field with kosten as member which should allow for the 'tasting'-interpretation would need more terms out of our typical semantic field S 6 and would therefore automatically exclude the 'price'-interpretation. So, our test would hold for kosten. In addition, the semantic field

S 8: kosten, beanspruchen, verschlingen, benötigen,...
with die Lektüre X zuviel Zeit as creating verbal context

shows that the 'having a certain price'-interpretation has not only to be interpreted with regard to money but can be seen in more general terms too. We can draw two conclusions out of this: first, the interpretations, for instance, developing from semantic field 8, are similar to each other and therefore the split up phrases of a homonymic phrase can themselves be polysemic.[11] This point depends on the following idea: Semantic fields may be used as a sufficient test for polysemic phrases as follows: If a semantic field which has the phrase as a member allows several interpretations for this phrase, then the phrase is polysemic. Obviously, this criterion leads to a fairly weak notion of polysemy. Secondly, the 'money'-interpretation is the basic one of all these 'price'-interpretations, because there is a semantic field, for instance S 7, which favours this interpretation compared to the others. A generalization of this would be: Let α be a polysemic phrase with the interpretations $I_1, \ldots, I_n$. I_{k_1} ($k_1 \in \{1, \ldots, n\}$) is more basic than I_{k_2} ($k_2 \in \{1, \ldots, n\} \setminus \{k_1\}$) iff there is a semantic field which allows for α the interpretation I_{k_1} but not I_{k_2}, but there is no semantic field which allows for α the interpretation I_{k_2} but not I_{k_1}.

These last remarks about further applications for a more precise notion of semantic fields may suffice for today. I hope that they have shown that semantic fields will not only be helpful for methodological questions like 'how to proceed in word semantics' but that they can also help to clarify traditional problems in word semantics, like the polysemy versus homonymy distinction. These questions touch theory of grammar as well as theory of language. In addition, the insights about the two notions

---

11) Example: kosten in the 'price'-interpretation is a polysemic word.

of similarity- the one between members of a semantic field and the other between interpretations of a lexical item - should be taken into account in philosophy of language. Enough reason, one should think, for contuinuing the work on semantic fields and related topics.