PROCEEDINGS OF THE '87 DEBRECEN SYMPOSIUM ON LOGIC AND LANGUAGE

EDITORS

I. RUZSA AND A. SZABOLCSI

1828

AKADÉMIAI KIADÓ * BUDAPEST

This volume constitutes the Proceedings of the '87 Debrecen Symposium on Logic and Language, held in Debrecen, Hungary, from August 25 to 28, 1987, under the joint auspices of the Department of Symbolic Logic and Methodology of Science of the Loránd Eötvös University and the Institute of Linguistics of the Hungarian Academy of Sciences.

These proceedings comprise the texts, sometimes revised, of almost all speakers of the Symposium. Also, two papers not read at the Symposium but pertinent to its topic are included.

AKADÉMIAI KIADÓ

BUDAPEST

1987

PROCEEDINGS OF THE '87 DEBRECEN SYMPOSIUM ON LOGIC AND LANGUAGE

EDITORS

IMRE RUZSA and ANNA SZABOLCSI

DEPARTMENT OF SYMBOLIC LOGIC
L.EÖTVÖS UNIVERSITY, BUDAPEST

INSTITUTE OF LINGUISTICS
HUNGARIAN ACADEMY OF SCIENCES

1828

AKADÉMIAI KIADÓ, BUDAPEST 1987

ISBN 963 462 238 0

PRINTED IN HUNGARY

PROCEEDINGS
OF THE '87 DEBRECEN SYMPOSIUM ON
LOGIC AND LANGUAGE

LOGIC AND LANGUAGE

SYMPOSIUM HELD FROM AUGUST 25 TO 28, 1987, IN DEBRECEN, HUNGARY

CONTENTS

PREFACE

The friendly collaboration of linguists and logicians in Hungary had begun over a decade ago. The occasional informal exchanges between the two parties look a more institutionalized form in 1986 under the auspices of the Institute of Linguistic of the Hungarian Academy of Sciences (thanks personally to Vice-Director Ferenc Kiefer) and the Department of Symbolic Logic and Methodology of Science of the L. Eötvös University. As the first result of the joint project, a two-day symposium was held in the city of Debrecen in September 1986. Most papers read at that symposium are published in No. 4 of *TERTIUM NON DATUR*, the annual of the Department of Symbolic Logic (in Hungarian).

The participants of the 1986 symposium decided that such meetings between linguists and logicians should continue with some regularity. The second meeting was held in August 25 through 27, 1987, with 19 participants from Hungary and 15 participants from abroad.

The '87 Debrecen Symposium on Logic and Language was sponsored by the Hungarian Ministry of Culture and Education and the Institute of Linguistics of the Hungarian Academy of Sciences. Thanks are due to our colleagues Katalin Bimbó, András Máté, Tamás Mihálydeák and László Pólos for their invaluable help with the organization of the symposium.

Seventeen out of twenty contributions to the '87 Symposium were submitted in time for us to include them in this volume. The collection is supplemented with the papers by A. Madarász and I. Ruzsa, not read at the Symposium but pertinent to the topics discussed therein.

The structure of the volume reflects the topical coherence of the contributions, with some global as well as local transition from "more linguistics" to "more logic".

The papers by VAN DER AUWERA, HUNYADI, MARÁCZ and KAŃSKY discuss problems related to the representation and interpretation of sentences involving quantifiers, negation and reciprocals in grammar. The contributions by DE MEY and LÖBNER complement these from the perspective of generalized quantifier theory. KÁLMÁN, PÓLOS and RANTALA go beyond the level of sentences by addressing specific problems in discourse representation and narratives. In another group of papers, CASADIO, SZABOLCSI, STRIGIN and RANTA discuss categorial grammar, combinators and type theory. PEREGRIN proposes a more general notion of intersubstitutivity than intensional logic. Finally, the papers by MADARÁSZ, WUTTICH, SMIRNOVA, WESSEL and RUZSA are primarily of logical character. They are concerned with problems belonging to "non-standard" semantics, including semantic games, semantic value gaps, overlap of truth and falsity, epistemic sentences and nonclassical negation; a set of issues relevant for the logical reconstruction of natural linguistic expressions.

Last but not least, we wish to express our gratitude to Katalin Bimbó for her extensive help in preparing this volume.

THE EDITORS

JOHAN VAN DER AUWERA

ARE ACTIVES AND PASSIVES TRUTH-CONDITIONALLY EQUIVALENT?

1. Introduction

What I will henceforth call the 'Equivalence Thesis' or 'ET' is the view that an active sentence has the same truth conditions as its corresponding passive sentence; they refer to the same state of affairs. Thus if (1) is true, then so is (2), and vice versa.

(1) John kisses Mary.

(2) Mary is kissed by John.

ET is a traditional point of view. For some time, it was also endorsed by Transformational Grammar, viz. during the period that the semantic interpretation was taken to apply to the underlying structure only. (1) and (2) were supposed to have an identical underlying structure. After the disappearance of Generative Semantics, however, ET was given up, primarily because of problems with quantifiers and negation[1]. Consider (3) and (4).

(3) Many arrows didn't hit the target.

(4) The target wasn't hit by many arrows.

Suppose that half of the arrows didn't reach the target. Then (3) seems true, and (4) false. Or take (5) and (6).

(5) Each student admires no politician.

(6) No politician is admired by each student.

Abstracting from any effects of special intonation, (5) seems to be a marked way of conveying that no student admires any politician, while (6) says that there isn't a single politician that is admired by each student.

Note that one doesn't have to be a transformationalist to give up ET. The most explicit rejection of ET is actually that of the non-transformationalist Keenan (1981), and a closely analogous view of the passive is also found in Generalized Phrase Structure Grammar (Gazdar e.a. (1985)). In the theory of Functional Grammar, however, as started by Dik (1978) (see also, De Groot (1985)), ET is still accepted, but the problem is that ET has not been confronted with the quantification and negation problems such as demonstrated in (3) and (4). That then is the main purpose of this paper.

2. Term Operators

Since the problems in (4) to (6) concern quantification and negation, I must briefly sketch how FG deals with them.*

Quantifiers like <u>many</u> in (3) are considered to be part of the NP. The same goes for the quantifier <u>all</u> in (7).

(7) <u>All</u> human beings are mortal.

While this may seem evident, we should not forget that predicate logic proceeds differently: the quantifiers would be outside of the clause or, in the jargon, outside of the sentential function.

(8) $\forall x\ (HB(x) \rightarrow (M(x))$

It is only recently that logicians have been exploring ways to respect the essential NP-internal nature of the quantifier- Generalized Quantifier theory (Van Benthem & Ter Meulen 1985). FG respects this in a different way: a quantifier like <u>all</u> has the same status as an article, a numeral, and even the number of the noun: they are all operators of the NP. An NP is called 'term' and thus

* 'FG' stands for 'Functional Grammar'.

the operators are 'term operators'.

As to negation, consider (9) and (10).

(9) He doesn't own books.

(10) He owns no books.

In (9) the negation operates on the predicate, whether it be own or the auxiliary do; in (10), however, it operates on a term (books). This, at least, seems to be the superficial structure. Of course, one can neglect this difference and propose an identical underlying structure, something like (11).

(11) ~ (he owns books)

In FG, however, respecting the surface structure is a principle of high importance. Thus even in the underlying structure of (10) the negation is a term operator, just like the quantifier all in (7), and in the underlying structure of (9), it is a predicate operator, just like time, aspect, and modality operators. Naturally, FG will have to explain why (9) and (10) have identical truth conditions, concretely, why a predicate operator can have a term in its scope. This gets done along the following lines. A predicate is not an isolated item, rather, it is the centre of what is called a 'predicate frame'. This predicate frame specifies the terms that can function as arguments. The predicate frame of the predicate give, for instance, will enumerate three arguments, an Agent, a Patient, and a Recipient. Specifying the combinatorial possibilities of these predicate frames, called 'nuclear', with 'satellites' is also done in a predicate frame, called 'extended'. Thus every predicate operator is always also a predicate frame operator and this way terms can fall in the scope of predicate operators. This, surely, is the intuition that (11) was to express, without, however, expressing the close link between negation and predicate.

Note also that nothing prevents me from abstracting from the difference between term and predicate and thus arriving at the conclusion that at THIS level of abstraction the negation of (9) and (10) are the same. Only, I would still maintain that the negations are different at the more concrete level.

3. Specificity and distributivity

Let's get back to the active-passive pairs that threaten ET.

(3) Many arrows didn't hit the target.

(4) The target wasn't hit by many arrows.

(5) Each student admires no politician.

(6) No politician is admired by each student.

I accept that (3) and (5) do not have the truth-conditions of (4), respectively (6). One can indeed side with latter-day transformationalists or with Keenan and reject ET. But this is not the only possible line of action. One could also try to develop the idea that (4) and (6) are not, despite appearances, the passives of (3) and (5). This approach is not new; it was used by Generative Semanticists. In Generative Semantics underlying structures, quantifiers and negations were placed outside of the 'clause proper', viz. in a higher clause. It is not denied that they occur within NPs in surface structure; but they reach these NPs only through a transformation of 'quantifier lowering'. The relevant point now is that the order of quantifier and negation in the underlying structure was different for (3) and (4), resp. (5) and (6). In (3) and (5) the quantifiers would precede the negations; in (4) and (6) it would be the other way round.

This approach, including the 'quantifier lowering' trick, has been given up, and justly so, and it is also impossible to reintroduce it in FG. <u>many</u>, <u>each</u>, and <u>no</u> belong with their NPs from the very start (from the underlying representation) - they are term operators.

Yet ET can be maintained in another way. It should be obvious that ET was conceived to characterize the relation between an active sentence and its passive 'counterpart'. The passive should thus be identical to the active, except for the passive - active distinction. This means that both sentences should have identical terms and predicates as well as identical term and predicate operators. The hypothesis I will now propose is that (3) and (4) and, again, (5) and (6) do NOT have identical terms, more specifically, that they do not have identical term operators. The relevant operators are those of distributivity and specificity.

Specificity characterizes a term as [± specific]. The distinction is typically illustrated with sentences such as (12) and (13).

(12) Mary wants to marry a Norwegian.

(13) Everybody in this room speaks two languages.

(12) is ambiguous: either Mary wants to marry a specific man and it so happens that he is Norwegian, or she doesn't yet know who she'll marry, although it is necessary that this man be Norwegian. In this first reading a Norwegian is [+specific], in the second [-specific]. It is only in the [+specific] reading that one can continue the sentence as in (14).

(14) Marry wants to marry a Norwegian, but you shouldn't confuse him with this other Norwegian, who she doesn't want to marry.

two languages in (13) is ambiguous, too. If the two languages are necessarily the same for every individual in the room, then one has a [+specific] reading. If the languages are not necessarily the same, one obtains the [-specific] reading. The continuation of (15) is available only for a [+specific] reading.

(15) Everybody in this room speaks two languages, but you shouldn't confuse them with this other pair of languages, which nobody speaks.

The second relevant operator is that of distributivity, illustrated in (16).

(16) The kids jumped on their horse.

(16) is ambiguous: either the kids jumped on one and the same horse and then <u>their horse</u> is [-distributive], or the horse the kids jumped on is not necessarily the same - in other words, the kids may have jumped on at least two horses, and then <u>their horse</u> is [+distributive]. Note that distributivity allows us to distinguish between two types of non-specificity: <u>a Norwegian</u> is [-specific] and [-distributive], whereas <u>two languages</u> is [-specific] and [+distributive].

A crucial point is that a term in isolation of the type <u>a Norwegian</u> is ambiguous and that the clause may disambiguate, but needn't. In (12) <u>a Norwegian</u> is ambiguous for specificity, but not for distributivity. In (17) <u>a Norwegian</u> is ambiguous for both operators.

(17) Everybody in this room wants to marry <u>a Norwegian</u>.

Let us now examine sentences (3) to (6) and in particular the terms <u>many arrows</u> and <u>no politician</u>. For <u>many arrows</u> only specificity is relevant: I claim that <u>many arrows</u> is [+specific] in (3), but [-specific] in (4). This is easily verified by the fact that only (3) can be continued in the way of (14) and (15).

(18) Many arrows didn't hit the target, but you shouldn't confuse them with the many arrows that did hit the target.

(19) *The target wasn't hit by many arrows, but you shouldn't confuse these with the many arrows that the target wasn't hit by.

For <u>no politician</u> both specificity and distributivity enter the picture. In (5) there is a relation between each student and a set of politicians and it is not required that this set is identical (for each student). This is analogous to the relation between everybody in this room and a set of two languages, not necessarily identical either (for each individual). In this reading two languages is [-specific] and [+distributive]. So I propose the same operators for no politician. It is true that the sets of politicians DO end up being identical, the reason being that they are empty. However, non-specificity does not forbid identity, it merely does not require it. So even in a [-specific] reading of (13) the two languages known by everybody in the room could be identical, the point of its non-specificity being that this identity is not obligatory.

Like (5), (6) involves a relation between a set of politicians and a set of students, but unlike (5), (6) requires the set of politicians to be identical. This is analogous with the relation between everybody in the room and one and the same set of two languages. In this reading two languages is [+specific] and [-distributive]; so I propose the same analysis for no politician. It is true that (6) can't be continued in the way of (14), (15), and (18), but there is a good reason for this: there simply can't be any confusion about the [+specific] set of politicians of (6): it is empty and thus unique.

If the above analyses are correct, many arrows is thus [+specific] in (3) and [-specific] in (4), while no politician is [-specific] and [+distributive] in (5) and [+specific] and [-distributive] in (6). (3) and (4), respectively, (5) and (6) do not therefore have identical terms, (4) and (6) are not the passives of (3), respectively, (5) and they cannot refute ET.

One could object to the above strategy that it does not yet explain WHY (4) and (6) are not the passives of (3), respectively (5) or WHY terms are sometimes [+] or [-specific] and how this hangs together with word order and intonation. This is true, but it cannot be a serious objection. (S)he who gives up ET will have exactly analogous problems. (S)he will have to specify the conditions under which simple actives and passives such as (1) and (2) do end up being truth-conditionally identical and why some less simple ones do not. The burden to explain why, say, no politician is [-specific] in (5) but [+specific] in (6) is just as heavy.

4. Perspectivization and Predicate Formation

To accept ET is to accept that actives and passives are equivalent with respect to their truth conditions. This is obviously not the same as to say that actives and passives are equivalent in ALL respects. FG addresses this question in an explicit way and it is proposed that actives and passives present one and the same state of affairs from a different perspective.[2] Thus (1) presents the situation of John's kissing Mary from the perspective of John, and (2) from Mary. Talking about one situation from two different perspectives is like taking a picture of it from two different angles.

The perspectivization hypothesis answers the question why a language should have the active - passive distinction at all. If one rejects ET, one obviously has to answer this question, too. Let us now see how Transformational Grammar, Keenan, and Generalized Phrase Structure Grammar fare in this respect.

First of all, I do not think that Transformational Grammar answers the question. Whatever the value of the Transformational efforts to give a precise, formal characterization of the passivization or 'NP Movement' rule, the theory says nothing as to why languages should have such a rule in the first place. Interestingly, the one feature of the transformational account that reappears in FG is that passivization is a clause level phenomenon. In Transformational Grammar, passivization is a clause level rule; in FG, passive choice is the choice of a perspective on the situation referred to by the entire clause.

Secondly, for Keenan and for Generalised Phrase Structure grammar, passivization is a kind of lexical rule. It is rule that has a predicate phrase as input and another yet related predicate phrase as output.[3] What is of relevance here is that this kind of approach is perfectly compatible with FG, and has already been suggested, in particular, by Vet (1985). All while accepting the FG view that the choice of passive vs. active is the choice of a different perspective on an identical situation, Vet (1985) argues for treating passivization as a 'predicate formation rule', i.e. a rule that has a predicate frame as input and another, related one as output.

5. Conclusion

In this paper I have argued that FG allows one to hold on to the view that actives and passives have identical truth conditions. This view is part of a more general account of the relation between active and passive. While ET is about the similarity between actives and passives, a perspectivization hypothesis deals with the difference. The perspectivization hypothesis also answers the question why languages should have the active - passive difference at all. In other theories such an answer is either lacking, in the case of Transformational Grammar or, if there is one, as for Keenan and Generalized Phrase Structure Grammar, it can easily be integrated into FG, too.

NOTES

1 As pointed out to me by Zbigniew Kański, the other problem that threatened ET was that the English present perfect was taken to presuppose the present existence of the subject, thus allowing (a), but not, or less clearly so, (b).

(a) Princeton has been visited by Einstein.
(b) Einstein has visited Princeton.

For an adequate analysis and references, see McCoard (1978, 60-64).

2 This was essentially also that of Jespersen (e.g. 1924, 167).

3 Because the input is a predicate PHRASE rather than a predicate, it is not a true lexical rule, but only 'a kind of'.

REFERENCES

De Groot, C. (1985) Functional Grammar Publications, Tilburg University, Tilburg.

Dik, S.C. (1978) Functional Grammar, North-Holland, Amsterdam.

Gazdar, G. e.a. (1985) Generalized Phrase Structure Grammar, Blackwell, Oxford.

Jespersen, O. (1924) The Philosophy of Grammar, Allen & Unwin, London.

Keenan, E.L. (1981) "Passive is Phrasal (not Sentential or Lexical)", in: T. Hoekstra e.a., eds, Lexical Grammar, Foris, Dordrecht.

McCoard, R.W. (1978) The English Perfect: Tense-choice and Pragmatic Inferences, North-Holland, Amsterdam.

Van Benthem, J. and A. ter Meulen, eds. (1985) Generalized Quantifiers in Natural Language, Foris, Dordrecht.

Vet, C. (1985) "Passive, reflexive, and causative predicate formation in French", in: A.M. Bolkestein e.a. eds, Predicates and Terms in Functional Grammar, Foris, Dordrecht.

BELGIAN NATIONAL SCIENCE FOUNDATION
& UNIVERSITY OF ANTWERP (UIA)

LÁSZLÓ HUNYADI

ON THE LOGICAL ROLE OF STRESS

In several languages stress has an important paradigmatic function: a difference in the position of stress in the same morpho-phonological sequence can result in different lexemes (cf. Russian mu'ka 'flour' and 'muka 'suffering'), or in different parts of speech of the same derivation (cf. English re'cord as a verb and 'record as a noun), or in different paradigmatic forms of one and the same lexeme (cf. Russian storo'ny as gen. sg. of storona 'side' and 'storony as nom. (acc. pl. of the same noun), etc. Hungarian, on the other hand, has a fixed paradigmatic stress always on the first syllable, and consequently, one cannot find the above features of stress there. Hungarian, in its turn, has elaborated a quite consistent syntactic system of stress that also has its clear logical functions. In what follows here, I will point out some of the important features of this system.

1. Neutral and non-neutral sentences are distinguished both syntactically and on the basis of stress. A neutral sentence has an only relatively fixed word order with more than one

constituent carrying (even) stress (for more detail cf. Kálmán et al.), as in (1):

(1) 'Péter 'el- ment a 'színházba.
Peter away went the theatre-to
'Peter went to the theatre'

A non-neutral sentence, on the other hand, has a distinct pre-verbal syntactic position for focus normally taking a heavy, distinctive stress, contributing to a relatively configurational character of the language (what is less configurational is the composition of topic as well as the post-verbal part; for more detail cf. É. Kiss 1979, 1981, 1987, Szabolcsi 1981), as in (2):

(2) "Péter ment el a színházba.
Peter went away the theatre-to
'It was Peter who went to the theatre'

(Without going into the details of non-neutral stress patterns, I will only note two — not contradictory but additional — cases: a) the focussed element can be preceded by a heavily stressed universal quantifier, in which case the focussed element may be unstressed, cf. (3)

(3) "Mindig (")Péter ment el a színházba.
always Peter went away the theatre-to
'It was always Peter who went to the theatre'

and b) a post-verbal element can also take heavy stress with the sentence now having two focus-like elements, cf. (4)

(4) "Péter ment el a "színházba.
Peter went away the theatre-to
'It was Peter who went — to the theatre')

2. Apart from this clearly syntactic function, stress in Hungarian also has an interesting logical function (not independently from its role in syntax, however). The presence or lack of heavy stress on the quantifier valamennyi (a composed formation from vala 'was' (appr.) and mennyi 'how much/many') determines whether it is a universal (stressed) or an existential (unstressed) quantifier, cf. (5) and (6), respectively:

(5) "Valamennyien eljöttek.
Q came
'All of them came'

(6) Valamennyien "eljöttek.
Q came
'Some of them came'

This is not only a property of stress with valamennyi, since an analytically universal quantifier, such as mindenki 'everyone' from minden 'every' and ki 'who' necessarily takes heavy stress in (7) (that is why (8) is out):

(7) "Mindenki eljött.
everyone came
'Everyone came'

(8) * Mindenki "eljött.

At the same time, an analytically existential quantifier, such as valaki 'someone' from vala 'was' (appr.) and ki 'who' necessarily lacks heavy stress, cf. (9) and (10):

(9) Valaki "eljött.
someone came
'Someone came'

(10) * "Valaki eljött.

The analiticity of mindenki 'everyone' and valaki 'someone' (and other quantifiers with these first compounds minden- and vala-) seems to be transparent indeed: mindenki = 'every(one) who...' and valaki = '(there) was (one) who...'. It remains, however, unanswered why one vala-quantifier, valamennyi has lost its analiticity.

3. The above function of stress was to distinguish between logical interpretations of one and the same morpho-phonological sequence, in which sense this function is somewhat related to the paradigmatic function of stress in other languages. There is, however, a clearly syntactic function of stress as well, logically interpretable as marking scope relations. Phonologically speaking, a scope-taking element has wide scope if it is heavily stressed and it has narrow scope if it loses its stress by passing it to another scope-taking element.

(From this it follows that an existential quantifier cannot be included into the scope of an element whose wide scope over the quantifier would be marked by passing its stress to it.) Syntactically, since stress in Hungarian is normally pre--verbal, i.e. wide scope is marked by moving out a wide-scope element to the left from its post-verbal position, the normal hierarchy of scope-relations is also left-to-right, cf. (11) and (12):

(11) "Mindig a színházba ment Péter.
always the theatre-to went Peter

'It was always the theatre that Peter went to'

(12) "Nem mindig a színházba ment Péter.
not always the theatre-to went Peter

'It was not always the theatre that Peter went to'

Now we can notice a principal difference between the quantifier mindig 'always' and the negative nem 'not'. Phonologically, mindig does not receive its stress from a following element in order to mark wide scope, since mindig has heavy stress in order to denote universal quantification (or in other terms: since it takes heavy stress, it denotes universal quantification; this is why in (3) a second heavy stress on Péter was allowed, although optionally, since the linear order would mark the same scope relation). On the other hand, the negative particle nem is not in an opposition with a positive particle so that they should be distinguished phonologically, too.

'Nem', in order to have wide scope, has to receive its stress from the element that is to be included in its scope. Thus, the existential valami is included in the scope of mindig 'always' in (13), but it is not negated in (14):

(13) "Mindig olvastam valamit.
always read-I something-acc.
'I have always read something '

(14) "Nem olvastam valamit.
not read-I something-acc.
'I did not read something'

Thus, the fact that (14) does not yield the sense 'I did not read anything' is not based on some peculiar logical property of the quantifier valami; it is rather the result of two interacting factors: a) the existential quantifier is, per definitionem, unstressed, thus unable to pass stress, and b) the negative nem requires stress passed over to it from its scope.

The question may now arise, how to express then 'I did not see anything'. Another, logically possible solution might be to have universal quantification with wide scope over negation. But it does not work in (15) either:

(15) * "Mindent nem olvastam.
everything-acc. not read-I

The analysis of minden may bring us closer to an answer. Minden derives from the connective mind 'both' and, as is known, both behaves like a positive polarity item; it cannot have wide scope with respect to negation. This is exactly the case with its Hungarian equivalent as well. I doubt that we can dig much deeper in this problem, probably we have to accept the existence of polarity items as such and build explanations on them as axioms.

With no more logically possible configurations of the two quantifiers, Hungarian chose the form of 'multiple negation' in the form of (16):

(16) "Nem olvastam "semmit.
not read-I nothing-acc.
'I did not read anything'

Semmi 'nothing', in its turn, can be analyzed in the following way: sem 'neither' derives from a composition of is 'also' and nem 'not' plus mi 'what' added. Thus, (16) is, in fact, a condensed series of conjunctions of negative statements. (For more detail, cf. Hunyadi 1981.)

The fact that the existential quantifier cannot be included in the scope of negation can be seen from the following examples:

(17) Megbuktam, mert "nem tudtam "valamit.
failed-I because not knew-I something-acc.
'I failed because I did not know something'

(18) Megbuktam, mert "nem tudtam valamit.
failed-I because not knew-I something-acc.
'I failed because I did not know something '

The unambiguous scope of the existential quantifier gives the luxory to the quantifier to reinterpret the opposition 'stressed/unstressed' in a unique way: if the quantifier is also stressed (cf. (17)), then this 'something' is not specific, whereas if it is unstressed (cf. (18)), this 'something' is specific. Thus, <u>with the existential quantifier, the opposition 'stressed/unstressed' indicates the opposition 'not specific/specific'</u>.

Since the universal quantifier <u>minden</u> 'everything' is a normally stressed element, it can pass its stress over to the negative <u>nem</u> 'not' to mark narrow scope with respect to negation (as we saw, the lack of stress on the existential quantifier, in its turn, excludes the possibility of narrow scope within negation). This stress-passing being the primary condition for the marking of scope-relation in Hungarian allowing for a relatively free movement of constituents, (20) has the same logical interpretation as (19), although mere linear order would contradict:

(19) "Nem mindent olvastam.
not everything-acc. read-I
'I did not read everything'

(20) Mindent "nem olvastam.
everything-acc. not read-I
'I did not read everything'

As a matter of fact, the stress of a constituent can be passed to a wide-scope element through word-boundaries as well (i.e. it is not restricted to adjacent constituents), cf. (21), which is synonymous with (19):

(21) "Nem olvastam mindent.
not read-I everything-acc.
'I did not read everything'

As we saw in (4) "Péter ment el a "színházba 'It was Peter who went — to the theatre', in certain sentences double contrast can be expressed by two heavy stresses, symmetrical around the verb. Whereas in (4) this second stressed element expresses contrast (similarly to focus), in (22) this second element is a quantifier and contrast is excluded:

(22) "Péter ment el "mindenhová.
Peter went away everywhere
'It was Peter who went — everywhere' (= nobody else went anywhere)

The fact is that mindenhová 'everywhere' is a universal quantifier that normally carries heavy stress, this being the phonetic condition for this logical interpretation. On the other hand, something can only have narrow scope if it is unstressed. (The reverse is not true, as we saw the case of the existential quantifier valami 'something' with negation in (14).) Thus, mindenhová 'everywhere' must have wide scope; the only question is, over what? It seems reasonable to assume a hierarchy of wide-scope elements with negation on top, the universal quantifier following and a non-logical element, focussed, as only having wide scope over elements in non--operator position. According to this, (22) will be synonymous with (23), although it has a reverse linear order; and (24) and (25) will be different simply because of the unstressed universal quantifier in it:

(23) "Mindenhová Péter ment el.
everywhere Peter went away
'It was Peter who went — everywhere' (= nobody else went anywhere)

(24) "Péter ment el mindenhová.
Peter went away everywhere
'It was (only) Peter who went everywhere' (others may have gone to less places if at all)

(25) Mindenhová "Péter ment el.
everywhere Peter went away
'It was (only) Peter who went everywhere' (others may have gone to less places if at all)

It is worth mentioning that a post-verbal stressed universal quantifier with wide scope can only appear in a sentence with a non-negative focus in the pre-verbal focus position. If the first heavy stress is taken by another quantifier (cf. (26)), the second quantifier has no hierarchical advantage; whereas if the focus is negative (as in (27)), the sentence is ungrammatical:

(26) "Mindig Péter ment el "mindenhová.
always Peter went away everywhere
'It was always the case that it was Peter who went everywhere'
(mindenhová 'everywhere' has wide scope over Péter 'Peter' but it does not have wide scope over mindig 'always')

(27) "Nem mindig Péter ment el "mindenhová.
not always Peter went away everywhere

The reason for the ungrammaticality of (27) is that this double stress pattern suggests that mindenhová 'everywhere' has wide scope over negation (since the quantifier does not pass over its stress) whereas the direct word order of these scope bearing elements results in ungrammaticality again, cf. (28):

(28) * "Mindenhová nem mindig Péter ment el.
everywhere not always Peter went away

(Let the reader be reminded that mindenhová has the component mind 'both' that behaves as a positive polarity item, thus we cannot reach deeper in findig the reason for the ungrammaticality of (28) and its counterparts in many other languages.)

Let us now come back to (15) * "Mindent nem olvastam el and (16) "Nem olvastam el "semmit 'I did not read anything'. One might ask whether Hungarian has this multiple negation in (16) (nem 'not' and semmit 'nothing, acc.') because of the restrictions on the universal and the existential quantifiers plus the fact that it has no equivalent of the English anything.

The answer makes the whole situation more complex: yes, Hungarian has <u>akármi</u> and <u>bármi</u> as (in most cases) equivalents of <u>anything</u>. Nevertheless, (29) is ungrammatical with two heavy stresses (as is the pattern of (16) as well as (30) with one stress only):

(29) * "Nem olvastam "akármit.
not read-I anything-acc.

(30) * "Nem olvastam akármit.
not read-I anything-acc.

Thus, we see a significant difference between <u>anything</u> and <u>akármi</u>: the latter can only appear with some (mostly overt) modal operator in a simple sentence. Thus in (31) the quantifier is, following the stress rule, included in the scope of negation (-<u>hat</u>- is a modal operator for 'can'):

(31) "Nem olvashattam akármit.
not read-can-past-I anything-acc.
'I could not read whatever I wanted'

Although the sentence is grammatical, nevertheless it does not have the requisite sense of universal negation (what would be suggested by simply following the glosses). But the reasons are principally semantic, firmly established in the logical structure of this and other Hungarian quantifiers, outside the frames of this paper.

4. As we have seen, Hungarian stress does not have a significant paradigmatic role, its role is rather syntagmatical. This is the way neutral and non-neutral sentences can be distinguished, the syntactic position of focus being phonetically marked, too. On the other hand, stress has at least two clearly logical functions: to distinguish between the universal and the existential quantifier and, a function on the syntagmatical scale, to assign scope-relations. The two functions may interact as in the case of quantification and negation. It is also mainly a matter of stress-restrictions that Hungarian has the so-called multiple negation for the expression of <u>any</u>: comparing English <u>any</u>-sentences to their Hungarian equivalents we have found that a) the Hungarian equivalent of <u>any</u> is more modally bound and b) there is a crucial logical difference in the inner semantics of <u>any</u> and its Hungarian counterpart that is demonstrated on the different assignment of scope with the same stress and sentence pattern.

As a concluding remark it can be emphasized that stress in Hungarian appears to be significant enough to be treated on its right place in the syntax and semantics of the language.

REFERENCES

Hunyadi, L. A kétszeres tagadásról a magyarban. *Általános Nyelvészeti Tanulmányok XV* (1981), 65-73.

É. Kiss, K. *A magyar szintaxis egy transzformációs generatív megközelítése*. PhD. diss., 1979.

É. Kiss, K. Structural relations in Hungarian, a "free" word order language. *Linguistic Inquiry 12*, 185-213.

É. Kiss, K. *Configurationality in Hungarian*. Akadémiai Kiadó, Budapest, 1987.

Szabolcsi, A. The semantics of topic-focus articulation. In: J. Groenendijk *et al.* (eds.), *Formal Methods in the Study of Language*. Amsterdam, 1981.

L. KOSSUTH UNIVERSITY

DEBRECEN, HUNGARY

LÁSZLÓ MARÁCZ

WH-STRATEGIES IN HUNGARIAN: DATA AND THEORY*

1. Introduction

In this paper I discuss some data on question (henceforth WH)-sentences in Hungarian and their consequences for the theory of grammar. Our primary concern will be instances of fronted WH-sentences in Hungarian. I will introduce this construction with the help of the following paradigms from English. Compare:

1a Mary saw John.
b Who saw John?
c Who did Mary see?

In the examples of (1b,c) we find examples of question-sentences derived from the single sentence (1a). In (1b) the subject phrase is questioned, while in (1c) the object phrase is questioned. If the sentence in (1a) is the embedded part of a complex sentence, and we want to question one of its constituents, then the question word must be fronted to the matrix sentence. Compare:

2a You think that Mary saw John.
b Who do you think saw John?
c Who do you think that Mary saw?

In (2b) the embedded subject of sentence (2a) is questioned, while in (2c) the object is questioned. Observe that the question words appear in the initial part, the so-called matrix sentence. This type of construction has been referred to in the literature as the fronted, extracted, or raised WH--construction. It has been noted that these fronted WH--sentences are rather limited in appearance. They may appear only if the matrix verb is a so-called "bridge-verb". Bridge verbs belong in general to the semantic class of verbs of knowing, saying and perception. For example, the verb "brag" does not qualify as a bridge for fronted WH-phrases as we may see from the comparison between (2b) and (3):

3 * Who do you brag saw John?

It may be concluded from this brief introduction that the questioning of an embedded constituent seems to cause more problems in natural language than the questioning of phrases in single sentences. In English a way out of this problem is provided by fronting the embedded question word into the matrix sentence. We have seen, however, that this phenomenon is subject to restrictions. Below I will present some more example sentences. The question arises whether all languages solve this problem in the same way. Taking Hungarian into consideration, the answer to this question is unambiguously: NO. In the literature on this topic two sorts of strategies have been noted for questioning a constituent of an embedded sentence: a strategy analogously to the fronting construction in English; and the so-called MIT-strategy. In section 2 I review some of the properties of the fronted WH-construction in Hungarian that have been discussed in É. Kiss (1981, 1985, 1987) and Horvath (1981, 1986). In section 3 I discuss the MIT-strategy. This strategy of forming embedded questions in Hungarian has been noticed first in De Mey and Marácz (1986). Its main characteristic is that the embedded question word remains in-situ in the Focus-position of the embedded clause. In the matrix sentence a question word dummy appears MIT ("what-ACC")[1] appears which represents the scope of the real question word in the embedded sentence. In section 4 I draw some conclusions from the existence of both types of strategies of embedded questions in Hungarian which are relevant for the study of the theory of grammar. The existence of both embedded question strategies provides empirical evidence for the Correspondence Hipothesis:

> 4. CORRESPONDENCE HYPOTHESIS: whenever there is a syntactic reflex of the assignment of (wide) scope, the dependency involved and fronted WH--constructions obey to a large extent the same governing and bounding conditions.

A consequence of the Correspondence Hypothesis is that there is no need to postulate a separate level for the representation of scope which has been called Logical Form (LF) in the linguistic literature. However, the unification of the properties of fronted WH-constructions and WH-in-situ-constructions has, somewhat disappointingly, NOT been a major tenet of research in the Government and Binding program in recent years (cf. Aoun et al. (1980), Huang (1982), Hornstein (1984), Lasnik and Saito (1984), Aoun (1986), Chomsky (1986), among others). Rather, on the basis of observations made in Huang (1982), to the effect that WH-in-situ in Chinese does not obey locality conditions, it has been concluded that wide scope assignment is not restricted by the Subjacency Condition; and thus Subjacency is not operative at the level of Logical Form (LF), calling the existence of LF into doubt. This is the opposite of the result I will argue for on the basis of Correspondence Effects in Hungarian.

2. Fronted WH-constructions in Hungarian

Consider the following sentences:

```
5a   KIT        gondolsz            *(hogy)
     who-ACC    think-2sg-indef       that
     --         látta          Jánost
                saw-3sg-def    John-ACC
     'Who do you think (*that) saw John?'

b    KIT        gondolsz            *(hogy)   János
     who-ACC    think-2sg-indef       that    John
     látott --
     saw-3sg-indef
     'Who do you think (that) John saw?'

c    KIVEL          gondolod            *(hogy) János
     who-INSTR      think-2sg-indef       that  John
     találkozott --
     met-3sg-indef
     'Who do you think (that) John met?'

6a   MELYIK FIÚT        gondolod        *(hogy)
     which  boy-ACC     think-2sg-def     that
     --   látta         Jánost
          saw-3sg-ACC   John-ACC
     'Which boy do you think (*that) saw John?'

b    MELYIK FIÚT       gondolod         *(hogy)
     which  boy-ACC    think-2sg-def      that
     János  látta  --
     John   saw-3sg-def
     'Which boy do you think (that) John saw?'

c    MELYIK FIÚVAL      gondolod        *(hogy)
     which  boy-INSTR   think-2sg-def     that
     János  találkozott --
     John   met-3sg-indef
     'Which boy do you think (that) John met?'
```

Examples (5-6) are instances of fronted WH-constructions, in Hungarian presented and discussed in Horvath (1981, 1986 chapter 4) and in É. Kiss (1981, 1985, 1987 chapter 3). The acceptability of these sentences is subject to dialectical variation. In fact, most of my informants entirely reject this strategy for question an embedded constituent in Hungarian (cf. also Komlósy (1986)). However, from the literature it is clear that native-speakers report instances of fronted WH--constructions. As É. Kiss (1981) points out, cases of fronted WH-constructions have been discussed by traditional linguists as well, e.g. Zolnay (1926). The phenomenon of fronted WH-

-constructions is especially frequent in the spoken language (cf. also De Groot (1981), Szalamin (1978), Szamosi (1976), and Anna Szabolcsi p.c.). There is a group of informants who accept the sentences in (5-6) only marginally. For this group of native-speakers there is even a preference hierarchy: the acceptability of fronted WH-constructions decreases from (5c,6c) > (5b,6b) > (5a,6a). In (5c,6c) a thematically determined or "oblique" argument is questioned; in (5b,6b) the object; while in (5a,6a) the subject. Apparently the phenomenon of fronted WH-constructions in Hungarian is for some speakers restricted by a so-called "Accessibility Hierarchy" (cf. Comrie (1981) for other examples):

7 oblique > direct object (nominative) > subject (nominative)

This hierarchy defines ease of accessibility to fronted WH--constructions. An oblique argument may be extracted more easily than any other argument of the verb, while a direct object may be extracted more easily than a subject. The question that arises, is: how are embedded questions formed in those cases in which fronted WH-construction is unavailable? We will return to this question in the following section. Let us turn to a discussion of the properties of the sentences (5-6). I will discuss the following more language-specific properties of fronted WH-constructions (cf. 8a-e): and some properties appearing in general with this type of construction (cf. 8f-h).

8a The obligatory presence of the complementizer HOGY ("that")
b The role of the anticipatory pronoun
c Case change of the fronted nominative subject
d Fronted WH-constructions are a subcase of Focus-fronting
e Morphological adjustment of the matrix verb
f The gap at the extraction site must remain non-overt
g Fronted WH-constructions are allowed by bridge verbs
h The meaning of Fronted WH-constructions

(a) THE COMPLEMENTIZER HOGY ("that") is obligatory in Hungarian.

In (5,a,b,c - 6a,b,c) the embedded subject, direct object, and an inherently selected argument, (the Hungarian verb TALÁLKOZ ("meet") selects an instrumental argument), are fronted from the subordinate sentence respectively. Notice that in contrast to fronted WH-constructions in English, the complementizer HOGY ("that") must be obligatorily present in order to avoid ungrammaticality. In English the complementizer THAT must be dropped in case of subject-raising, whereas the complementizer is optional in the case of object raising. This phenomenon has been called the THAT-TRACE EFFECT in Chomsky and Lasnik (1977). Thus, if English possesses a That-Trace Effect, Hungarian has an ANTI-THAT-TRACE EFFECT.

(b) THE ANTICIPATORY PRONOUN of the embedded clause may not be spelled out.

Hungarian embedded clauses are normally introduced by a Case-marked variant of the anticipatory pronoun AZ ("that"). In some instances the pronoun and the clause may form a discontinuous constituent. For example, in case the embedded sentence is focussed the clause introduced by HOGY has to be displaced (cf. Kenesei (1984), (1985) for details).

9 azt tudtam hogy el fogsz
that-ACC knew-1sg-def that away will-2sg
jönni
come-infinitive
'I knew that you would come.'

The anticipatory pronoun may not be spelled out in case an embedded constituent has been fronted in a WH-sentence:

10 * KIT gondolsz AZT hogy János látott

(c) Case change of the EMBEDDED NOMINATIVE SUBJECT.

Hungarian subjects are nominatively marked. Note that in (5a,6a) the fronted embedded subject undergoes a Case change. It ends up accusatively marked. Note, further, from (5-6) that non-nominative fronted WH-phrases retain their Case during this fronting process.

(d) Fronted WH-constructions are a subcase of FOCUS-FRONTING. Fronted WH-constructions in Hungarian are a special case of Focus-fronting. In fact any argument of the embedded sentence may be raised from an embedded clause into the matrix sentence, but only if it lands in the Focus-position (F). Hungarian syntax possesses a fixed Focus-position left-adjacent to the verb. Compare:

11 $_F$MARIT gondolod hogy láttam
Mary-ACC think-2sg-def that saw-1sg-def
'It is Mary who you think that I saw.'

(e) MORPHOLOGICAL ADJUSTMENT of the matrix verb.

Verbs in Hungarian are conjugated in two different ways. They display either definite (glossed as def) or indefinite conjugation (glossed as indef). The definite paradigm is used in case of the accusative argument of the verb is definite; otherwise the indefinite paradigm is used. However, in a number of cases this informal rule is not obeyed. Embedded sentences introduced by AZT ("that-ACC"), for example, trigger the definite conjugation as may be seen from sentence (9). The question words KIT ("who-ACC") and MELYIK NP-ACC ("which NP-ACC") behave differently. Comparing (5a,b) and (6a,b) respectively, accusative WHO-phrases trigger indefinite conjugation but WHICH-phrases trigger definite conjugation. Note that in the case of an indefinite embedded nominative subject or accusative object is fronted the matrix verb is in the indefinite conjugation form. Because of the morphological adjustment in the case of indefinite fronted WH-phrases, there is an agreement correspondence between the matrix verb and the embedded verb in the case of fronted embedded accusative objects. The matrix verb and the embedded verb are both conjugated definitely in case a definite WH-phrase is raised, while they are conjugated indefinitely when an indefinite WH-phrase is raised.

(f) THE GAP at the extraction site must remain non-overt

If it is assumed that the fronted WH-phrase in (5,6) is an argument of the verb of the embedded clause, then the question arises whether its position, i.e. the extraction site, must remain non-overt or whether it may be filled by a personal pronoun? The counterparts of the sentences in (5a,6a) respectively are ungrammatical with an overt pronoun Ő ("he") spelled out at the extraction site. Compare:

12a KIT gondolsz *(hogy) ő látta Jánost
b MELYIK FIÚT gondolod *(hogy) ő látta Jánost

(g) The phenomenon of WH-fronting is allowed by BRIDGE VERBS.

In Hungarian, as is the case in a number of other languages, fronted WH-construction is only allowed in the context of verbs belonging to the semantic classes of PERCEPTION VERBS, VERBS OF SAYING, and VERBS OF KNOWING. Bridge verbs in Hungarian assign accusative Case to their objects.

(h) The fronted is assigned WIDE-SCOPE.

A felicitous answer to the WH-questions in (5,6) would be respectively: PÉTERT ("Peter-ACC"); PÉTERT ("Peter-ACC"); PÉTERREL ("Peter-INSTR"). From this we conclude that a WH-phrase in a WH-sentence is assigned scope over the other constituents in the clause.

Summarizing, the fronted WH-construction in Hungarian represents a continuum. It may be rejected completely, it may be accepted unrestrictedly, or it is restricted by an accessibility hierarchy (cf. 7.). With respect to the second and third dialect/idiolect it has the properties in (a)-(h) as discussed in this section. Let us turn now to a discussion of the question put forward in the beginning of this section: how are embedded questions formed in those cases in which fronted WH-construction is unavailable?

3. MIT-strategy

In De Mey and Marácz (1986) it has been reported that the most common strategy to form embedded questions in Hungarian is to employ the so-called MIT-strategy. Compare the structural counterparts of the sentences in (5,6) in the MIT-strategy:

13a MIT gondolsz hogy Jánost KI látta
what-ACC think-2sg that John-ACC who saw-3sg
'Who do you think saw John?'

b MIT gondolsz hogy János KIT látott
what-ACC think-3sg that John who-ACC saw-3sg
'Who do you think that John saw?'

c MIT szeretnél hogy János KIVEL
what-ACC like-COND-2sg that John who-INSTR
beszéljen
speak-3sg-SUBJ
'With whom would you like that John should speak?'

14a MIT gondolsz hogy Jánost
what-ACC think-2sg that John-ACC
MELYIK FIú látta
which boy saw-3sg
'Which boy do you think saw John?'

b MIT gondolsz hogy János MELYIK FIúT
what-ACC think-2sg that John which boy-ACC
látta
saw-3sg?'
'Which boy do you think that John saw?'

c MIT szeretnél hogy János
what-ACC like-COND-2sg that John
MELYIK FIúVAL beszéljen
which boy-INSTR speak-3sg
'With whom would you like that John should speak?'

The reason we discussed the more marked construction of WH--fronting first, is, because of the fact it has, surprisingly, received more attention in the literature; and has caused more discussion. A number of native-speakers tend to interpret the sentences above as consisting of two parts. In those cases the WH-phrase MIT is the object of the matrix verb as in (13-14) asking for the contents of thought or communication. The second part is an indirect question expressing the issue on which an opinion or statement is being asked. In such cases sentence (13a) could be paraphrased as follows: What is your opinion on the following question: What do you think: Who saw John? (cf. De Mey and Marácz (1986)). Furthermore, all native-informants report that in such cases the complementizer must be dropped. Probably in these cases we have another strategy for solving the problem of questioning embedded constituents. I belive, however, that this strategy does not belong to the sentence-grammar as such and hence I will not discuss it further at this place. Let us turn to a discussion of the MIT-strategy. The MIT-strategy has the following properties:

15a The question word remains IN-SITU
b The anticipatory pronoun may not spelled out
c The scope marker MIT is assigned accusative Case
d The complementizer HOGY ("that") is optional
e The MIT-strategy displays locality effects
f The MIT-strategy is allowed by bridge verbs
g The meaning of the MIT-sentences

(a) The question words remain IN-SITU in the Focus-position of its clause.

In the surface position of the fronted WH-sentence variant, i.e. the matrix Focus, now a dummy question word MIT ("what-ACC") appears which represents the WH-in-situ in the matrix sentence.

(b) THE ANTICIPATORY PRONOUN may not be spelled out in the MIT-strategy.

The MIT-strategy has this property in common with fronted WH-constructions (cf. (b) of section 2). Compare:

16 * MIT gondolsz AZT hogy Jánost KI látta

(c) The MIT-phrase bears the CASE which would be assigned in declaratives to the anticipatory pronoun and to a raised nominative WH-phrase in the fronting-strategy (cf. (b) and (c) respectively of section 2).

(d) The complementizer HOGY ("that") is optional in MIT-constructions.

Recall that in the fronted WH-construction variant the complementizer always has to be present (cf. under (a) of section 2). Consider:

17 MIT gondolsz KI látta Jánost
what-ACC think-2sg who saw-3sg John-ACC
'Who do you think saw John?'

The optionality of the complementizer in the MIT-constructions depends on the position of the real question word in the embedded sentence. In the case of the question word is in initial position as in (17) native-speakers tend to drop the complementizer, whereas if the real WH-word is not in initial position as in (13) and (14), the complementizer is spelled out.

(e) The MIT-strategy displays LOCALITY effects.

In the case of multiple embeddings the Focus-positions up to the Focus-position of the matrix sentence must be filled with the dummy WH-phrase MIT ("what-ACC"). Compare:

18 MIT gondolsz hogy Mari *(MIT) mondott
what-ACC think-2sg that Mary what-ACC said-3sg
hogy János KIT látott
that John who-ACC saw-3sg
'Who do you think that Mary said that John saw?'

Observe that the MIT-phrase must be repeated in each clausal domain. This shows that the MIT strategy is subject to a locality effect.

(f) The MIT-strategy is allowed by BRIDGE VERBS.

The MIT-strategy is allowed with the same verbs as the fronted WH-constructions, that is with VERBS OF SAYING, VERBS OF KNOWING, and PERCEPTION VERBS (cf. (g) of section 2).

(g) The WH-IN-SITU in the MIT-strategy takes wide scope.

From a comparison of the English translations in (1,2) and (13,14) respectively, it can be observed that fronted WH-constructions and their equivalents in the MIT-strategy have the same meaning. A felicitous answer to both the WH-questions with the fronted WH-variant and their counterparts in the MIT-strategy would be respectively: PÉTERT ("Peter-ACC"); PÉTERT ("Peter-ACC"); and PÉTERREL ("Peter-INSTR") (cf. (h) section 2). From this I conclude that the MIT-phrase functions as a scope marker in the sense of Baker (1970). It represents so to speak the scope of the embedded real question word in the matrix sentence.

Summarizing, in this section I have discussed a strategy to form embedded questions in Hungarian in which the real question word remains in-situ in the embedded clause. Native-informants report that the use of the MIT-strategy is by far the most common strategy to form embedded questions in Hungarian.

4. Correspondence Effects

Elsewhere (cf. Marácz (1987), to appear), I have presented an analysis of fronted WH-constructions in Hungarian within Chomsky's (1986b) BARRIERS-framework. The intuition behind this analysis is that the bridge verbs have a lexical property to make the domain of the embedded clause transparent for long distance relations. This is indicated by the fact that the anticipatory pronoun may not be spelled out both in case of fronted WH-constructions and in the case of the MIT-strategy. Furthermore, it can be shown that both strategies are subject to island conditions or locality effects such as the Complex Noun Phrase Constraint, the Sentential Subject Condition, and the Adjunct Condition. The locality effect in the case of the MIT-strategy is stressed by the repetition of the MIT-phrase in each clausal domain. These phenomena support the idea that fronting of the WH-phrase in the raising-variant or the percolation of the scope of the real question word in the case of the MIT-strategy proceeds SUCCESSIVE CYCLICLY. It might be hypothesized that these locality effects resolve the complexity of the formation of embedded questions. The existence of two different WH-question strategies in Hungarian and some of their properties discussed above have some consequences for the theory of grammar.

As noted above the fronted WH-constructions and their MIT-counterparts do not differ in meaning, that is, the question words take widest possible scope in both strategies. This parallelism and the other properties that fronted WH-constructions and the MIT-strategy in Hungarian have in common reported above provides support for the Correspondence Hypothesis:

> 19 CORRESPONDENCE HYPOTHESIS: whenever there is a syntactic reflex of the assignment of (wide) scope, the dependency involved and fronted WH- -constructions obey to a large extent the same gov- erning and bounding conditions.

As indicated briefly above, governing conditions involved are the selection and L-marking (cf. Chomsky (1986b) for this concept) of a subordinate clause by a bridge verb, while the principle of Bounding Theory involved is the Subjacency Condition (cf. Chomsky (1986b) for details), to be more precise 0-subjacency. Note that the Correspondence Hypothesis is not a rule of the grammar but has only descriptive status. In fact, it states that the grammar of WH-gaps and the grammar of scope are constrained by the same syntactic principles. For this reason the Correspondence Hypothesis represents the null-hypothesis. However, the unification between the grammar of WH-gaps and WH-in-situ has surprisingly not been a major tenet of research in the Government and Binding framework in recent years (cf. Aoun et al. (1980), Huang (1982), Hornstein (1984), Lasnik and Saito (1984), Aoun (1986), Chomsky (1986), among others). Rather, on the basis of observations made in Huang (1982), who argues that WH-in-situ in Chinese does not obey locality constraints, it has been concluded that wide scope assignment is not restricted by Bounding Theory and thus the Subjacency Condition is not operative at the level of Logical Form (LF). This is the opposite of the result we have reached on the basis of Correspondence effects in Hungarian. In Huang-Chomsky's terminology wide scope assignment in Hungarian is restricted by the Subjacency Condition and consequently Subjacency holds at LF. From the point of view of language acquisition, it would be unmotivated to escape this contradiction by setting parameters at LF with respect to Bounding Theory (cf. Chomsky (1986a), p 220). Furthermore, in several other studies the Correspondence Hypothesis has received empirical support from other languages than Hungarian, such as Italian and German (cf. Longobardi (1986) and Van Riemsdijk (1983); Bayer (1984) respectively). Recall that the scope of the WH-word in the fronted WH-construction in Hungarian corresponds directly to its position in syntax, whereas the scope of the WH-in-situ in the MIT-strategy is represented at another position than where the WH-phrase is physically realized. However, the real WH-phrases have wide scope in both strategies. Therefore, it is legitimate to put forward the following question: why are there two possible strategies which result in the same scope reading? In the literature several answers to this question have been suggested. For example, according to Williams (1986), natural languages have the following two syntactic quantificational schemes at their disposal, namely an adjunction and an in-situ scheme respectively:

20a $[Q\ N][\ldots t_i \ldots]_i$
   b $[\ldots [Q\ N']_i^i \ldots]_i$

Fronted WH-constructions are an instance of pattern (20a), interpreted cyclically, in which the WH-phrase is actually at the position of operator :i, whereas the MIT-strategy patterns as in (20b) in which the WH-in-situ is bound by the operator :i. The MIT-phrase is in fact an overt representation of :i. It must be admitted, that it is not quite clear at the present state of research why natural languages should have the quantificational schemes in (20) and only those. Whatever the exact reason for the parallel existence of fronted WH-construction and WH-in-situ is, it is obvious that the Correspondence Hypothesis represents the null-hypothesis and that Correspondence Effects are the unmarked case. What seems to be at stake then, is, that the transfer of scope features is constrained by the same principles as the Case and thematic features. Hence, this leads to the following generalization defended in Koster (1987): the transfer of grammatical features (Case, thematic roles, scope, etc.) is restricted by the same syntactic principles.

I derive the following consequences from the Correspondence Hypothesis. First, the WH-in-situ facts of Chinese-Japanese should be reconsidered. The null-hypothesis dictates that it might be worth to investigate whether these languages display locality effects in such cases as well. Studies along these lines have been carried out (cf. Pesetsky (1984) and Koster (1987)). If their conclusions are correct, then, languages such as Chinese with wide scope WH-in-situ but lacking the phenomenon of overt scope markers will provide direct insight into the opacity properties of Universal Grammar, because there is no overt evidence for the child available to "construct" locality effects. Second, if wide scope assignment is restricted by the Subjacency condition and Subjacency is a syntactic constraint, as is generally assumed, then a separate level for the representation of scope, i.e. LF, is unnecessary, that is, LF is identified with syntax (see Williams (1986) for a similar conclusion). Note that the rule of Quantifier Raising in this framework is replaced by the adjunction of the abstract operator :i to the sentence at the level of syntactic representation.

5. Concluding Remarks

In this paper I have discussed strategies of embedded question formation. Native-speakers consulted report that those strategies are subject to dialectical/idiolectical variation. Speakers of Hungarian who reject fronted WH-constructions for forming embedded questions can only rely on the MIT-strategy. Following Chomsky (1986a, p. 75), I may suppose that such dialectical variation is associated with a parameter: MOVE WH. If this parameter is real, it may be hypothesized that it is associated with various other phenomena in Hungarian such as the licensing of parasitic gaps, the setting of the indefinite/definite conjugational paradigm in multiple fronted WH-constructions, etc. I will leave the assessment of the inductive strength of this parameter as a task for further research. The presence of these WH-strategies in the grammar of Hungarian has consequences for the theory of grammar. The existence of both strategies yield empirical support for the Correspondence Hypothesis which excludes the existence of an independent level for the representation of scope, i.e. LF.

NOTES

* I am extremely indebted to István Bodnár, Piroska Csúri, Zsuzsa Gécseg, András Jánossy, László Kálmán, Márta Maleczki, László Pólos, Zoltán Szabó, and Tibor Szécsényi for dicussion on the data presented in this paper. I am extremely indebted to Sjaak de Mey for helping me to evaluate the data collected and

for detailed and careful comments on an earlier version. I want to thank Eric Hoekstra for a critical survey of the text. I alone am responsable for the conclusions and claims made above.

1 In this paper, the following abbreviations will be used: ACC-accusative Case; INSTR-instrumental Case; def-definite conjugation; indef-indefinite conjugation; COND-conditionalis mood; SUBJ-subjunctive mood.

REFERENCES

Aoun, J. (1986). Generalized Binding. Dordrecht: Foris.

Aoun, J., N. Hornstein & D. Sportiche. (1980). 'Some Aspects of Wide Scope Quantification'. Journal of Linguistic Research 1, 69-95.

Baker, C. L. (1970). 'Notes on the Description of English Questions'. Foundations of Language 6, 197-219.

Bayer, J. (1984). 'Comp in Bavarian Syntax'. The Linguistic Review 3, 209-274.

Chomsky, N. (1977). 'On WH-Movement'. In: P. W. Culicover, T. Wasow & A. Akmajian (eds.), Formal Syntax. New York: Academic Press.

Chomsky, N. (1982). Some Concepts and Consequences of the Theory of Government and Binding. Cambridge: MIT Press.

Chomsky, N. (1986a). Knowledge of Language. New York: Praeger.

Chomsky, N. (1986b). Barriers. Cambridge: MIT Press.

Chomsky, N. & H. Lasnik. (1977). 'Filters and Control'. Linguistic Inquiry 8, 425-504.

Comrie, B. (1981). Language Universals and Linguistic Typology. Oxford: Blackwell.

Groot, de C. (1981). 'Sentence-Intertwining in Hungarian'. A. M. Boldestein et al. (eds.), Predication and Expression in Functional Grammar. New York: Academic Press.

Groot, de C. (1981). 'On Theme in Functional Grammar'. In: T. Hoekstra, H. v. d. Hulst & M. Moortgat (eds.), Perspectives on Functional Grammar. Dordrecht:Foris.

Hornstein, N. (1984). Logic as Grammar. Cambridge: MIT Press.

Horvath, J. (1981). Aspects of Hungarian Syntax and the Theory of Grammar. PhD. dissertation UCLA.

Horvath, J. (1986). FOCUS in the Theory of Grammar and the Syntax of Hungarian. Dordrecht: Foris.

Huang, C. T. (1982). Logical Relations in Chinese and the Theory of Grammar. PhD. dissertation MIT.

Kenesei, I. (1984). 'Word Order in Hungarian Complex Sentences'. Linguistic Inquiry 15, 328-342.

Kenesei, I. (1985). 'Subordinate Clauses: Types and Structures'. In: I. Kenesei (ed.), Approaches to Hungarian I. Szeged: University Press.

É. Kiss, K. (1981). 'Structural Relations in Hungarian, a "Free" Word Order Language. Linguistic Inquiry 12, 185-213.

É. Kiss, K. (1985). 'Parasitic Chains'. The Linguistic Review 5, 41-74.

É. Kiss, K. (1987). Configurationality in Hungarian. Dordrecht: Reidel.

Komlósy, A. (1986). 'Focussing on Focus in Hungarian'. In: W. Abraham & S. de Mey (eds.), Topic, Focus, and Configurationality. Amsterdam: Benjamins.

Koster, J. (1987). Domains and Dynasties: The Radical Autonomy of Syntax. Dordrecht: Foris.

Lasnik, H. and M. Saito (1984). 'On the Nature of Proper Government'. Linguistic Inquiry 5, 235-289.

Longobardi, G. (1986). In Defense of the "Correspondence Hypothesis": Islands Effects and Parasitic Gaps in Logical Forms. Ms. Scuola Normale Superiore.

Marácz, L. K. (1987). 'Locality and Correspondence Effects in Hungarian. Groningen Papers in Theoretical and Applied Linguistics, nr. 7. University of Groningen.

Marácz, L. K. (to appear). 'Locality and Correspondence Effects in Hungarian. In: G. Cinque (ed.), Annali di Ca Foscari: Proceedings of the 1987 GLOW conference. Venice.

Mey, de S. & L. K. Marácz (1986). On Question Sentences in Hungarian. In: W. Abraham & S. de Mey (eds.), Topic Focus, and Configurationality. Benjamins: Amsterdam. 253-276.

Pesetsky, D. (1984). WH-in-Situ: Movement and Unselective Binding. Ms. U.Mass/Amherst.

Riemsdijk, Van H. (1983). 'Correspondence Effects and the Empty Category Principle. In Y. Otsu et al. (eds.), Studies in Generative Grammar and Language Acquisition. Tokyo.

Ross, J.R. (1967). Constraints on Variables in Syntax. MIT dissertation.

Szalamin, E. (1978). 'Questions Concerning the Grammar of Modern Spoken Hungarian (in Hungarian). Magyar Nyelv 74, 293-303.

Szamosi, M. (1976). 'On a Surface Structure Constraint in Hungarian'. In: J. McCawley (ed.), Syntax and Semantics. vol.7. 409-425.

Van der Auwera, J. (1983). 'On the Relativization of Embedded Subjects'. Studies in Language, 7. 277-281.

Williams, E. (1986). 'A Reassignment of the Functions of LF'. Linguistic Inquiry 17, 265-299.

Zolnay, Gy. (1926). 'Mondatszövődés' (Sentence-Intertwining). Értekezések a nyelv és széptudományok köréből, XXXIV, 8. Budapest.

INSTITUTE OF GENERAL LINGUISTICS

UNIVERSITY OF GRONINGEN, THE NETHERLANDS

Pesetsky, D. (1987). [illegible] Movement and Unselective Binding, ms. U.Mass/Amherst.

Riemsdijk, van H. (1988). 'Crossover Effects and the Empty Category Principle'. In T. Otsu et al. (eds.), [illegible]

Ross, J.R. (1967). Constraints on Variables in Syntax, PhD dissertation, MIT.

[illegible] Modern [illegible] 73, 299-308.

[illegible] and Semantics, vol.7, 409-425.

Van der [illegible] Subjects', Studies in Language 7, [illegible]

Williams, E. (1986). 'A Reassignment of the Functions of LF', Linguistic Inquiry 17, 265-299.

Zolnay, [illegible] 'Mondatátszövődés', Értekezések a [illegible] XXIV, 9. Budapest.

INSTITUTE OF [illegible]

UNIVERSITY OF [illegible]

ZBIGNIEW KAŃSKI

LOGICAL SYMMETRY AND NATURAL LANGUAGE RECIPROCALS[1]

1. Introduction

In this paper I shall address some issues concerning the applicability of the notion of symmetry as defined in mathematical logic to those natural language expressions that may most naturally invite such application, namely so called reciprocal expressions. I shall accept the standard definition of symmetry as a property of relations (where a relation R is a set of ordered pairs):

(A)(i) A relation R on a set S is symmetric if, for any two elements x and y of s if xRy then yRx

(ii) Asymmetric: For all x and y in S, if xRy then not yRx

(iii) Non-symmetric: neither symmetric nor asymmetric (cf. Partee 1978:10)

It is more difficult to avail oneself of a well-established parallel linguistic definition of reciprocity. Thus one can start with a morphological criterion and study a set of structures with one or more designated morphemes, call them "reciprocal elements", such as each other/one another in English[2] and reserve the term "reciprocity" for well-formed expressions with those designated elements. By this criterion, (1)-(6) would all be examples of reciprocity in English:

(1) They love each other.

(2) These monks lash each other/one another.

(3) The guests followed each other/one another.

(4) The guests followed each other/one another into the room.

(5) We had beds on top of each other/one another (in a sleeper).

(6) Pike eat each other/one another.

On the other hand, English wffs such as (7)-(8) and similar would not fall under thus defined reciprocals:

(7) They kissed.

(8) They collided.

What this approach would miss is that, intuitively, (1) and (7) are much more related in terms of some common sense notion of reciprocity and symmetry than are (1) and (4)-(6). In fact, (7) is more strongly symmetric or reciprocal in terms of the event it describes than any of (1)-(6), while (4)-(6) cannot in any non-ad hoc way be reconciled with a common sense notion of reciprocity, and even less so with logical symmetry. Thus (4) can be truthfully asserted only if there is not a single pair of guests x, y such that if x followed y then y followed x. Similarly, in (5).

On the other hand, if one starts one's linguistic description with some ready semantic definition of reciprocity, say based on symmetry defined on relations as in (A), one would ultimately have to consider sentences such as (4)-(6) either as absurd or else as completely unrelated to (1), except by a homophonous/polysemous expression each other/one another. I will try to demonstrate in this paper that examples (1) to (8) represent a semantically discrete continuum, with examples of type (4)-(5) at its one extreme, and examples of type (8) at its other extreme. The logical discreteness warrants possible distinct morpho-syntactic expressions for particular distinct points on the semantic scale in one language, witness (7) and (8) with no overt "reciprocal element" in English; the fact that we deal with a continuum justifies no "surface" morphosyntactic differences between expressions of its discrete points, as is the case with English (1) on the one hand, and (5)-(6) on the other. In Polish, all semantic types of expressions covered by our English examples (1)-(8) may be expressed by means of one designated element, namely the reflexive pronoun, which, interestingly enough, expands the continuum to cover both reciprocals and reflexives.[3]

In the remainder of this discussion I shall retain the mixed informal term "reciprocal expressions" as a shorthand to cover all types exemplified in (1)-(8), i.e. including those types where the occurrence of a designated "reciprocal element" does not induce common sense reciprocity or anything close to logically defined symmetry, and those types where no designated reciprocal element occurs but whose interpretation is related to symmetry or common sense reciprocity. However, in a more explicit and uniform formulation, the types exemplified above will be claimed to be related in terms of reduction or neutralization of logical ordering of terms with respect to a relation R. The scalarity alluded to above will accordingly be presented in terms of different logically definable degrees of order neutralization.

A few other introductory remarks are in order before we proceed any further. Firstly, I will assume without argument that transitive verbs, including their basic and derived expressions in the sense of type-theoretic categorial grammar, can be represented in terms of two-place relations (in intension) in Montague-style semantics. Our discussion will be limited to basic and non-basic expressions of the category TV, i.e. syntactic functions NP/IV of type $\langle e, \langle e, t\rangle\rangle$. I will also take it for granted that there are language specific ways of representing the logical order (place) of arguments with basic expressions of binary relations. Thus English represents the ordering in terms of syntactic position with respect to a TV while "free-word-order" languages such as Polish typically represent the ordering by means of morphological system of case-marking. Likewise, I will assume after Dowty (1982) that syntactic rules in a categorial format refer only to logical ordering whereas it is a job of syntactic operations to specify language-specific ways of mapping the logical ordering of arguments into surface structure (linear ordering and/or case marking etc.).

Secondly, the logical definition of symmetry (A)(i) qua property of binary relations is not directly applicable to the range of above English examples and their equivalents in other natural languages for several reasons. As it stands, the definition refers to binary relations which are extensions of some basic or non-basic expressions of two-place predicates. Thus, relative to a specified universe of discourse, a given relation is once and for all invested with the property of symmetry (A)(i) or its derivatives (A)(ii) and (A)(iii). If we transfer this definition directly to English, we might for instance want to treat as symmetric relation the extension of the basic TV <u>resemble</u> in the set of all individuals, or the relation <u>be a siter of</u> in the set of all female individuals. This transfer causes numerous problems if it is interesting at all.

There are very few expressions (especially basic expressions of TV) that we could honestly treat in terms of symmetric relations sensu stricto. Those that appear to be most susceptible to such treatment tend to misbehave in some relevant tests. While we can still claim that (10) can be inferred from (9) by postulated symmetry, analogous inference would be problematic from (11) to (12)

(9) Jim $\left\{\begin{matrix}\text{resembles}\\ \text{looks like}\end{matrix}\right\}$ Tom.

(10) Tom resembles Jim.

(11) Tom $\left\{\begin{matrix}\text{resembles}\\ \text{looks like}\end{matrix}\right\}$ a donkey.

(12) A donkey resembles Tom.

Moreover a reciprocal construction (13) is odd:

(13) Tom and a donkey resemble each other/one another.

Turning back to the heterogenous set of reciprocals in (1)-(8), it is easily seen that those examples among them that do allow interpretation in terms of common sense reciprocity or symmetry do not necessarily contain expressions of relations invested with the property of symmetry in the sense discussed above. Thus <u>love</u> or <u>lash</u> cannot be sensibly postulated to denote symmetric relations in the domain of, say, all humans. Clearly, definition (A)(iii) of <u>non-symmetric relations</u> applies to these much more comfortably. In (1), for instance, by means of a special designated element, the non-symmetric <u>love</u> relation is made to pick up just that subset of pairs characterized by it for which both orders are true. Therefore, symmetry in this case cannot be taken to be an "independent" property of the relational expression involved, but derived by a morphosyntactic rule with its semantic correlate. We will be further concerned exclusively with this "contingent", derived sense of symmetry and reciprocity whereby it is a property not so much of relational expressions in them as a property of one-place predicates construed from those (usually non-symmetric and <u>less</u> usually <u>asymmetric</u>) relational expressions by specified rules.

Finally, even the narrow sense of symmetry as defined above will be shown to have only a limited application to the set of reciprocal constructions in linguistics. In the next section it will be demonstrated that some significant attempts to interpret natural language reciprocals in terms of truth conditions for thus modified logical symmetry must fail in the face of examples such as (4)-(5) on the one hand, and, more surprisingly, examples of type (8), on the other, which represent the two extremes of the scale to be proposed. The former elude analysis in terms of symmetry, however adjusted, for the simple reason that the relational expressions involved in them can best be assumed to denote <u>asymmetric</u> relations (cf. <u>follow</u>) in the sense defined in (A)(ii), and therefore one cannot even talk about a <u>subset</u> of pairs for which both orderings are true with respect to those relations. The latter, of the <u>collide</u> type preclude a consistent treatment in terms of symmetry in spite of the intuitive feeling of strong symmetry involved. The reason is that symmetry is defined in terms of implication whose antecedent and consequent represent two reverse orderings of variables with respect to R; obviously, both the antecedent and consequent must be well formed formulas independently of each other. This last condition, however, is not met with expressions of the <u>collide</u> type. Thus (14) cannot be split into two (conjoined) sentences (15), for these sentences will not be well formed:

(14) A truck and a bus collided.

(15)+ A truck collided a bus and a bus collided a truck.

Hence, collide and similar verbs cannot be treated as basic expressions of two-place relations between individuals in the same sense as love and kiss can, and therefore the property of symmetry is not applicable to the former as it can be to the latter. Verbs of the collide type should best be treated as basic expressions denoting properties of sets (cf. Bennett's (1975) category IV), which makes them closely related to verbs such as gather. Thus one end of our continuum shades off into apparently a different area of syntactic and semantic study.

2. Weak Reciprocity Is too Strong

Langendoen's 1978 paper The Logic of Reciprocity is the most explicit and consistent attempt in the literature on the subject known to me, which provides a uniform all-inclusive logical interpretation of a set of reciprocal sentences in English. He is concerned only with what he calls Elementary Reciprocal Sentences (ERSs) which are sentences of the form ARr in which A denotes a set of cardinality ≥ 2, R is a relation on AxA, and r is a reciprocal element, e.g. each other and one another. (Langendoen 1978:177)

After considering several interpretive proposals in the literature, notably Fiengo and Lasnik (1973), and a broad range of English data, Langendoen concludes that proposals are all too strong in the sense that they logically preclude several attested interpretive options available for English ERSs. He then puts forward his logical translation termed Weak Reciprocity (WR) which has the advantage of being inclusive enough to cover the hitherto precluded interpretations as well as all those accounted for in the earlier literature:

B: $(\bigwedge x \in A)(\bigvee y,z \in A)(x \neq y \wedge x \neq z \wedge xRy \wedge zRx)$

(Langendoen 1978:179)

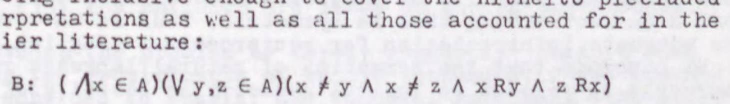
Let us take (1) repeated here for the reader's convenience and assume that they in this ERS denotes a set A of cardinality 3.

(1) They love each other/one another.

$A = \{a,b,c\}$

Then WR allows (1) to be truthfully asserted for four configurations (and others), where arrows symbolize the ordering of arguments with respect to R=love'.

(16) (i) (ii) (iii) (iv)



(16)(i) is a situation whereby for each individual x in A y=z in the formula for WR, and accordingly each individual in A loves another individual in A and is loved by that same individual, and moreover this reciprocal relation holds for any possible pair in A. Naturally, this situation or interpretation is closest to logically defined symmetry (in the narrower sense). In (16) (ii), each individual in A loves another individual in A and is loved by that same individual, but not all possible pairs of distinct elements are thus related (a neither loves nor is loved by c). In (16)(iii) y=z only for some individuals, namely a and b, but not for c (c loves a but a does not love c; c is loved by b but does not love b). (16)(iv) represents the weakest configuration allowed by WR: for no elements in A, y=z, i.e. no individual in A loves and is loved by the same individual, although each loves and is loved by someone. This allows for asymmetric relations to be involved in WR! For the minimum logical requirement of WR for ERSs is that for each element x of A there be a distinct element y of A such that xRy is true <u>and</u> there be an element z distinct from x in A such that zRx is true, where y need not but may be identical to z.

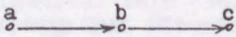
Although Langendoen does not explicitly mention it, WR is tantamount to the simple requirement of identity of Domain and Range of R in ERSs, i.e. the identity of both projections characterized by R in ERSs. How does this condition or the interpretation of ERSs relate to symmetry? If a relation R is symmetric it will always characterize identical projections. However, the converse implication does not hold: identity of projections does not entail that R characterizing them is symmetric. Thus if WR were the adequate interpretation for reciprocals in English, we might conclude that the semantics of natural language reciprocals is related to logical symmetry as identity of projections of relations is to logical symmetry. Analogously, WR for ERSs is weaker than the common sense notion of reciprocity: thus although one can truthfully assert a formally reciprocal sentence (1) about a situation shown in (16)(iv), one would not readily admit that if Tom loves Mary, Mary loves Jim and Jim loves Tom, the three are fortunate enough to experience reciprocal love relations.

WR in Langendoen's formulation has an important advantage in that it allows for asymmetric Rs to be constituents of ERSs, which we showed at the outset to be desirable considering the natural language data. Thus, sentence (3), repeated here, is covered by WR, even if we justifiably consider the relation <u>follow</u> in the set of individuals (or events) to be asymmetric.

(3) The guests followed each other/one another.

If there are three guests, (3) can be true in the situation shown in (16)(iv), that is, if the guests move in a circle: each guest follows another guest and is followed by another guest but none of them follows and is followed by the same individual. The relation will characterize here three ordered pairs (x,y) such that in no case (y,x) will also be in its extension but both its projections will be identical and will each comprize the three guests.

The problem with WR is that it precludes the possibility of a truthful assertion of sentences such as (4) with the same asymmetric relation but most naturally describing a configuration such as (17) (again assuming cardinality 3):

(17) 

Here, only b follows an individual in A and is followed by another individual; a follows someone but no one follows her, while c is followed but does not follow anyone. WR is not satisfied; the projections are not identical: the Domain (left projection) will be a set {a,b} and the Range (right projection) will be a set {b,c}.

Langendoen (1978) is aware of the problem and cites other examples to show that (4) is by no means an isolated case:

(18) The boxes are nested inside one another.
(Langendoen's (31))

(19) The children are lined up behind one another. (32)

He evades the problem by merely suggesting that in some cases the requirements of WR are suspended. These cases are confined to spatial and temporal relations that order the elements of A in the following ways: from top to bottom, from outside to inside, from front to back, from left to right or from right to left, from earlier to later (1978: 192).

This list and the principle are not exhaustive, and even if they were, they fall more under some intracategorial pragmatically defined felicity conditions. These conditions, however, should not prevent in principle a semantic and syntactic categorial rule for ERSs from covering situations of type (17).

I will demonstrate presently that the spatial and temporal relations listed by Langendoen are by no means the only ones that can enter ERSs interpretable in terms of non-identical projections and that therefore WR is not weak enough. But let us first return to the list and try to see what is intuitively achieved by the ERSs with Rs which order elements in space and/or time in contradiction to the same relation in non-reciprocal sentences. Compare our ERS (5) with a non-reciprocal (20) where the latter may in fact be the precise description of the situation allowed but not directly described by the former.

(5) We had beds on top of each other/one another.

(20) Mary's bed was on top of mine.

If we assume that *we* in (5) is a two element set consisting of the denotations of *Mary* and *mine* in (20), and that *on top of* is an asymmetric relation R then the only possible configuration that both sentences can truthfully assert is (21):

(21) $a \rightarrow b$

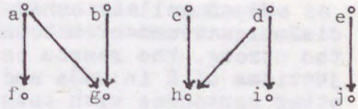
The obvious difference between the two sentences is that while (20) explicitly asserts in *which* order the two individuals are alligned, i.e. which individual replaces which variable in the ordered pair (x,y), sentence (5) only asserts *what kind* of ordering is involved and is non-committal as to which particular order holds true. Hence, in (5) as well as in Langendoen's examples cited above as (18) and (19), the particular ordering of elements with respect to R is not retrievable (perhaps because it is conversationally irrelevant). We will henceforward say that the order of arguments with respect to R is *neutralized* or *reduced*, where neutralization of two possible orders does not entail the truth of both orders, although conjunction of two orders is an instance of order neutralization. Note that the negation of (20) leaves open the possibility that my bed was on top of Mary's, while the negation of (5) precludes both the latter interpretation and the one provided by (20). This inevitably suggests that the logical interpretation of ERSs should be based on alternative rather than conjunctive orderings.

Let us consider some examples to which Langendoen's list does not apply, yet which can be interpreted in terms of non-identical projections and therefore such an interpretation will be outside the scope of WR. Take our (2) and assume that these are, say, ten monks involved, and that *lash* denotes a non-symmetric relation in the set of all humans.

(2) These monks lash one another.

My intuition tells me that among various possible situations which (2) can truthfully describe (including all types shown in (16), there is an extreme case whereby there is not a single monk *x* in the set who both lashes another monk *y* in this set and is lashed by yet another monk *z* (where $y=z$ or $y \neq z$). This situation may be represented as, for instance, (22):

These monks' = {a,b,c,d,e,
(22) f,g,h,i,j}



Thus (2) may be truthfully asserted about the situation whereby the ten monks are divided by R into two subsets of the same cardinality such that the job of five monks is to do the lashing and the passive job of the other five is to suffer the lashing. The projections characterized in this situation not only would not be identical but would not have a single common element! Of course, ERSs are not designed to describe directly this particular situation and the sentence will not be uttered to convey that particular allignment, but rather to escape commitment to any particular allignment. This could be done much more efficiently by saying that five monks lash five monks, and still more precisely by identifying referentialy each five flanking the R. The point is that ERS (2) is non-committal as to who lashes whom in the set and by being non-committal does not exclude in principle the configuration shown in (23). All that needs be verified to establish the contingent truth of (2) is whether each member of the group is involved in the lashing as either a lasher or a lashee (or both, since lash' is postulated to be non-symmetric).

Even the last requirement may be weakened as the cardinality of the set A increases and these monks is taken to refer quasi-generically to a particular order of monks, say Camedules, which is characterized by practicing lashing even if not all its members do actually practice it. The next natural step on this way is provided by generics of the type illustrated in (6) above. Clearly, we do not want to require that each specimen x of the species pike eat another specimen y of the species and should eventually be eaten by another specimen z where $y \neq z$ (or still less probably, $y = z$). All we want to say by (6) is that pike are characterized by cannibalism. But this is just to say that both eaters and their prey are members of the same kind expressed by the generic subject of (6).

It may be concluded from the preceding discussion that the primary and most inclusive function of ERSs is the neutralization of the logical ordering of individual constants with respect to the logical ordering of variables specified by binary relations R. The neutralization thus conceived may involve all three types of relations: symmetric, non-symmetric, and asymmetric, and the range of available interpretations will depend, among other things, on which of these three types a given relation R is postulated a priori to represent. By neutralizing the order of constants with respect to an ordered pair of variables reciprocal constructions in effect turn binary relations into characteristic functions of sets. This will be shown to be a formal correlate of the sense whereby the involvement in lashing (in either of the two "roles" with respect to this relation) characterizes the set of ten monks in our example.

WR as well as other "stronger" logical translations/disambiguations of natural language reciprocals is still too strong. The reason is that it requires identity of projections of R in ERSs and thereby identifies ERSs with any other sentences with both plural subject and object arguments ARB except that in ERSs A=B (cf. Langendoen 1978:§4). I have argued that this requirement should be abandoned under the pressure of numerous, however "marked", counterexamples.

3. Order-Reduction Rule

The following discussion is largely based on the Montague-style format of syntactic and semantic rules, and particularly on Dowty (1982) where rules called relation-reducing rules are proposed to account for Unspecified Object Deletion and passives, cited here as (C) and (D) respectively.

(C) S5: $\langle F_5, \langle TV \rangle, IV \rangle$
$F_5(\alpha) = \alpha$ in English
sem. operation $\lambda x (\bigvee y)[\alpha'(y)(x)]$
$\alpha \in \{$<u>eat</u>, <u>read</u>, <u>write</u>, ...$\}$

(D) S6: $\langle F_6, \langle TV \rangle, IV \rangle$
sem. operation $\lambda y (\bigvee x)[(\alpha'(y))(x)]$
Eng. $F_6(\alpha) = be \frown \alpha'$, where α' is the passive form of α

In both cases reduction affects the number of syntactic arguments, turning a TV into an IV. However, the remaining and the only argument in each case is still identified with one and only one logical position with respect to the order defined by a relation R. Thus the semantic operation in S5 identifies the subject argument of the derived IV with the left argument of R, and the semantic operation in S6 identifies the subject of the derived passive IV with the right argument of R.

Now I suggest that the general rule for ERSs should belong to the family of relation-reducing "detransitivizing" rules such as (C) and (D) in that a TV denoting a two-place relation is also turned into an IV of a special kind denoting in effect a one-place predicate. Contrary to (C) and (D), the subject of the predicate thus derived will not be identified with any of the two particular logical arguments of R for the simple reason that it will consist of an <u>unordered</u> pair (or a set = a union of unordered pairs) comprising both terms which R orders in one of the two possible directions. Thus neither the subject set as a whole nor its particular elements can be identified by virtue of the derived syntax with either of the two variables ordered by R rather than the other. In this case relations R are reduced because order (a defining characteristics of R) is reduced or neutralized by a rule. The following rule, analogous to (C) and (D) has all the features discussed above:

(E) S7: $\langle F_7, \langle TV \rangle, \overline{IV} \rangle$

sem. operation:

$$\lambda S \bigwedge x \left[S(x) \rightarrow \bigvee y \left[S(y) \wedge x \neq y \wedge ((\alpha(y)(x)) \vee (\alpha(x)(y))) \right] \right]$$

$F_7(\alpha) = \alpha$⁀each other or α⁀one another

$\overline{IV}$ in (E) is a syntactic category of one-place verbs whose arguments are always plural and which thereby denote characteristic functions of sets (or properties of sets in intensional semantics) (cf. Bennett 1975). Its basic expressions in English include gather, collide, etc. In (E) $\overline{IV}$ is derived from TV by a relation-reducing rule.

Note that we could append (E) with yet another syntactic operation whereby $F_7(\alpha) = \alpha$ to account for examples such as (7) They kissed. I will show presently that this move is undesirable for it will unduly disregard some intuitive and empirical differences between the latter kind of constructions (call them Ø-reciprocals) and the ones covered by rule (E) (call them r-reciprocals, where r=each other or one another).

First, note that only TV's denoting non-symmetric relations enter Ø-reciprocals, and only a few such TV's can occur in them, e.g. kiss in (7). The second and more difficult problem is whether the Ø-reciprocals with TVs that allow them are identical in meaning with their corresponding r-reciprocals.

Surface-syntactically and morphologically, English Ø-reciprocals with TVs denoting two-place relations are identical to basic expressions of $\overline{IV}$, i.e. expressions denoting properties of sets such as collide, intersect, in that no accusative direct object is present in both cases, and both require plural nominative subjects. The reason to treat collide, intersect, etc. as basic expressions of $\overline{IV}$ was that no conjunctive or disjunctive logical ordering (potentially mapped in English into surface linear subject vs. direct object orderings) is retrievable. In the case of Ø-reciprocals built on basic expressions of TV, such as kiss or embrace, however, whose basic semantics and syntax imposes logical ordering (x kiss y, x embrace y) the retrieval of ordering (conjunctive or disjunctive) is theoretically available, as it is with r-reciprocals built on the same verbs. Thus the options reduce to (i) including Ø-reciprocals in rule (E) and consequently treating them as synonymous with r-reciprocals; (ii) treating Ø-reciprocals as separate lexical items, homophonous with their corresponding TVs, interpreted as basic expressions of $\overline{IV}$ and as such not synonymous with r-reciprocals derived by rule (E) from the homophonous TVs. I will argue that the latter option is more feasible.

Langendoen (1978: 189-91) dismisses (correctly) some earlier transformational attempts to relate Ø-reciprocals to r-reciprocals by means of a meaning-preserving transformation of each other-deletion, for the truth conditions of the two kinds of reciprocals do not seem to be identical. However, I find his solution unconvincing. Among other problems that L.'s analysis faces there is his tacit and unjustified assumption that only basic or non-basic TV's denoting symmetric relations can form Ø-reciprocals (Langendoen's Elementary Covert Reciprocal Sentences). His examples, for instance, include (be)similar but we never learn what category (type) this expression is a member of in the first place; for syntactically it is never a Transitive Verb (mapped into a Z-place relation R); rather, it is a basic expression of a property of sets and as such does not even pose the problem of logical ordering (and therefore symmetry).

It is only when one of two designated prepositions, with, to appears, that a pair of singular arguments may be syntactically ordered with respect to the predicate, but logically we will always deal with symmetric ordering in these cases. (The same is true of basic non-relational, property-denoting expressions of IVs, such as walk; with with a necessarily symmetric relation can be formed: a walks with b ⟷ b walks with a.) What with seems to be doing (at least in one of its meanings) is to syntactically split (order) elements which all satisfy a given one-place predicate and as such are logically unordered with respect to that predicate by definition. Therefore x Verbs with y where Verb=IV or IV does not just happen to be a symmetric relation. It denotes a symmetric relation because the meaning of with ensures that the denotation of its argument NP satisfies the predicate satisfied by the denotation of the subject argument. Thus the secondary syntactic ordering with with does not have the same function as primary ordering imposed by basic and derived Transitive Verb.

We seem to be facing a somewhat paradoxical situation in Logic vs. English: those expressions of English which are most readily interpreted as (logically) symmetric relations, e.g. identical, collide, intersect are not members of the relational syntactical category TV and therefore are not mapped into type $\langle e, \langle e,t \rangle\rangle$ of 2-place relations between individuals, i.e. sets of ordered pairs of individuals. But talking about logical symmetry (and the very definition (A) of symmetry) makes sense only with reference to the latter type, that is with reference to logical order of arguments. A natural conclusion follows: what is defined in matematical logic in terms of a property of two-place relations (symmetry), in English is achieved directly through syntactic categorization IV and corresponding denotation type (properties of sets).

All this suggests that in a strictly logical vocabulary of predicate calculus one cannot seriously refer to English Ø-reciprocals in terms of "relations", symmetric or not. Langendoen is right in dismissing semantic identity between r-reciprocals and Ø-reciprocals where both are possible but at the same time the only semantically interesting case where the two kinds of constructions could be compared are those involving relational expressions (TVs) with primary logical ordering, such as kiss, classified as non-symmetric and which allow for both r-reciprocals and Ø-reciprocals. Not only does L. not consider the latter but his solutions advanced for "symmetric" relations such as similar is, I think, of little linguistic interest. He proposes that Ø-reciprocals be interpreted in terms of strong reciprocity (adapted for subsets which does not bear on the issues presented here):

(F) $(\bigwedge x,y \in A)\ (x \neq y \rightarrow xRy)$

(F) was intended for Ø-reciprocal sentences with symmetric relations, such as They are similar. However, it is not clear to me what English expressions can be translated into R of (F). If similar is so translated, then the language of semantic representation (here, 1st order predicate calculus) would have to licence x Similar y (cf. xRy in (F)) as a well-formed formula. If similar to is so translated, then x similar to y would be a well-formed formula in the syntax of disambiguated language (of semantic representation) and it would be structurally parallel to the object language sentence, but then Ø-reciprocal in the object language would not contain the full expression of the relation in question, for only (23b,c) are grammatical in English:

(23) a. + They are similar to.

b. They are similar to each other.

c. They are similar.

But Langendoen requires that the structure of Ø-reciprocals be AR where A denotes a set of cardinality ⩾ 2 and R is a symmetric relation. Thus if one wants to keep AR as the general syntactic scheme for Ø-reciprocals and at the same time interpret it in terms of (F), then neither similar, identical, equal nor similar to, identical with, equal to can be consistently translated into non-logical constants replacing R in (F).

Langendoen's solution fails in another important area: if Ø-reciprocals (his ECRSs) were interpreted along the lines of strong reciprocity (F) it should be another productive derivative structure vis-a-vis each other/one another that would distinguish structurally between intended weakly symmetric and strongly symmetric interpretation of any non-symmetric relation. But the Ø-object structure is productive with only that class of basic expressions whose primary relational status is dubious, as was shown in the previous paragraph, and which are best analysed as denoting properties of sets with no primary (logical) ordering imposed on members of those sets. Consequently, if similar, identical etc. are taken to be basic expressions of category IV denoting properties of (unordered) sets, the notion of strong symmetry is not only too weak and spurious, but altogether inapplicable.

On the other hand, Ø-reciprocals (AR structures) are not productive with TVs denoting 2-place relations which have logical ordering of arguments given by definition. Only a restricted set of TVs {kiss, embrace, ...} allow the AR structure. This is strange since it is precisely in this area that explicit distinction between derived weak and strong symmetry might be most functional if it were communicationally relevant at all (note that strong reciprocity entails weak reciprocity). Why are there no Ø-reciprocals (AR) interpretable in terms of strong symmetry where R is replaced by kill′, see′, like′ ? Surely, there is nothing in these (non-symmetric) relations that inherently prevents them from having such subsets of their fields that meet the TC for strong symmetry defined by (F); hence I can see no reason why Ø-reciprocal predicates could not be derived from those TVs just as r-reciprocals interpreted in terms of weak-reciprocity can be productively derived from those TVs. TC for strong reciprocity/symmetry requires the truth of conjunctive orderings for any arbitrary pair of distinct elements of the subject set. But if the set has only 2 elements, there is no way of distinguishing between Strong and Weak Symmetry. Hence, if L.′s claims are accepted, the (a) and (b) sentences below should be synonymous while (c) and (d) sentences should have different TC, weak and strong symmetry respectively, which are different for sets of cardinality >2.

(24) a. Mary and Jane are similar to each other.

b. Mary and Jane are similar.

c. Mary, Jane, and Sue are similar to one another.

d. Mary, Jane and Sue are similar.

(25) a. Tom and Mary are kissing each other.

b. Tom and Mary are kissing.

c. Tom, Mary, and Sue are kissing one another.

d. Tom, Mary and Sue are kissing.

(24)(a and b) are synonymous indeed while I have not found sufficiently significant empirical support for different interpretations of (24) (c) and (24)(d), following from L.′s theory. I claim that the synonymy of (24)(a) and (b) follows from the "inherent" symmetry of similar or, as I have suggested above, from the fact that similar is of category IV with no primary logical ordering. (24)(a and c) are results of two syntactic operations: secondary order introduction (similar to NP) and secondary order neutralization (via each other/one another).

Examples in (25) are more criterial if only because the primary relational status of the denotation of TV kiss is undisputable. Hence the synonymy of (25)(a and b) on the one hand, and the weakly and strongly reciprocal interpretation of (25)(c and d) respectively, should support L.'s theory. Yet (25)(a and b) are not synonymous, while (25)(d) is pragmatically odd.

(25)(a) is fine if Tom is Mary's father, for instance, while (25)(b) describing this situation would smack of incest. For kiss with Ø-object acquires a specialized, idiomatic meaning of an amorous, erotic event. (25)(d), unlike (25)(c), describes three people involved in one amorous kiss. The logical difference between (25)(a and b) that would correspond to the use of such sentences consists in the role of logical ordering of variables in the two examples with respect to the verb kiss. In (25)(a), the semantic representation and interpretation would be reducible to the truth conditions for the conjunction (or at least alternative) of both possible orderings kiss' (m,t) kiss' (t,m), and the verification of (25)(a) will accordingly depend on the independent verification of each of the orders. In the case of (25)(a) the two uni-directional events may be simultaneous or alternate; with simple aspect/tense the two events may be separated by a span of time of arbitrary duration.

None of this applies to (25)(b or d). No syntactic and/or semantic reduction to a conjunction (or alternative) of two independently verifiable orders is possible. Hence, not two events but only one event is described and the issue of simultaneity and/or temporal distance does not even arise.

These characteristics are not to be predicted by an interpretation in terms of (F). They will be predicted if we assume that the idiomatic shift of the intransitive kiss as in (25)(b) with respect to the productively detransitivized kiss of (25)(a) is matched by a categorial shift from TV to IV. However, the idiomatic (amorous) kiss will belong to IV not as a result of a relation-reducing rule (E) because no compositional Boolean semantic operation referring to primary order of variables in the transitive kiss is applicable which would correctly represent the difference between (25) (a and b). Kiss is (25)(b) would therefore be a basic expression of IV, denoting a property of sets. It denotes not a bi-directional but a non-directional event. In short, strong symmetry will be still too weak for kiss of (25)(b) as it will be too weak for expressions such as collide, coincide, intersect. Let me demonstrate that this decision receives additional support from the behaviour of each other reciprocals with negation and embedded under some prepositional attitude verbs.

Higginbotham (1980) notes that negative reciprocal sentences such as (26) below are ambiguous as between merely denying their positives on the one hand, and as asserting that neither of a, b loves the other. on the other hand.

(26) They don't love each other.
(they stands for a two-element set {a, b})

If I understand H. correctly, the first interpretation would "merely" deny reciprocity of love between a and b (without denying that, say, a loves b) while the other interpretation does not even presuppose a claim of reciprocity, but is a shortcut to assert that neither a loves b nor b loves a.

If Ø-reciprocals such as (25)(b) are to be interpreted formally in terms of strong reciprocity, as Langendoen insists they should, then their negations should display the same ambiguity. Thus one interpretation of (27) should be that neither Tom is kissing Mary nor Mary is kissing Tom; and one interpretation of (28) should be that it is never the case that either of Tom and Mary kisses the other.

(27) Tom and Mary are not kissing.

(28) Tom and Mary never kiss.

(29) Tom and Mary never kiss each other.

Yet (27) and (28) do not seem to be interpretable in this way. Rather, non-directional, simultaneous involvement in the event is negated in both. If we interpret never as negated existential quantification over moments of time, may assert that at no single point of time it is true that either Tom is kissing Mary or Mary is kissing Tom. On the other hand, (28) will be interpreted only in one way: at no single point of time is the couple (two-element set) involved in a non-ordering kiss' event. (29) will be false (in the second of the two interpretations) if, for instance Tom sometimes kisses Mary on the cheek when he returns from work and Mary never responds; (28) will remain true in this situation. Hence (28) and (29) cannot have identical truth conditions. Since for two-element sets Weak and Strong Reciprocity make identical predictions, (28) cannot be interpreted either in terms of Weak or, more significantly, Strong Reciprocity.

The other relevant test is based on another observation in Higginbotham (1980) whereby reciprocals embedded under some propositional attitude verbs whose subject set is coreferential with the Subject of the embedded reciprocal may, again, be ambiguous.

(30) Tom and Mary think they love each other.

One interpretation would be "narrow" in that the object of both Mary's and Tom's thought would be the reciprocity of their love relation. (Tom thought that they (Tom and Mary) loved each other and Mary thought that they (Tom and Mary) loved each other): the other, wide or distributive, interpretation has it that Tom and Mary have two different thoughts: Tom thinks he loves Mary and Mary thinks she loves Tom. Now compare (31) and (32):

(31) $[\text{Tom and Mary}]_i$ think they_i kiss each other too often.

(32) $[\text{Tom and Mary}]_i$ think they_i kiss too often.

I believe only (31) is liable to receive the "two different thoughts" interpretation. Since, again, in the case of two-element subject sets, strongly and weakly symmetric interpretations are equivalent, ∅-reciprocals cannot be interpreted in terms of Strong Reciprocity, for (31) and (32) should have the same range of interpretations, which they do not. Hence, ∅-reciprocals cannot be interpreted in terms of strong symmetry. The conclusion is rather straightforward: ∅-reciprocals are not "symmetric" in any formal sense (Weak to Strong). The reason is that ∅-reciprocal sentences are not derived syntactically from verbs of category TV denoting two-place relations between individuals. Rather, they are derived from basic verbs of Bennett's category $\overline{IV}$ denoting properties of sets.

4. Degrees of Order-Reduction

We have shown that the weakest possible formal definition of reciprocity (WR) related to logical symmetry is still too strong for one class of English reciprocals; the strongest possible formal definition of reciprocity (SR) is still too weak for another class of English reciprocals. The former class of English expressions has a form of "reciprocal expressions" marked by a designated lexical/morphological reciprocal element, but the only interpretation that can be accorded to them contradicts both the common sense notion of reciprocity and the formal-semantic definition of symmetry. The latter class of English expressions lacks a designated reciprocal element (for ∅ cannot be considered to be such an element), though here, paradoxically, the common sense reciprocity may reach its peak; indeed, as we argued in the preceding section, the reciprocity in this class is too good to be true in terms of truth conditions for (Strong) Symmetry, for conjunction of two orderings of arguments with respect to a relation entailed by Strong Reciprocity is not directly representable for the class of ∅-reciprocals.

In this paper I have been arguing in effect that the requirement of conjunctive ordering, hawever stretched, should be abandoned, with all its implications, in favour of a weaker and pragmatically more motivated common denominator of and semantic effect achieved by reciprocals, i.e. order reduction or neutralization. This is to say that the overall job of a reciprocal rule (or rules) is to undo via syntactic and/or lexico-morphological devices the logical ordering imposed by expressions whose defining characteristic is precisely to impose such ordering, i.e. TVs denoting sets of <u>ordered</u> pairs (or intensions thereof).

Since logical order of arguments is expressed via language specific morphosyntactic means (word order, case), neutralization of order needs also be "visible". This is achieved in English through the designated "reciprocal" element or zero-object, and the plurality of the argument. Syntactically speaking, an order-neutralized TV is that kind of IV that takes a plural argument; the latter plural argument really comprises singular or plural argument NPs that would flank the TV prior to the neutralization rule(s). Hence a corresponding type shift from characteristic functions of ordered pairs to characteristic functions of sets.

Now the requirement that the set-denoting plural argument <u>comprise</u> the potentially ordered individual or set arguments is clearly distinct from the requirement that the argument set be coextensional with both projections of a relation R denoted by the TV. We showed that the latter constraint, underlying Langendoen's WR is too strong. The sense whereby the argument set comprises elements potentially ordered by R was expressed more formally in terms of alternative ordering (AO). This treatment makes reciprocals related directly not so much to symmetry qua property of relations (however one adapts and stretches the notion of symmetry) but rather to the relation theoretic concept of the field, i.e. union of domain and range of a relation R. Thus our relation reducing rule (E) produces a predicate $\overline{IV}$ that denotes a characteristic function of sets S such that S is really the field of relation R in S.

The range of data discussed and the formal notion of relation reduction that accounts for the data suggest that order reduction may be scalar and that the position on the scale depends on the verb involved (whether it is postulated as denoting an asymmetric or non-symmetric relation) and various pragmatic factors (e.g. licencing weaker commitment as to membership in the two projections of some R as the cardinality of the field of R increases).

(i) Exclusive Alternative Ordering (EAO): follow each other

(ii) Inclusive Alternative Ordering (IAO): lash each other

(iii) Conjunctive Ordering (CO): kiss each other

(iv) Order Cancellation (OC)/Lexical Shift:

kiss-∅, embrace-∅

(v) No primary order - basic expressions of $\overline{IV}$: collide

(i)-(iii) are covered by options allowed by our order reduction rule (E). (i)=EAO is the weakest possible order reduction, naturally associated with relations most resistant to such reduction, i.e. asymmetric relations (cf. follow in (4)). But as we showed in section 2, EAO is in principle available also for nonsymmetric relations. The point is that with nonsymmetric relations, the order reducing rule (E) does not in itself specify any particular interpretive options out of (i)-(iii), all other things being equal; the semantic effect of (E) is neutral with respect to (i)-(iii), which seems to have a natural pragmatic counterpart in the irrelevance of membership in particular projections of R. The interpretation in terms of (iii)=CO will be most obvious with subject-sets of cardinality 2 (where non-commitment to any particular ordering if only one were true might be considered uncooperative). Langendoen's WR is the strongest order reduction on the scale covered by rule (E), i.e. it is compatible with CO where CO is understood in the non-connected sense whereby xRy and zRx must be true (where $x \neq y$, $x \neq z$).

(iv)=OC is the next point on the scale plotted from the weakest to the strongest reduction of order. OC is reserved for those TVs whose $\overline{IV}$ homophonous conterparts tend to be invested with extra (idiomatic) meaning and where directionality is obliterated (irretrievable in terms of (i)-(iii)). I have used the term Order Cancellation suggestive of some operation on relations but I can think of no formal expression of such an operation and have decided to treat verbs exemplifying OC as really basic expressions of $\overline{IV}$ (which corresponds to the lexicographic practice of classifying e.g. kiss as a TV and an IV, with possible additional, unpredictable meaning difference). Hence, the relevant examples should best be thought to represent a transitional type of order reduction, between the rule-governed CO and class (v) to which the term order-reduction no longer applies, for it comprises basic expressions of $\overline{IV}$ with no primary order (non-relational expressions).

It is clear that symmetry as defined in mathematical logic applies (significantly modified, as in Langendoen's definition of WR) only to (iii)=CO on our scale, while the latter scale covers possible semantic effects that should be associated with the equivocal linguistic term "reciprocals." Hence, in identifying a morpheme or lexical item in a natural language as "a reciprocal element/word/morpheme" it is neither sufficient to rely on our common sense notion of reciprocity nor to adduce the formal definition of symmetry, for the semantic spectrum covered by "reciprocal" expressions and sentences seems to be wider and correlate with our notion of order reduction on the one hand, and basic expressions of catefory $\overline{IV}$ (denoting properties of sets), on the other.

NOTES

[1] A significant part of research reported in this paper was conducted in spring 1987 when I was a Sidney Holgate Fellow at Grey College, University of Durham.

[2] Following Langendoen (1978), I will ignore the prescriptivist difference between the usage of each other and one another, as indeed most of native speakers of English seem to do.

[3] For a treatment of reflexives in terms of relation reduction, related to the present approach to reciprocals, see Kański (1986).

REFERENCES

Bennett, M. R. (1975) Some Extensions of a Montague Fragment of English, reproduced by Indiana University Linguistics Club.

Dowty, D. (1982) "Grammatical Relations and Montague Grammar," in P. Jacobson and G. K. Pullum, eds., The Nature of Syntactic Representation, D. Reidel, Dordrecht, 79-130.

Fiengo, R. W. and H. Lasnik (1973) "The Logical Structure of Reciprocal Sentences in English," Foundations of Language 9, 447-468.

Higginbotham, J. (1980)"Reciprocal Interpretation," Journal of Linguistic Research 3, 97-117.

Kański, Z. (1986) Arbitrary Reference and Reflexivity, University of Silesia Press, Katowice.

Langendoen, D. T. (1978) "The Logic of Reciprocity," Linguistic Inquiry 2, 177-197.

Partee, B. H. (1978) Fundamentals of Mathematics for Linguistics, Greylock Publishers, Stamford, Conn.

DEPARTMENT OF ENGLISH AND GENERAL LINGUISTICS
UNIVERSITY OF SILESIA, SOSNOWIEC, POLAND

SJAAK DE MEY

TRANSITIVE SENTENCES AND THE PROPERTY OF LOGICALITY

1. Introduction.

In this paper we give an introduction to the semantics of transitive declarative sentences. Due to limitations put on the size of the papers in this volume, however, only one small aspect of it, so-called logicality of dyadic quantifiers, can be dealt with in some detail. For the background knowledge used in this paper we refer to van Benthem (1986, 1987) and De Mey (to appear).

In section 2 we discuss two different analyses of declarative transitive sentences. In section 3 properties of quantifiers, e.g. conservativity and logicality, are discussed. In section 4 we show that the property of logicality should be defined differently than is done traditionally and that a new definition can be phrased in terms of two properties of dyadic quantifiers which are subsequently discussed in section 5. Section 6 discusses a number of linguistically interesting dyadic quantifiers and their properties. Section 7 summarizes the main results.

2. The intransitive analysis and the absorption analysis.

A more or less tacit assumption of formal semantics has been that the structure of declarative transitive sentences can be analyzed in exactly the same way as intransitives, that is as

(1) DET(N,IV)[1]

We will call this the intransitive analysis. In a transitive sentence the phrase consists of a transitive verb, a direct object and/or one or more prepositional objects. So, spelled out in more detail, for transitives, (1) should look like this:

(2) $DET_s(N_s, \{x \mid DET_o(N_o, TV_x)\})$

where TV_x is the image set of x under TV, that is, the set $\{y \mid \langle x,y\rangle \in TV\}$.

This analysis has much to recommend itself as the same analysis of determiners can be applied to subject and object determiners in transitives as to subject determiners in intransitives. Nevertheless, it must be wrong: not all transitive sentences can be analyzed in accordance with the intransitive analysis.

Another way of analyzing transitives was put forward in Higginbotham & May (1981). We call this the absorption analysis. Under an absorption analysis the subject and object determiners are taken together ('absorbed') into a determiner-complex yielding the following (semantic) structure:

(3) $\langle DET_s, DET_0 \ldots, DET_0 \rangle \, (N_s \times N_0 \times \ldots \times N_0, TV)$

We will use DET both for the denotations of determiners and determiner complexes.

Determiner complexes denote relations between sets, just like determiners. In this case, however, the sets are n-place relations between individuals, instead of being just sets of individuals. That is, the denotation of a determiner complex is a binary relation[2] between n-place relations. We will call such binary relations n-adic quantifiers. Henceforward we will restrict attention to two-place transitives, where the object may either be a direct object or a prepositional object. Consequently, the DETs we are interested in are dyadic quantifiers.

The important question, of course, is whether a sentence expresses different truth conditions under the two analyses. If all sentences would express the same truth conditions under both analyses there would not be much point to working out the absorption analysis. However, this is not the case.

Let us call a determiner complex 'separable' iff

(4) $\langle DET_s, DET_0 \rangle \, (N_s \times N_0, TV)$ iff
$(DET_s^s \, (N_s, \{x \mid DET_0^s (N_0, TV_x)\}))$

If and only if a sentence has a separable determiner-complex under an absorption analysis, does it express the same truth-conditions under both analyses. It can be readily checked now that not all types of transitive sentences have separable determiner complexes under an absorption analysis. More in particular, certain types of sentences can only be sensibly analyzed using an intransitive analysis, others have only use for an absorption analysis, whereas still others admit both analyses both express different truth conditions under each of these.[3]

Sentences with an inherently distributive subject determiner such as 'each' or 'no' (sg) can only be analysed using an intransitive analysis. Reciprocal sentences, sentences with a dependent plural subject such as 'Unicycles have wheels' or (Dif,Dif)-sentences such as 'Different students read different books' admit only an absorption analysis. Sentences such as 'Three boys ate four apples' are ambiguous between a distributive reading meaning that each of three boys ate four apples, and a total reading meaning that three boys ate a total of four apples. Compare the Hungarian sentences 'Három fiú négy-négy almát evett' and 'Három fiú négy almát evett'.[4]

As Keenan (to appear) points out, the question of whether a determiner complex is separable or not hinges on the question of whether we can determine the image sets of each of the members of the domain under the transitive verb without having to look at the image sets of other such members. If I have 'Three boys ate four apples' under a distributive reading the important fact is that each of three boys has an image set of cardinality four or more under the verb 'eat'. It does not matter, so to say, whether boy a ate different apples than boy b. This is not so in the case of 'different student read different books', 'Unicycles have wheels' or 'All the boys helped each other'. In the case of 'three boys ate four apples' under a total reading the important point is that a totality of four apples is eaten by three boys, irrespective of the number of apples that each individual boy ate.

Let us realize the importance of the separability-issue. Only sentences that admit both types of analyses exhibit the classical scopal ambiguities and only under an intransitive analysis. 'Each student read some books' cannot be made to mean that some books were read by each student. In the cases of reciprocals, the 'unicycle'-type of sentence and the 'dif--dif'-type of sentence it simply does not make sense to ask which NP has largest scope. In the case of 'three boys ate four apples' there is scopal ambiguity only under a distributive reading.

Also the consequences for the theory of generalized quantifiers are of relevance. The fact that transitive sentences may have unseparable determiner complexes means that we need dyadic quantifiers alongside monadic ones. A whole new area has been added to the theory.

3. Properties of dyadic quantifiers.

DETs have been demonstrated to possess interesting properties. Knowing these properties of course results in a better insight into the semantics of the sentences of which the DETs are the main operators.

A property that is assumed to be common to determiner-denotations of all types is CONSERVATIVITY. Furthermore, a property that a large subset of DETs of all types have is LOGICALITY. We call a relation Q between subsets of E logical iff Q is closed under automorphisms of E. That is, if π is an automorphism (1-1 mapping) of E onto E then $\langle A,B\rangle \in Q$ iff $\langle \pi(A), \pi(B)\rangle \in Q$. The intuitive sense of this is that logical quantifiers cannot discriminate between sets of equal cardinality.

Furthermore, monadic DETs have characteristic properties which can be described by making use of notions such as 'transitive', 'symmetric', 'connected' 'reflexive', 'non--symmetric', 'antisymmetric' and asymmetric'. Compare for an extensive discussion Zwarts (1984).

With dyadic quantifiers, however, such notions are not very instructive. Dyadic ALL is the partial order just as monadic ALL and all members of TRANS are transitive by definition. So, we have to look for more illuminating properties. As it turns out, a very useful notion for our purposes is 'being closed under a particular form of automorphisms'. There are two concepts residing under this more general notion: 'being closed under α-automorphisms' where α ranges over a hierarchy of automorphisms of various degrees of strength; and 'being closed under conversion'. We discuss both notions in section 5 in more detail. Both notions appear to be interesting properties in relation to scope ambiguities; scope reversal, e.g., amounts to conversion of the relation.

A quantifier Q is closed under conversion of its members or, simply, 'closed under conversion' iff $R \in Q$ iff $conv(R) \in Q$. Obviously, conversion is also an automorphism. SYM and RECIPR are closed under conversion, since $R = conv(R)$ iff R is symmetric. Also REFL is closed under conversion.

It should be borne in mind, however, that the most interesting thing to know about a set of relations is not that it is closed under a particular type of operation but to know what the maximal set of operations is under which it is closed. Although, e.g., SYM is closed under conversion it may be (and certainly is) also closed under more embracing sets of operations of which conversion is only one member.

4. Logicality.

Monadic DETs which are closed under automorphisms of E are called 'logical'. An automorphism of a set A is a 1-1 mapping of A onto A. Logical DETs are sensitive to certain logical properties of their arguments only, such as quantifiers. Generalizing this definition we have to say that dyadic DETs are logical iff they are closed under automorphisms of $E \times E$. SYM, then, is not logical as SYM is not closed under arbitrary automorphisms of $E \times E$. The same holds of REFL, TRANS and RECIPR.

It would appear now that this definition of logicality is too coarse. What we try to capture when we call a quantifier logical is the fact that it is sensitive to logical properties of its members only. In other words, a quantifier cannot distinguish between two different relations when both have the property required.

Now, far more quantifiers are logical in this sense than in the sense of our definition. SYM, TRANS and REFL certainly seem to be logical in an intuitive sense. Also quantifiers such as <ALL,SOME> as in 'All students read some books' are logical in a very elementary sense, but <ALL, SOME> is not logical in the sense of our definition either.

An alternative and more interesting way of putting the same point is by pointing to the structure that is inherent in a relation. Quantifiers closed under operations that preserve the structure of the relation will be have to be called logical then.

Let us therefore try and device a new concept of logicality. We do this by discussing the property of dyadic quantifiers of being closed under certain weaker types of automorphisms. We will have to examine under exactly which types of automorphisms classes of quantifiers are closed.

5. Two properties of dyadic quantifiers.

5.1 The HM-hierarchy.[5]

Higginbotham & May (1981) discuss the following types of automorphisms of E x E. We will cite them under names different from the ones HM used themselves.

A - pair-automorphisms
m is a pair-automorphism iff m is an automorphism of E x E

B1 - L-automorphism
m is a L-automorphism iff there are automorphisms p and q_a of E such that $m(\langle a,b\rangle) = \langle p(a), q_a(b)\rangle$

B2 - R-automorphisms
m is an R-automorphism iff there are automorphisms q_b, p of E such that $m(\langle a,b\rangle) = \langle q_b(a), p(b)\rangle$

C - LR-automorpism
m is a LR-automorphism iff there are automorphisms p and q of E such that $m(\langle a,b\rangle) = \langle p(a), q(b)\rangle$

D - I-automorphisms
m is an I-automorphism iff there is an automorphism p of E such that $m(\langle a,b\rangle) = \langle p(a), p(b)\rangle$

In B1 and B2 it is assumed that the choice of q_a and q_b depends on **a** and **b** respectively. We come back to them in a moment. In what follows we will call automorphism belonging to one of the lower classes (that is, either B1, B2, C or D) 'HM-automorphisms'.

This is a hierarchy in the sense that automorphism of a lower category are, at the same time, members of the higher categories. D is lower than C, C is lower than either B1 or B2, whereas both B1 and B2 are lower than A. This means that when a quantifier Q is closed under a lower type of automorphisms Q is also closed under automorphisms of a higher type.

I-automorphisms obviously are LR-automorphisms and LR-automorphisms are both L-automorphisms and R-automorphisms. (As a R-automorphism is, so to say, the mirror-image of an L-automorphism, however we will pay attention to representatives of this type only when there is a special reason to do so). Finally, L-automorphisms and R-automorphisms can be easily shown to be pair-automorphisms.

Relations closed under L-automorphisms have the following interesting property. Assuming that $m(\langle a,b\rangle) = \langle p(a), q_a(b)\rangle$ it holds for each $a \in dom(R)$ that its map p(a) has the same number of successors under m(R) as **a** has under R. L-automorphisms preserve for each $a \in dom(R)$ the cardinality of the image set R_a.

It may be useful to look at some figures. Here are the numbers of the different automorphisms:

		card(E) = 2	card(E) = 3
I-automorphisms:	n!	2	6
LR-automorphisms:	n!xn!	4	36
L-automorphisms:	$(n!)^{n+1}$	8	1296
R-automorphisms:	$(n!)^{n+1}$	8	1296
pair-automorphisms:	(n x n)!	24	362880

The difference between the number of HM-automorphisms and pair-automorphisms grows gigantically under extension of the domain. It is obvious, then, that defining logicality in terms of being closed under HM-automorphisms results in far more restrictive a notion than when we define it in terms of being closed under arbitrary automorphisms.

5.2 Conversion.

There are two different forms of conversion which we should carefully keep apart. On the one hand we have the converse m^{-1} of a pair-automorphisms m, on the other hand we have the converse of a binary relation R. The converse m^{-1} of a pair automorphisms m of type X is again a pair automorphism of type X as $m^{-1}(m(R)) = R$. The converse conv(R) of a relation R can be thought of as arising from applying a pair automorphism conv to R as card(R) = card(conv(R)).[6]

Conv is not a HM-automorphism. Let us go into the reasons why. So far, we have met with two types of automorphisms. Pair-automorphisms do not obey any furher restrictions beyond cardinality. HM-automorphisms are all in one way or other based on automorphisms of E. That is, pairs are mapped to other pairs such that either each member (I- and LR-automorphisms) or at least one member of each pair (L- and R-automorphisms) is mapped to another individual in accordance with a permutation of E.

With conv, however, reflexive pairs, that is pairs of the form ⟨x,x⟩, are treated differently from irreflexive pairs, that is pairs of the form ⟨x,y⟩ with x ≠ y. conv maps irreflexive pairs according to a certain pair-automorphism m such that m(⟨x,y⟩) ≠ (⟨x,y⟩). Reflexive pairs, on the other hand, are invariably mapped onto themselves. So, conv obeys restrictions but these are different for pairs belonging to different subsets and cannot be stated in terms of the HM-hierarchy.

6. Some dyadic quantifiers.

6.1 Two dyadic Squares of oppositions.

So far we used as examples quantifiers such as SYM, TRANS, REFL and RECIPR. Let us look now at some more dyadic quantifiers in order to see under exactly which structure preserving operations they are closed.

We consider dyadic quantifiers which result from applying absorption to the logical DETs of the square of oppositions ALL (A), NOT_ALL (¬A), SOME (E, or ¬A¬), and NO (N, or A¬). These dyadic DETs[7] can be classified in two 'dyadic' squares of oppositions:

```
I:                                      II:

AA/N¬A ----- E¬A/¬AA                    A¬A/NA ----- EA/¬A¬A
  |              |                         |              |
AN/NE  ----- EE/¬AN                     AE/NN  ----- EN/¬AE
```

or:

```
I:                                      II:

AA    ----- ¬AA                         A¬A    ----- ¬A¬A
 |            |                           |            |
AA¬   ----- ¬AA¬                        A¬A¬   ----- ¬A¬A¬
```

Membership of such dyadic DETs is determined by the cardinality of the relation, completely in the case of the type-I DETs, partially (necessarily but not sufficiently) in the case of the type-II DETs, in accordance with the following schema:

```
I:                              II:

n² ..... max: n² - 1            max: n(n-1) ... min: n
.        .                      .               .
.        .                      .               .
0 ...... min: 1                 min: n ........ max: (n + 1) (n - 1)
```

(As to the latter figure, all but one may have V-d all: (n-1) x n, and one may have V-d as many as all but one: (n-1)).

Consequently, the type-I DETs are closed under pair-automorphisms but for the type-II DETs this would be too weak a requirement: type-II quantifiers are closed under L-automorphisms.

Before we discuss type-II quantifiers we notice that the square of type-I quantifiers is identical to the square of monadic quantifiers up to the type of their arguments. AA and A differ only in that AA takes relations as arguments and A sets. The same holds of ¬AA and ¬A, EE and E, AN and N.

As for the type-II quantifiers, let us first define the set FU of functions having E as their domain, that is, $R \in FU$ implies that for each $x \in R$ $|R_0| = 1$. We now have the following definitions:

$R \in AE$ iff there is a S: $S \leq R$ & $S \in FU$

$R \in NA$ iff there is a S: $S \leq (E \times E - R)$ & $S \in FU$

$R \in EA$ iff there is no S: $S \leq (E \times E - R)$ & $S \in FU$

$R \in EN$ iff there is no S: $S \leq R$ & $S \in FU$

The monadic square was built on the notion of empty difference, or subset. That is, the members are the subset relation, its external negation, its internal negation and its dual. The dyadic square I is built on the notion of the Cartesian product. Square II is built up on the notion of function but here we have to add that the members of such DETs are not only members of FU and its (FU's) internal and external negation, but also their supersets. That is, such quantifiers are monotonically increasing.

6.2 Characteristic sequences.

In order to gain a deeper insight into the essence of all this we should look at what automorphisms do to the structure of a set. Certain types of structure may be preserved under permutations, others may not. Particularly the structure present in subset need not be preserved.

Let us first look at the Boolean structure of a set. A set has a fixed number of subsets. Moreover for each n there is a fixed number of subsets of cardinality n. The Boolean structure of a set is preserved under automorphisms, that is after permutation there is still the same number of subsets, and, also, there is, for each n, the same number of subsets of cardinality n. It would appear that monadic quantifiers are mainly characterized by this Boolean structure. This explains why logicality of monadic quantifiers can be aptly defined in terms of cardinality.

A different situation obtains with structures inherent in subsets of the members of 2-place relations R. Let us examine an example and look at what exactly is preserved under L-automorphisms. For that purpose we introduce the notion of a characteristic sequence. Let us think of a relation R as a sequence of numbers $(p_0, p_1, p_2, \ldots, p_n)$ where card(dom(R)) = n and p_i is the nunber of members of dom(R) that have i R-successors. Let us call such a sequence a characteristic sequence (CS). If a set S of relations is closed under L-automorphisms, we can look upon S as a family of equivalence classes of relations, where each such equivalence class is characterized by one and the same characteristic sequence. We may think of such a sequence as representing the image structure of a relation.

Now we can say that L-automorpisms preserve the CS of the members of a quantifier. Consequently, the cardiality of the range need not be preserved. E.g., suppose $E = \{a,b\}$, $R = \{\langle a,b\rangle,\langle b,a\rangle\}$, $\pi(R) = \{\langle a,a\rangle,\langle b,a\rangle\}$. However, when the cardinality of the domain is not preserved, the CS of such a relation is not preserved either, as the numbers of the CS add up to the cardinality of the domain.

Membership of the quantifiers of squares II can now also be expressed in terms of conditions on characteristic sequences in accordance with the following schema:

II:

$$
\begin{array}{lcl}
p_n = 0 & \text{--------} & p_n \geq 1 \\
\vdots & & \vdots \\
p_0 \neq 0 & \text{--------} & p_0 > 0
\end{array}
$$

Finally, consider

(5) Three boys ate four apples

The property to be preserved by pair-automorphisms is that all of the members of <THREE-FOUR> have a characteristic sequence in which $(p_4 + p_5 + \ldots + p_n) \geq 3$. No other similarities between CSs are required. That is, in order to be a member of <THREE-FOUR>, only a part of a relation has to be considered, the part which we will symbolize by "4(R)":

$$4(R) = \{\langle x,y\rangle \mid card(R\ ) \geq 4 \ \&\ y \in R_x\}.$$

<THREE-FOUR> is best considered a member of EE.

6.3 Inherent structure.

We set out to investigate dyadic quantifiers. Our aim was to find out which properties may help us to draw a map of the area occupied by dyadic DETs.

The important point is the inherent structure of a class of relations. Such structures determine the type of operation the class is closed under. Let us list the different types of structures we came across investigating dyadic quantifiers.

a - cardinality.

Type-I quantifiers are closed pair-automorphisms as they preserve nothing but the cardinality of a relation. Cardinality is preserved both under conversion and automorphistic permutation.

b - image structure, that is the way the relation relates the memebers of its domain and subsets of the range.

We should distinguish here between two subtypes:

b1 - the numerical image structure representable by a characteristic sequence which sums up for each different cardinality the number of members of the domain which are related to an image set of this cardinality. Dyadic numerical quantifiers such as <THREE-FOUR> belong here.

b2 - the 'shape' of the relation which is describable by predicates such as 'one-to-one', 'many-to-one', etc. The type – II quantifiers such as AE belong here.

Neither numerical image structure nor shape is preserved under conv, but is preserved under L-automorphisms.

c - the 'architecture' of the members of a relation, e.g. the architecture of symmetric, reflexive, transitive or reciprocal relations. Architecture is characteristically preserved under conversion and also under strong forms of automorphistic permutation such as I-automorphisms.

As to the type of structure preserved under conversion, there are two cases. A quantifier may be insensitive to conversion of its members because besides their cardinality and besides the fact that they are relations of a given type, e.g. they are dyadic, the members of the quantifier have no structure in common. Or the members do have a certain type of structure in common and it is preserved under conversion. The first possibility is realized in the type-I quantifiers, the second in, e.g., RECIPR, SYM, REFL and TRANS.

NOTES

* I want to express my gratitude to Johan van Benthem for insightful discussion of the contents of this paper and for his encouragement.

1. In model theoretic semantics sentences are interpreted in a model which is a pair ⟨E, ‖ ‖⟩ of a set or universe of discourse E and an assignment function '‖ ‖' assigning denotations (extensions, interpretations) to well-formed expressions. Denotations are set theoretic constructions whose definition may make use of E. N and IV are properties, that is, subsets of E. DET is a binary relation between properties, that is, between unary relations between individuals, members

of E. We will call such relations monadic quantifiers.

2. Clark & Keenan (1985/86) point out taking determiner-complexes to be binary relations between binary relations leads to certain difficulties. However, DETs can just as well be taken to be sets (unary relations) of relations. There is no essential difference between dyadic DETs and arbitrary sets of binary relations except that the former are denotations of natural language expressions. In what follows we use the term 'quatifier' for any set of relations. So we have the set SYM of symmetric relations between individuals, the set TRANS of transitive relations between individuals, and the set REFL of reflexive relations between individuals, which are all (dyadic) quantifiers in our sense.

A linguistically interesting class of relations is furthermore RECIPR, the set of reciprocal relations between individuals. For an informal idea of a reciprocal relation we can use the following: if R is reciprocal then it holds that each x who R's some y is R'ed by some z, where $x \neq y$ and $x \neq z$. It follows that every x who is R'ed by some y R's some z himself. More formally, for each $x \in dom(R)$ there are y, z such that $\{\langle x,y\rangle, \langle z,x\rangle\} \leq R$ and $(x \neq y \ \& \ x \neq z)$. Reciprocity is weak-symmetry as each symmetrical relation is reciprocal but not vice versa.

For more details, compare De Mey (to appear).

3. It should be borne in mind that this means that such a sentence does not have a separable determiner-complex under an absorption analysis.

4. Compare for a far more extensive analysis of such facts De Mey (to appear).

5. This hierarchy is derived from Higginbotham & May (1981). It should be added that the authors do not present it as such.

6. A combination of automorphisms is also an automorphism. Also, we should realize that the set {R, m(R)} consisting of a dyadic relation R and its permutation product under some automorphism m is a dyadic quantifier. The same holds of the closure of a set S of relations under a class of automorphisms m. We refer to such a set as |m|(S), the closure of S under a certain type of automorphisms of which m is a characteristic representative.

7. The names chosen for these quantifiers are self-explanatory. E.g., the DET AA is the denotation of the determiner-complex <all,all> as we find in 'All students read all books' (under an absorption analysis, of course). AA and N¬A are the same quantifier as the sentence cited is synonymous to 'No students did not read all books', at least as far as the truth conditions are concerned.

REFERENCES

van Benthem, J. (1986): Essays on Logical Semantics, Dordrecht.

van Benthem, J. (1987): Polyadic quantifiers. University of Amsterdam, ITLI Prepublication Series 87-04.

Clark, R. — E. L. Keenan (1985/86): The Absorption Operator and Universal gramar, In: The Linguistic Review 5, 2, pp. 113-135.

Higginbotham, J. — R. May (1981): Questions, quantifiers and crossing, In: The Linguistic Review 1, 1, pp. 41-79.

Keenan, E. L. (to appear): N-ary Quantifiers.

de Mey, S. (to appear) : On the Semantics of Plurals.

Zwarts, F. (1984): Determiners: a relational perspective, In: Alice G. B. ter Meulen, ed., Studies in Model-theoretic Semantics, Dordrecht, pp. 37-62.

DEPARTMENT OF GENERAL LINGUISTICS

UNIVERSITY OF GRONINGEN, THE NETHERLANDS

SEBASTIAN LÖBNER

THE CONCEPTUAL NATURE OF NATURAL LANGUAGE QUANTIFICATION

Basically, there are two ways of describing a thing. I can describe a certain person A by saying what he is like, say, a tall and fat man in his thirties, or by describing the role he plays in some connection, introducing him, e.g., as the person who sold me the house in which I am living now. The latter would be a functional description, the former a sortal description. The two ways of referring to things though radically different, cannot be completely separated. A certain role or function presupposes certain qualities and, on the other hand, certain distinctive qualities may lead to a special role. An effective description will contain both functional and sortal features of the object. In this paper, I will try to formulate suggestions about the nature of natural language quantification. In the first part, I will say something about the function of quantifiers, and in the second, about what quantifiers are like. Because of the limited space available, reference will be made occasionally to other papers where I discussed some of the points relevant here in more depth and detail.

1. The case of definite plural and mass terms

Let me start the functional description of quantificational expressions with the consideration of simple sentences without quantifiers. From a logical point of view, the simplest sentences relevant here are combinations of a one place predicate with a definite argument term, i.e. sentences of the logical form p(a). This type is represented by sentences with a definite subject and a simple verb phrase. In the tradition of Montague (1973) and Barwise & Cooper (1981), definite NPs were treated as quantifiers along with genuinely quantifying NPs such as every mouse. But a closer analysis of the syntactic and semantic properties of definite NPs shows that they are individual terms in the sense of predicate logic rather than second order predicates. While this corresponds directly to intuition with respect to definite singular count terms, it might appear counterintuitive when applied to definite plural and mass or collective terms. Doesn't a term such as the children or the government refer to more than one individual and cannot hence be considered an individual term? The answer is no. The objection is invalid since it is due to a confusion of ontology and conceptual/logical content. From a logical point of view, definite plural and mass terms refer to what they refer to as one object, i.e. as an individual, regardless if it consists of several distinguishable parts. For a detailed argumentation, the reader is referred to Löbner (1987a) and, in particular, Löbner (1985). Link (1983) has provided a technical frame in which this analysis can be formulated.

I regard a predicate as a conceptual device which applied to an argument may yield one of two truth-values, say, 1 for "true" and 0 for "false". Natural language predicates, the meanings of verbs, nouns, adjectives and other expressions, obviously do not yield a truth-value for every argument whatsoever. They are conceptual instruments developed and appropriate for certain purposes, but inapplicable for others. There are, thus, in general truth-value gaps for every predicate, i.e. cases where for certain arguments the predicate yields neither of the two truth-values. For every predicate p, there is an opposite predicate p', the negation of p, which may or may not be lexicalized, if p is. p' yields a truth-value in exactly those cases where p does, but always the opposite one. Independently of predicate negation, sentences of the logical form p(a) (and, of course, of any other form) can be negated using an operator denoted here by the sign '-'. Sentence negation converts the truth-value of the sentence, provided it has a truth-value. There is a different sense of sentence negation, corresponding to a different notion of falsity, which assigns the value 1 also to those cases where the negated sentence lacks a truth-value. But this is not the type of negation which I am referring to here. Rather "negation" here always means the strong, presupposition-preserving variant. (Cf. Horn 1985 for the distinction between "internal", i.e. presupposition-preserving, negation and "external" negation.) Exactly what kind of presuppositions is preserved will be explicitly stated below. Obviously, in case of sentences of the form p(a), sentence negation and predicate negation exert the same effect on the truth-conditions: -p(a) is true, false, or truth-valueless iff p'(a) is true, false, or truthvalueless, respectively.

Hence, if we combine a definite plural or mass term with a simple predication, the effect of sentence and predicate negation should be the same, provided the analysis is correct. To check this, consider the following situation. Four pawns of a normal game of chess, and nothing else, is what the following sentences are about. These pawns, as usual, are each either white or black. With respect to these pawns, therefore, white and black yield opposite truth-values. The four sentences to be checked are

(1)	the pawns are white	d+p	(3)	the pawns are not white	-(d+p)
(2)	the pawns are black	d+p'	(4)	the pawns are not black	-(d+p')

I chose a neutral notation to combine d, which corresponds to the definite plural term, and the predicate term, in order not to anticipate the decision upon the logical status of d. Let case 1 be such that the pawns are black and let two of them be black and two white in case 2. Apparently we get clear truth-values in the first case, but not in the second:

(5)	case 1 ABCD		case 2 ABCD
	0	(1)	?
	1	(2)	?
	1	(3)	?
	0	(4)	?

In the mixed case, (1) and (2) are clearly not true; hence, (3) and (4) are not false. If (1) were false, (3) were true. But if the pawns are not white, they must be black, whence (2) would be true, which it isn't. Analogously, (2) cannot be false. Hence, both (1) and (2) lack a truth-value and consequently (3) and (4). It thus turns out that sentence negation and predicate negation do not differ if the argument place is filled by a definite plural term. The same is true of definite mass terms. Think of a sentence such as (6), together with its corresponding negations, in an analogous situation:

(6) the food is vegetarian

Apparently, besides eventual sortal restriction, predicates do not yield a truth-value when provided with an argument which is not homogeneous in terms of the relevant truth-criteria. In the case of definite plural terms a broad truth-value gap opens between the clearcut positive and negative cases.

(7)	**false**	(no truth-value)	**true**
	0		●
	00	0●	●●
	000	00● 0●●	●●●
	0000	000● 00●● 0●●●	●●●●

But this phenomenon is not restricted to definite plural or mass terms, as is shown by examples such as

(8) the Japanese flag is red (9) Istanbul is in Europe

The truth-value gaps are due to the existence of the following presupposition which applies to every predication whatsoever:

(10) **presupposition of argument homogeneity** (PAH)
The argument of a predication is homogeneous with respect to the predication.

PAH is what I would like to call a structural presupposition, as opposed to specific presupposition induced by certain lexical items. Another structural presupposition, e.g., is the presupposition of non-ambiguity of definite terms. PAH will play a role in different connection below again.

2. Fill the gap with quantification

The function of nominal quantification, now, is to bridge the gap between global falsity and global truth. If we replace the definite plural term in (1)-(4) by a quantificational NP, the truth-value gaps in the mixed cases are filled, as is demonstrated in (1q)-(5q):

(1q)	all pawns are white	$q+p$
(2q)	all pawns are black	$q+p'$
(3q)	not all pawns are white	$-(q+p)$ or $q'+p$
(4q)	not all pawns are black	$-(q+p')$ or $q'+p'$

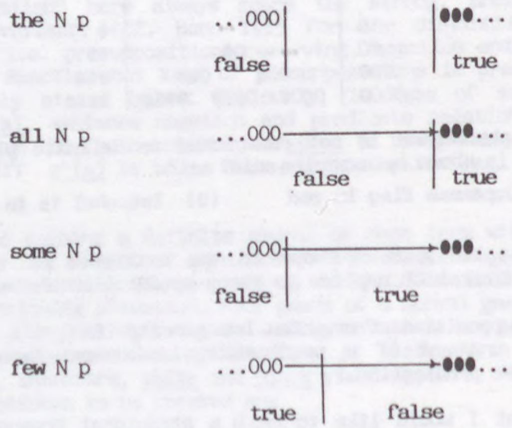
In addition to the disappearence of the truth-value gaps, (5q) shows that in the mixed case predicate and sentence (or equivalently quantifier) negation have different effects on the truth-value.

(5q)

case 1 ABCD		case 2 ABCD
0	(1q)	0
1	(2q)	0
1	(3q)	1
0	(4q)	1

If a predication is applied to an object which is complex in terms of the predication (in that it is possibly inhomogeneous), the two extremes of global truth and global falsity span a natural scale of possible cases. Quantificational sentences yield truth-values for all possible cases on the scale, in particular for those between the extremes. They usually do this in just cutting the scale into two parts, a negation and a positive range, as in the cases displayed in (11).

(11)



The quantifiers resulting in a bisection of the scale are those called monotone (cf. Barwise and Cooper 1981, van Benthem 1984, Löbner 1987b). They either provide a lower bound but no upper bound for the extent to which the predication applies or an upper bound but no lower bound. In addition more complex partitions of the scale can be constructed using expressions such as <u>some but not all</u>, <u>three or seven</u> etc.

The function of nominal quantifiers is thus the differentiation of an otherwise global application of a predicate to a complex object. The cases considered so far involve reference to a certain object such as a collection of pawns. In the count term cases this object constitutes what is traditionally called the domain of quantification. The underlying definite reference to the domain is implicit in sentences such as (1q) but explicit in other quantificational sentences.

(12) all the pawns are white
(13) some of the pawns are white
(14) the pawns are all white
(15) the pawns are partly white

The last sentence, and possibly the one before, contain an adverbial quantifier, which in view of the function of quantifiers appears to be the most natural way to express quantification. Note that sentence (15) is ambiguous between a group reading roughly equivalent to (13) and a distributive reading under which each single pawn is to some part white.

In Löbner (1987b) I have called this type of quantification "referential" as opposed to "generic" quantification which does not involve reference to the domain of quantification but the consideration of a totality of abstract cases. The difference between those two types, however, is not relevant for the following discussion and will not be pursued further. So far, this may suffice as a functional description of quantification. In what follows, I will try to give a sortal specification, sketching what appears to be the conceptual characteristics of natural language quantification. The discussion will start with an analysis of non-nominal quantifiers, which exhibit these characteristics in a more perspicuous manner.

3. From FALSE to TRUE (or vice versa): the dynamic characterization

There is a set of basic non-nominal quantifiers in natural languages for which I have coined the term "phase quantifiers" (cf. Löbner 1987b). These operators can be understood dynamically in the sense that they express the transition from a negative to a positive section (or phase) on some scale or vice versa, or the lack of a transition. Let me start with temporal quantifiers which illustrate the idea in a very direct way.

3.1 Transitions in time

In a sense, the concept of phase-quantification is prototypically represented by the basic meaning of <u>already</u> and its correlates <u>not yet</u>, <u>still</u> and <u>no more</u>. I have presented an extensive analysis of the German <u>schon</u> ("already") in its various uses elsewhere (Löbner, to appear) and I will therefore restrict myself here to a very brief sketch of the main idea. The basic use of <u>already</u> is the one as a sentence adverb in imperfective sentences.

(16) it is already dark

Imperfective sentences are predicates about a time of reference t' (cf. Löbner, to appear, Löbner 1987c). The logical structure of a simple imperfective sentence such as

(17) it is dark

is just a first order one-place predication p with a temporal argument, i.e.

(18) p(t')

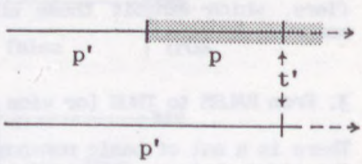
if, for the sake of simplicity, we neglect the contribution of tense to the sentence meaning. (17) is true if (16) is, but what already adds to the meaning of (16) is the condition that it is dark after not being dark. (17) is false iff it is not dark **yet**. The use of already presupposes a possible development from a state at which it is not dark to a state of darkness, i.e. from a negative phase in terms of the embedded sentence to a positive phase. (16) is true if the positive phase is reached at time t' and false if t' is in the initial negative phase. In general, a sentence of the form

(19) already(t', p)

is true if the reference time t' lies beyond the transition point of an initial negative phase in terms of p to a succeeding positive phase. It is false if t' lies before any transition point, i.e. if up to t' no transition has taken place. These conditions are illustrated in the following diagram:

(20) **already**(t',p) **true**

false



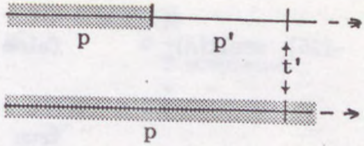
In contrast, the simple sentence (17) does not concern any development in time, but only one time t' and the respective state at that time in terms of p. Again, quantification - accepting for the moment that (16) is a case of quantification - means a differentiation between plain truth and falsity.

The adverb still, in those uses which are correlated to already, is in a twofold way opposite to already. The sentence

(21) it is still dark

presupposes a development from **darkness** to the contrary and states that the reference time t' is **not** beyond the transition point. Thus, still in this use functions as the dual of already, being equivalent to simultaneous negation of the embedded predicate p and negation of the whole. In the diagram this is reflected by the exchange of the positive and the negative phase, combined with the exchange of the truth-values for the whole statement.

(22) still(t',p) false

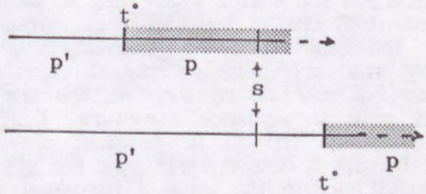


true

Another instance of temporal phase quantification is tense itself. The basic tense distinction is the one between past and non-past (cf. Comrie 1985), marked in English and in most other languages by the past and present tense, respectively. Tense locates the situation s expressed by the sentence relative to the time of utterance t°, either as past or non-past, non-past being a proximal and non-earlier phase with respect to t° and past a non-proximal, i.e distal, and earlier phase. Let p denote proximity in this sense and p' distality (see Löbner 1987c for details on the analysis suggested here.) [1]

(23) PRESENT(s)

PAST(s)



As time goes by, situations pass from the non-past to the past, and tense specifies whether the transition has taken place.

3.2 Transitions on other scale

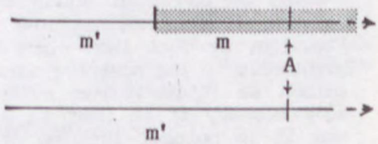
A further group of relevant cases are the scalar adjectives. They involve all sorts of scales, such as size, weight, length etc., which again are cut into two phases according to criteria which are strongly context-dependent (cf. Bierwisch & Lang (eds.) 1987 for an extensive analysis). A sentence such as

(24) A is big

expresses that according to some such criterion, A is marked in size in the sense of being bigger than comparable unmarked cases.

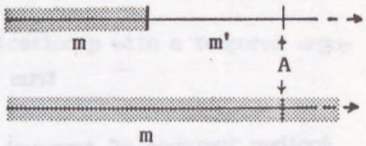
(25) big(A) true

false



(m for markedness). The antonymous adjective small is dual with respect to big, stating that the object is not bigger than unmarked cases.

(26) small(A) false true



Antonymous pairs of adjectives are furthermore related by the condition that the upper bound of markedness must be lower on the scale than the lower bound. The scale is parted into three sections marked - unmarked - marked. Let me mention in this connection that <u>many/few</u> and <u>much/little</u> are further instances to be analysed in this sense (cf. Löbner 1987b).

3.3 Nominal quantifiers

There are two scales which play an eminent role in natural language semantics. One is the time scale already dealt with, and the other is the quantity scale. Quantity specifications are incorporated into the morphology of nominal expressions in many languages and quantity quantifiers such as the English <u>every</u> and <u>some</u> play a key role in sentence structure. I will argue that the standard quantifiers (or to be precise: determiners, in the terminology of Barwise & Cooper 1981) <u>some</u> and <u>every</u> can also be seen as phase quantifiers in the sense illustrated in the previous section. Let B be a domain of quantification and P a predication. Let me abbreviate universal quantification over B with respect to P as $\underline{\forall(B,P)}$ and existential quantification as $\underline{\exists(B,P)}$. If one evaluates $\underline{\forall(B,P)}$, one has to check the elements of B one by one with respect to the predication P. The sentence is true if we encounter no counterexample throughout and false otherwise. Similarly, $\underline{\exists(B,P)}$ turns out false if we find no positive cases and true if we do. Van Benthem (1987) has therefore proposed to model $\forall$ and $\exists$ as finite automata which perform such an evaluation procedure. Any such procedure is based on some enumeration of the elements of B. (In case of mass term quantification we can think of an enumeration of the elements of a partition of B instead.) An enumeration is a scale. Obviously, it does not matter, which scale we use (in van Benthem 1984 this property of natural language quantifiers is called QUANT). What matters is the occurrence of positive or negative cases and not, when they occur. We are therefore free to restrict the consideration to canonical enumerations which have at most one switch between positive and negative cases. In case of existential quantification, we must then start with the negative cases and proceede eventually to the positive ones, whereas for the universal quantification we proceede from evidence to counterevidence. This order is necessary if we want to keep the outcome open when we start, and it is natural in view of epistemic considerations. Proceeding this way, we will or will not have crossed the critical transition between opposite cases when we arrive at the end of the enumeration. The resulting picture is analogous to the others given above:

(27) ∃(B,P) true 0 ~ 0 ~ 0 ~ 0 | 0 ~ 0

p' p ↑ B completed ↓

false 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0

p'

∀(B,P) false 0 ~ 0 ~ 0 ~ 0 | 0 ~ 0

p p' ↑ B completed ↓

true 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0

p

3.4 only

Only is a focussing particle. It operates on a partition of the sentence into foreground and background or focus and what I would like to call "cofocus". In general, the expression in focus is a predicate, which applies to something which is defined by the cofocus. Consider one example:

(28) Mary only loves John and two Italian linguists

Under the interpretation thought of here, only focusses on John and two Italian linguists. This is a predicate over something which is defined by the cofocus Mary loves –, (as ιx(Mary loves x)). Only adds the restriction to the sentence that the cofocus applies to nothing more than is explicitly specified in the focus. In view of the broad variety of uses of only, it generally appears reasonable to define "more" in turns of implicative scales: With respect to the cofocus C, A is more than B ("A C-more B"). To return to our example, in terms of the cofocus given here, "John and two Italian linguists" is more than "John and an Italian linguist", and this, in term is more than John alone, since if Mary loves John and two Italian linguists it follows that she loves John and an Italian linguist, and so on. We can now think of an evaluation procedure for only in the following way. The focus predicate F licenses certain objects in the role defined by the cofocus. We start from some object licensed by F and proceede in the upward direction of the implication scale up to the point, when the actual value of the cofocus is reached. If we remain within the range of objects licensed by F, the outcome is positive; if however we leave that range, the sentence is false. Let c be short for ιx(Cofocus x), and ⊃ for C-more, F^* for "licensed by F":

(29) only(c,F) false 0 ⊂ 0 ⊂ 0 ⊂ 0 | 0 ⊂ 0

F^* $F^{*\prime}$ ↑ c ↓

true 0 ⊂ 0 ⊂ 0 ⊂ 0 ⊂ 0 ⊂ 0

F^*

4. The formal representation

These results can be formulated in mathematical terms along the following lines. We will first define the crucial notion of an admissible chain and then give a general definition of phase quantifiers, which applies to all cases mentioned above.

4.1 Admissible chains

In all the cases considered above the predication quantified applies to elements of a partially or linearly ordered domain. The domain of quantification itself is in every case a chain, i.e. a linearly ordered subset. It has an upper bound, e.g. the time of reference or the last element of the enumeration, and it contains at most one transition between opposite elements with respect to the crucial predication.

(30) Definition

If P is a predicate with domain D(P), $<$ a partial ordering on D(P), e an element of D(P), then c is an **admissible** chain in terms of P, e, and $<$, for short $c \in AC(>,e,P)$, iff

(1) c is a chain with respect to $<$, i.e. a linearly ordered subset of D(P).

(2) e is the maximum of c.

(3) (optional) c starts with a negative phase of P: for some $x \in c$: if $x' < x$ and $x' \in c$, then $P(x')=0$.

(4) monotonicity: P is a monotone increasing function on c (in terms of 0 and 1): for every $x,x' \in c$: if $x < x'$, then $P(x) \rightarrow P(x')$.

Applied to the cases discussed above, the respective admissible chains are as follows:

already(t',p)	p-monotone time intervals (t,t'] for some t earlier than t'
still(t',p)	p'-monotone time intervals (t,t'] for some t earlier than t'
big(A)	m-monotone chains $\langle X,\ldots,A\rangle$ for some X smaller than A
some(B,P)	P-monotone enumerations of B
only(c,F)	F*'-monotone accumulations of c, F* being defined as: $F^*(x) \leftrightarrow \exists y(y \supset x \;\&\; F(y))$

The optional third condition is required in case of presuppositional quantifiers such as *always* and *only*, but not in the other cases. The crucial condition is the monotonicity constraint. Due to that, AC(>,e,P) is always a homogeneous class with respect to the property of containing positive elements. If any admissible chain contains positive elements in terms of P, then every chain does, because in this case the maximal element e common to all chains must be positive.

The monotonicity condition does not allow a transition from P-positive to P-negative elements in the chain. Hence, admissible chains can only differ in consisting of one or two phases in terms of P, and if there are two phases, then the negative phase comes first. Admissible chains have just the minimal length required in order to present a nontrivial alternative.

4.2 Phase quantifiers

The meaning of a phase quantifier, now, is that the respective admissible chains do or do not contain positive elements. To formulate this accurately, we need one further notion, a quantifier $\exists\forall$ tantamount to "all and any":

(31) *Definition*
For any first order predicate logic formulae ϕ and ψ: if
(i) $\exists x\ \phi$
(ii) $\exists x(\phi \& \psi) \leftrightarrow \forall x(\phi \rightarrow \psi)$
then $\exists\forall x(\phi : \psi) \leftrightarrow_{df} \exists x(\phi \& \psi)$

In particular, $\exists\forall x(B(x):P(x))$ means that "the Bs are P". The sentence is true iff all Bs are P and false iff all Bs are not P. Obviously, $\exists\forall$ is not always defined. It presupposes the homogeneity of B with respect to P. $\exists\forall$ is self-dual, i.e.

(32) $-\exists\forall x(\phi : \psi) \leftrightarrow \exists\forall x(\phi : -\psi)$

We are now in the position to define the four possible phase-quantifiers, given any predicate P, a partial ordering $<$ on the domain of P, and an element e out of the domain:

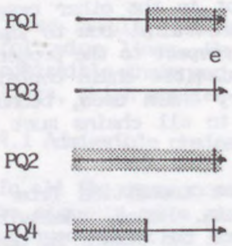
$PQ1(>,e,P) \leftrightarrow_{df} \exists\forall c(c \in AC(>,e,P): \exists x(x \in c \ \&\ P(x))$

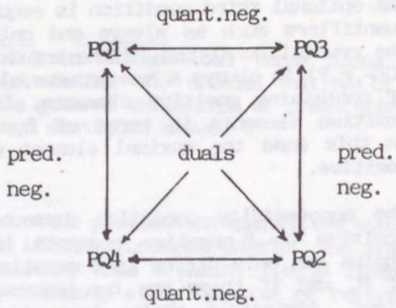
$PQ3(>,e,P) \leftrightarrow_{df} \exists\forall c(c \in AC(>,e,P): -\exists x(x \in c \ \&\ P(x)) \quad = PQ1'(>,e,P)$

$PQ2(>,e,P) \leftrightarrow_{df} \exists\forall c(c \in AC(>,e,P'): -\exists x(x \in c \ \&\ P'(x)) \quad = PQ1'(>,e,P')$

$PQ4(>,e,P) \leftrightarrow_{df} \exists\forall c(c \in AC(>,e,P'): \exists x(x \in c \ \&\ P'(x)) \quad = PQ1(>,e,P')$

The four phase quantifiers represent the four possibilities displayed below and form a duality square:





Applied to the examples above, the definition yields:

already(t',p)	=	PQ1(later,t',p)
still(t',p)	=	PQ2(later,t',p)
not yet(t',p)	=	PQ3(later,t',p)
no more(t',p)	=	PQ4(later,t',p)
PRESENT(s)	=	PQ1(later,s,prox)
PAST(s)	=	PQ2(later,s,prox)
big(A)	=	PQ1(bigger,A,marked)
small(A)	=	PQ2(bigger,A,marked)
∃(B,P)	=	PQ1(after,b#,P)
∀(B,P)	=	PQ2(after,b#,P)
only(c,F)	=	PQ2(C-more,c,F*)

(b# = last element of B in the enumeration)

In Löbner (1987b) I have emphasized the importance of duality relations for the semantic analysis. The four types are in a natural way related to monotonicity and persistency properties (see Löbner 1987b: 76f). It appears that the type assignment is also significant in a different sense: crosslinguistic evidence suggests that the four types form a descending chain in terms of the frequency of proper lexical items and an ascending chain in terms of markedness in several regards. According to the definition of $\underline{\exists\forall}$, the phase quantifiers are only defined if the set of admissible chains is not empty and homogeneous with respect to the property of the existence of P (or P') -positive elements in the chain. AC can only be empty due to violation of the optional condition (3). In case of <u>already</u> and <u>only</u> this condition in fact yields the relevant presuppositions (cf. Löbner to appear for <u>already</u>). The homogeneity condition is always fulfilled due to the monotonicity constraint on admissible intervals.

In a sense, PAH is at work here again. These quantifiers are essentially predicates about the admissible chains. The chains that come into consideration are all alike with respect to the predication. We can chose a single chain for evaluation, and the result will not depend on the choice. But this is only possible if the set of relevant cases is kept homogeneous.

NOTES

1 The analysis reported here is a slight modification of the ideas presented in Löbner (1987c), but is in accordance with Löbner (to appear). I consider situations as pairs of a facts and a time component. Under the perfective aspect, the factual component is an event, which is located relative to the time of utterance t^*. Under the imperfective aspect, the temporal component, i.e. the reference time called t' above, is located on the time scale.

REFERENCES

Barwise, Jon & Robin Cooper: 1981, 'Generalized Quantifiers and Natural Language', **Linguistics and Philosophy** 4: 159-219.

van Benthem, Johan: 1984, 'Questions about Quantifiers', **Journal of Symbolic Logic** 49: 443-466.

- : 1987, 'Semantic Automata', in J. Groenendijk, D. de Jongh, M. Stokhof (eds.), pp. 1-25.

Bierwisch, Manfred & Ewald Lang (eds.): 1987, **Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven**. Akademie-Verlag, Berlin.

Comrie, Bernard: 1985, **Tense**, Cambridge University Press, Cambridge.

Groenendijk, Jeroen & Dick de Jongh & Martin Stokhof (eds.): 1987, **Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers**, Foris, Dordrecht.

Horn, Laurence R.: 1985, 'Metalinguistic Negation and Pragmatic Ambiguity', **Language** 61: 121-174.

Link, Godehard: 1983, 'The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretical Approach', in R. Bäuerle, Chr. Schwarze, A. von Stechow (eds.) **Meaning, Use, and Interpretation of Language**, De Gruyter, Berlin, pp. 302-323.

Löbner, Sebastian: 1985, 'Definites', **Journal of Semantics** 4: 279-326.

- : 1987a, 'Natural Language and Generalized Quantifier Theory', in P. Gärdenfors (ed.) **Generalized Quantifiers: Linguistic and Logical Approaches**, Reidel, Dordrecht, pp. 181-201.

Löbner: 1987b, 'Quantification as a Major Module of Natural Language Semantics', in J.Groenendijk, D. de Jongh, M. Stokhof (eds.), pp. 53-85.

- : 1987c, 'Ansätze zu einer integralen Theorie von Tempus, Aspekt und Aktionsarten', in V. Ehrich, H. Vater (eds.) **Temporalität**, Niemeyer, Tübingen.

- : to appear, 'German schon - erst - noch: An Integrated Analysis', **Linguistics and Philosophy**.

Montague, Richard: 1973, 'The Proper Treatment of Quantification in Ordinary English', in J. Hintikka, J. Moravcsik, P. Suppes (eds.) **Approaches to Natural Language**, Reidel, Dordrecht, pp. 221-242.

SEMINAR FÜR ALLGEMEINE SPRACHWISSENSCHAFT

UNIVERSITÄT DÜSSELDORF, DÜSSELDORF, FRG.

LÁSZLÓ KÁLMÁN

GENERICS, COMMON-SENSE REASONING, AND MONOTONICITY IN DISCOURSE REPRESENTATION THEORY

0. Introduction

It is a commonplace today that so-called "context-change" theories of semantics, conceiving of sentence meaning not as purely propositional but as contributing to contexts both a proposition and other constraints of various nature, capture important generalisations not sufficiently emphasised in standard model theoretic frameworks. I do not deny this fact; what I shall try to demonstrate in what follows is that the so-called Kamp—Heim framework of context-change semantics still oversimplifies the relationship of sentence meaning to context and interpretation.

In Section 1 of this paper I shall review, following the ideas of Partee (to appear), issues related to "semantic competence" and its status in model theoretic semantics as well as a possible compromise between basic principles of model theoretic semantics and a "mentalistic" view of semantic competence. In Section 2 I shall present a class of phenomena pointing to the fact that a "semantic competence" level of representation mediating between sentence meaning and interpretation is called for. The phenomena in question are generic sentences on the one hand and common-sense (default) reasoning on the other. I shall briefly characterise the mappings involved in the transition from sentence representation to interpretation, and conclude that discourse knowledge is underdetermined by the input sentences and the propositional content determining truth conditions. Finally, in Section 3, I shall outline the essential properties of the model proposed as related to phonological and syntactic processes. I shall argue that we may be able to model default reasoning without recurring to (non-monotonic) default logic in the same way as we can use (monotonic) unification while taking care of apparently non-monotonic features of linguistic processes.

1. Semantic competence and model theoretic semantics

The analysis of current semantic trends presented in this section is basically the same as the one in Partee (to appear). I cover some of the most important points of that analysis in order to emphasise certain aspects that remain in the background of Partee's paper but will be of crucial importance in the following. All quotations in this section are from her paper.

According to Partee, "a strictly model theoretic approach with no intermediate representational level is committed to denying that semantic competence involves in any systematic way the syntactic manipulation of expressions in some language of semantic representation". Although this observation is undoubtedly true, we have to point out that the Kamp—Heim framework does not depart from model theoretic semantics in this respect. Both Kamp (1981) and Heim (1982) posit an intermediate representational level, which they call **discourse representation structure** (DRS) and **card file**, respectively, but neither theory allows "in any systematic way the manipulation of expressions" in those representation languages.

There are two extreme interpretations of "semantic competence" that we have to consider when assessing the Kamp—Heim position. The first, "conceptualist", view, related to the Chomskyan tradition, claims that, "as far as the semantics of natural language goes, criteria of adequacy of a given theory derive ultimately from the attempt to account for what is 'in the head' of the competent native speaker of a language". The second, "anti-psychologist", view, inherited by the Fregean tradition, emphasises that "translating from one uninterpreted language to another gets one no closer to the first thing about the **meaning** of a sentence, namely its truth conditions"; its aim is, thus, to account for speaker-independent truth-conditional phenomena.

In my understanding of Partee's paper, she suggests that a feasible compromise between these two extremes is to stick to the claim that "all that really matters for semantics is what the language user **thinks** the sentence means and that in turn can be characterised by the state of his mind viewed as a machine with a lot of internal states functionally related to each other", and yet be concerned with "the question of how one can attribute informational content to internal states of a computer", and use the answer thereto to provide internal mental states, described in terms of propositions, with model theoretic interpretation.

The alternative described above can actually be found in current artificial intelligence and computational linguistic literature. The studies in question (e.g. Etherington—Reiter (1983), Touretzky (1982)) either emphasise the need for using non-monotonic logic to model common-sense reasoning, or design models of mental processes equivalent to some non-monotonic logic. This way, the solution of the dilemma related to the two extreme views of semantic competence is directly linked to our approach to reasoning and the role of systems of beliefs in interpretation: irrespective of whether we are interested in the issues raised by artificial intelligence and computational linguistics, we have to account for the linguistically relevant mechanisms of reasoning that seem not to fit within the monotonic model theoretic framework.

In the next section I shall propose a solution to this anomaly by assuming that the intermediate representations are accessible for other sorts of knowledge and for mechanisms different from sentence processing proper. That is, I shall argue that reasoning, interaction with background knowledge, and some kinds of disambiguation proceed on the DRS level of representation rather than on the model theoretic level, and their outputs are not always subject to interpretation.

2. Generics and background knowledge in a context-change framework

2.1. A version of discourse representation theory

The version of discourse representation theory (DRT) I proposed in Kálmán (1986) heavily relies on the computational metaphor. In that version, sentences are translated into **DRS-changing instructions** that are first **incorporated** into the DRS and then performed any time. A DRS is a richly structured set of **discourse referents** conceived of as **pegs** in the sense of Landman (1986) with **predicates** assigned to them. The semi-lattice structure of discourse referents is treated in detail by Pólos (1987). Both DRS-changing instructions and predicates are **programs** whose execution may effect changes in the DRS. A program can only be run if it is sufficiently specific. Programs can be classified, in terms of their effects, into knowledge-retrieving, discourse-referent-introducing, predicate-assigning, etc. programs. Although there are no type distinctions among programs, their relationship to discourse referents can also be classified informally: a discourse referent can correspond to the location, some argument, etc. of a predicate.

For example, the translation of an indefinite noun phrase (NP) such as a man can be represented as follows:

(1) Representations of a man

a. Kamp—Heim:

x_n

cond: (i) man'(x_n)

(ii) x_n "new"

b. Kálmán (1986):

INTRODUCE(man')

((1a) is a rough representation of the idea that processing a man leads to the introduction of a new constant having the property "man".) Although the execution of the program in (1b) may lead to a result analogous to (1a) in an empty DRS, its content is clearly different. This becomes evident if we consider its eventual interaction with other programs. For example, if an indefinite NP is part of a predicate verb phrase (VP), then the corresponding discourse-referent-introducing instruction will be part of the predicate assigned to the discourse referent corresponding to the subject of the sentence. I shall represent this by a " : " sign. So (2) below will be translated as something like (2'):

(2) Ambiguous sentence

John and Mary carried a piano upstairs.

(2') DRS representing (2)

John-and-Mary' : carried-upstairs'(INTRODUCE(piano'))

Note that (2') is as "ambiguous" as (2) itself. The two readings of the original sentence can be paraphrased as in (3):

(3) a. Paraphrase of (2)

'John carried a piano upstairs and Mary carried a piano upstairs'

b. Paraphrase of (2)

'There is a piano that John and Mary carried upstairs'

Whether (2) is disambiguated or not is a matter of previously or subsequently acquired discourse or other knowledge. There are several ways in which executing programs corresponding to that knowledge can take care of incorporating the translation of either (3a) or (3b) in the DRS; the details of these are irrelevant for the present discussion. The only point is that, in the Kamp—Heim framework, the only way of accounting for the ambiguity of (2) is to allow two alternative representations; consequently, deciding which one is correct can be a matter of previous contextual knowledge, but cannot depend on subsequently incorporated information. I believe this is not adequate in terms of predicting "what is going on in the speakers' mind" although that treatment might account for the possible truth conditions of (2).

In sum, if we allow DRS-changing instructions as programs or "problems" to be resolved to be included in the DRS (instead of including the solution of those "problems" and the effects of those programs only), then a DRS will eventually contain **underspecified representations**. This has the immediate advantage of allowing subsequent information to disambiguate previously incorporated information, and less direct consequences bearing on models of reasoning and non-truth-conditional aspects of linguistic processing to be treated in the next sections. Of crucial importance is the fact that, in the approach proposed above, a DRS may contain information that contains "unresolved problems" not subject to interpretation.

2.2. Donkey sentences and generics

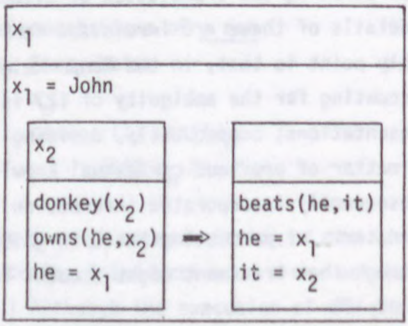
One of the main achievements of the Kamp—Heim theory seems to be the correct treatment of so-called "donkey sentences" such as the one in (4):

(4) Donkey sentence

If he owns a donkey, he beats it.

The effect of (4) on the DRS can be represented as follows in Kamp's notation, provided that the antecedent of he in (4) is John:

(4') DRT representation of (4)



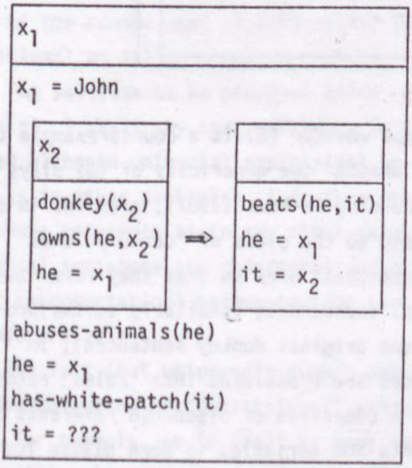
Kamp's claim is that universal accessibility conditions allow the anaphora in the right-hand side of the conditional to find their antecedents in either the left-hand side or outside the conditional; anaphora in the left-hand side can find their antecedents outside the conditional; the converse is not true in any of these cases. For example, (5a) but not (5b) is a valid continuation of (4), as illustrated in (5'), where I merged (4') and the representation of (5a,b).

(5) Valid and invalid continuations of (4)

a. He always abuses animals.

b. # It has a white patch on its forehead.

(5') DRT representation of (4) and (5a,b)



Several apparent counterexamples to the accessibility conditions have been proposed since the theory was put forward. One of them is the case when the generic donkey sentence is followed by another generic, as in (6):

(6) Generic donkey sentence followed by another generic

Every man who loves a woman respects her. He treats her as his equal.

As a matter of fact, (6) is not a counterexample provided we treat the second sentence as belonging to the right-hand side of the conditional:

(6') DRT representation of (6)

x_1 x_2 man(x_1) woman(x_2) loves(x_1,x_2)	⟹	respects(x_1,her) her = x_2 treats-as-equal(he,her) he = x_1 her = x_2

The important issue here is not whether (6) is a counterexample to accessibility conditions but whether the genericity of (6) plays any role in this porblem. According to Hess (1987), pronouns in (6) and in similar sentences belong to the class of "descriptional" anaphora as opposed to "denotational" ones in that they refer to "stereotypes" rather than real individuals (similarly to the proposal of Bartsch (1979) for the original donkey sentences). In Hess' theory, generic sentences are translated into "rules" rather than ordinary propositions. He conceives of discourse referents as the generalisation of constants and variables in Horn clause logic.

There is an important drawback in Hess' analysis, apart from the fact that he needs to multiply the types of discourse referents. Consider the following sentences:

(7) a. Reproduction of the first sentence in (6)
Every man who loves a woman respects her.
b. Non-generic universal sentence
Every hunter who came with us yesterday killed a deer.

In (7b) we commit ourselves to a claim on the actual set of hunters who came with us yesterday; the quantified NP in this case stands for a group referent, hence we can only use plural ("denotational") anaphora, which the Hessian account takes care of.

(8) Valid and invalid continuations of (7b)

a. They cut them up and carried them home.

b. # He cut it up and carried it home.

It is not clear, however, why (7a) can have a group reading as well, and why the sequence in (6) is correct on the rule reading only (as a matter of fact, a plural continuation would work for both readings, but this is irrelevant for the present problem).

In the Kamp—Heim framework, on the other hand, it is unclear how it can be avoided for (8b) to be included in the right-hand side of the conditional of (7b) or for (7b) to be interpreted as a conditional at all.

The solution to be proposed below relies on the same intuition as Hess'. I shall say that reference in "rules" is not to be resolved although universal expressions in generics are of the same type as in other sentences, i.e. they would "evaluate to" group discourse referents as in any other case when resolved. In sum, universal sentences are underspecified at the outset, and thus their representations belong to that part of the DRS that is not to be interpreted.

The fact that universals should not be interpreted until they are disambiguated as "denotational" (thanks to the sentence adverbial, for example, as in (7b)) is best shown by the fact that their "propositional content" can easily be overridden, by except for...-clauses, for instance. Therefore, they can function as default statements that can be used in reasoning but should not be assigned truth conditions. This points to the fact that apparently non-monotonic reasoning processes take place on the DRS level rather than the propositional level.

2.3. Common-sense reasoning and monotonicity

Implicit in what I said so far is the existence of a sequence of four mappings involved in semantic processes: (i) **translation**, the mapping from some representation of the sentence and the lexicon to a DRS-changing instruction (i.e. the **semantic representation** of the sentence); (ii) **execution,** the mapping (from the class of DRSs to the class of DRSs) embodied in the DRS-changing instruction; (iii) **filtering,** the mapping from a DRS to a set of propositions to be actually interpreted; (iv) **interpretation,** the mapping from this set of propositions to a set of possible worlds. There are well-motivated and well-established assumptions on the nature of translation, e.g. compositionality, and of interpretation, e.g. its model theoretical character, which I shall not discuss nor question here. As for "filtering", I shall assume that this mapping is essentially governed by specificity, i.e. that a piece of information in the DRS gets into the set of propositions to be interpreted just in case it is fully specified in some sense (e.g. it should not contain unresolved reference problems). In the rest of this section I shall be concerned with the nature of "execution".

If a speaker uses the verb marry in a conversation, this allows the participants to use a number of anaphoric expressions (the bride, the wedding ceremony, the dowry, etc.) exactly as if the antecedent had been introduced. That is, components of the semantic representation of a sentence (in this case, the mysterious marry' looked up in the lexicon) somehow point to a stack of knowledge stored somewhere (either in the lexicon itself or in some separate background knowledge representation). The fact that linguistic phenomena such as the **accomodation** of the definite descriptions above require that the speaker have permanent access to that stack of knowledge led Bartsch (1987) to the conclusion that the knowledge in question is directly incorporated in the DRS.

In fact, as I argued elsewhere (Kálmán (1987)), the mechanisms of accessing background knowledge to resolve reference problems seem to work in exactly the same way as the ones serving to access discourse knowledge. But incorporating background knowledge in the DRS whenever it is referred to would lead to enlarging our DRS to an uncontrollable extent: incorporated propositional knowledge may contain new references to background knowledge, e.g. the default location of a wedding ceremony may be a church, and nothing would preclude the incorporation of all the background knowledge connected to church' in the sequel. Nevertheless, since the transition from discourse knowledge to background knowledge is extremely smooth, it is rather useful to conceive of background knowledge as a sort of "larger", "older", "less active" discourse knowledge our DRSs are actually embedded into.

Obviously, we have to avoid the interpretation of entire and possibly inconsistent systems of beliefs. Since most of our background knowledge is of the generic or otherwise underspecified kind, this is automatically resolved under the treatment of generics proposed in the previous section. In Bartsch (1987), the piece of information incorporated in the DRS as a result of processing a lexical item, called "MP—SF" (meaning postulate—scenic frame), just has the label "Default" by virtue of which it can be overridden, and this is exactly what we get under the present treatment.

The conclusion is that linguistic expressions licensed in a certain context depend not only on the portion of DRS that is to be interpreted. For example, the anaphoric expressions licensed by marry depend on our background knowledge of this word, whether or not this is included in the DRS in the form of "default" assertions or not. In any case, reasoning is needed to access the appropriate piece of knowledge, and this involves "default" assertions as well. In the case of marry, for instance, any particular instance of marriage will as a default **inherit** the assertion of the existence of its protagonists. In sum, linguistically relevant, though truth-conditionally irrelevant pieces of reasoning must proceed on the DRS level (before "filtering").

Predicate inheritance, i.e. the repeated application of modus ponens, is one of the most important machanisms of common-sense reasoning. According to Etherington—Reiter (1983), however, inheritance systems with defaults and overriding exceptions give rise to non-monotonicity. Thus, "execution" should be non-monotonic: if some assignment "s : p" is entailed in a DRS, then it is not necessarily the case that this assignment is also entailed by the image of the DRS under an arbitrary DRS-changing instruction. On the other hand, since only underspecified assignments can be overridden, this may happen to underspecified assignments only, hence "filtering" will eliminate non-monotonicity. If a proposition p is entailed by a set of propositions resulting from "filtering", then the set of propositions resulting from a subsequent mapping will also entail p.

I have to emphasise that I neglect the fact that human beings change their minds in rather radical ways, especially when they interact with each other. DRS-changing instructions explicitly violating monotonicity (e.g. those changing an assignment "s : p" to, say, "s : NOT(p)") must not be considered as influencing the reasoning process but rather as restarting it. Another important point about default reasoning is that conclusions drawn from default assertions preserve their overridable nature. Consider the following DRS:

(10) DRS allowing default conclusion
  (i) elephants : grey
  (ii) Clyde : elephant

The conclusion "Clyde is grey", although specific in itself, may be overridden by an explicit assertion of, say, Clyde's being pink. That is, reasoning is preserved together with the conclusion:

(11) DRS (10) after drawing the conclusion
  (i) elephants : grey
  (ii) Clyde : elephant; grey by (i)

(12) DRS (11) after incorporating "Clyde : pink"
  (i) elephants : grey
  (ii) Clyde : elephant; pink
  (The previous conclusion by (i) is to be cancelled because the newly incorporated assignment is more specific.)

To sum up, since common-sense reasoning relies on default knowledge in the DRS and in background knowledge, and we want to avoid for default knowledge to be interpreted and thus contribute to truth conditions, it is necessary to allow reasoning mechanisms to operate on the DRS-level representation, therefore restrict the strictly model theoretic and strictly monotonic machinery to the interpretation phase of processing.

3. Defaults and monotonicity in phonology and syntax

Some of the issues mentioned so far are familiar from phonology and syntax, others seem to be unrelated to those domains of grammar. A feature certainly shared by all those components is the extensive use of representations describing different linguistic phenomena.

A phonological representation describes an articulated sequence of partially overlapping articulatory gestures associated with a linguistic unit; a semantic representation describes the meaning of a linguistic unit; and syntactic representations describe patterns of meaningful linguistic units. One thing we require to hold for all of them at this level is monotonicity: enrichment of, i.e. increasing information on, a representation should lead to non-decreasing information on the objects described. This is usually assumed for phonological and syntactic representations, and holds for the sets of propositions representing a piece of knowledge or for a DRS-changing instruction representing the meaning of a sentence.

Another well-established principle is that some notion of "full specification" plays an important role in all domains of grammar.

Only the fully specified parts of a phonological representation can be "phonetically implemented" and, similarly, sentence represenations can be implemented only after "lexical insertion" and "phonological interpretation". What "implementation" would mean in semantics is rather unclear; nevertheless, logical interpretation, although not an observable human activity, may require the full specification of a set of propositions. On the other hand, only a fully specified DRS-changing instruction can be implemented as a sentence with that meaning. For a statement of this principle, cf. Sanders (1971).

A third major principle grammatical components should share consists in the overriding power of specific rules and representations over less specific ones.

Let me illustrate this principle by examining its version as stated in Kiparsky (1973) for phonology. The Elsewhere Condition states that, if the structural description of a phonological rule properly includes the structural description of another rule, then the one with more specific application conditions applies whenever possible, and the other one will apply just in the remaining cases ("elsewhere"). Thus, Kiparsky's phonology could be non-monotonic in a certain sense. Given two rules, a and b, with b being more specific than a, then, provided a applies to a representation x, and thus a(x) is "entailed", it could be the case that enriching x will lead to the applicability of b (and the inapplicability of a); in that case, b(x') rather than a(x) would "follow" (where x' is the enriched representation). The reason why this does not happen and why phonology is still well-behaved lies exactly in the principle of full specification. In the above case, "if not(x') then a(x)" rather than a(x) is "entailed", where "not(x')" means that x is fully specified w.r.t. a and b, i.e. it cannot be enriched to yield x'. In Lexical Phonology, for instance (Kiparsky (1982)), where this principle holds, enriching representations is done step by step, and phonological rules apply in blocks after each step of enrichment, when representations are fully specified w.r.t. the rules of the given level.

Since there is no evidence for similar levels in semantics, we probably have to use a different mechanism, e.g. the one proposed in the previous section (reasoning preservation). The fundamental difference of the two systems, however, lies not in this fact, but rather in the lack of interference with background knowledge in phonology. In any case, monotonic and apparently non-monotonic aspects of phonology nicely parallel what is going on in semantics under the approach proposed in this paper.

Syntax also seems to break up into a monotonic and an apparently non-monotonic domain. The interference of syntactic and even semantic knowledge on particular lexical items and different other entities on the continuous scale of idiomaticity up to fairly regular syntactic patterns seems to follow principles similar to common-sense reasoning and thus give rise to the same kinds of effects of non-monotonicity (cf. Small (1981)), whereas the unification operation whose input is a set of representations with such conflicts already resolved is assumed to be monotonic in all current theories. It is not at all clear for the moment whether the system of syntax is comparable to either phonology (as in the transformational tradition) or semantics (as suggested in the first part of this paragraph) or both (as in Sanders (1971), who proposes a uniform treatment of all three components by putting forward a principle equivalent to the Elsewhere Condition, called Proper Inclusion Precedence).

In conclusion, monotonicity seems to characterise the edges and interfaces of the modules of grammar, whereas their internal machineries seem to obey the Elsewhere Condition with possibly module specific mechanisms probably ensuring monotonicity.

ACKNOWLEDGEMENT

I would like to thank Edith Moravcsik and Anna Szabolcsi for their valuable comments on an earlier version of this paper.

REFERENCES

Bartsch, R. (1979) "The Syntax and Semantics of Subordinate Clause Constructions and Pronominal Coreference," in F. Heny and H.S. Schnelle, eds., Selections from the Third Groningen Round Table: Syntax and Semantics, Academic Press, New York, 23-59.

Bartsch, R. (1987) "Lexical Frames and Discourse Representations," paper read at the Conference on Logic and Linguistics, Stanford University, July 11.

Etherington, D.W. and R. Reiter (1983) "On Inheritance Hierarchies with Exceptions," Proc. AAAI-83, 104-108.

Heim, I. (1982) The Semantics of Definite and Indefinite Noun Phrases, doctoral dissertation, UMass, Amherst.

Hess, M. (1987) "Descriptional Anaphora in Discourse Representation Theory," in Proc. ACL 3rd European Chapter Conference, Copenhagen.

Kamp, H. (1981) "A Theory of Truth and Semantic Representation," in J.A.G. Groenendijk et al., eds., Formal Methods in the Study of Language I, Matematisch Centrum, Amsterdam, 277-322.

Kálmán, L. (1986) "Discourse Representation and Suggestion Semantics," ms., ISSCO, Geneva; in Hungarian: Tertium non datur 4(1987):101-116.

Kálmán, L. (1987) "Handling Word Meaning in DRT," paper read at the Symposium on Lexical Semantics, Stanford University, August 2.

Kiparsky, P. (1973) "'Elsewhere' in Phonology," in S.R. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle, Holt, Rinehart, and Winston, New York.

Kiparsky, P. (1982) "From Cyclic to Lexical Phonology," in H. van der Hulst and N. Smith, eds., The Structure of Phonological Representations. Part I, Foris, Dordrecht, 131-176.

Landman, F. (1986) "On Pegs and Alecs," in Towards a Theory of Information: The Status of Partial Objects in Semantics, Foris, Dordrecht.

Partee, B.H. (to appear) "Possible Worlds in Model Theoretic Semantics: A Linguistic Perspective," in S. Allén, ed., Possible Worlds in Humanities, Arts, and Sciences, Proc. of Nobel Symposium 65, Walter de Gruyter and Co., Berlin and New York.

Pólos, L. (1987) "DRT and the Structured Domains (Typed or Type-Free?)" in this volume, 111-120.

Sanders, G.A. (1971) "On the Symmetry of Grammatical Constraints," in 7th Regional Meeting of the Chicago Linguistic Society, University of Chicago, Dept. of Linguistics, 232-241.

Small, S. (1981) "Viewing Word Expert Parsing as Linguistic Theory," Proc. IJCAI-81, 70-76.

Touretzky, D.S. (1982) "Exceptions in an Inheritance Hierarchy," ms.

INSTITUTE OF LINGUISTICS
HUNGARIAN ACADEMY OF SCIENCES

LÁSZLÓ PÓLOS

DRT AND THE STRUCTURED DOMAINS

(TYPED OR TYPE-FREE?)

1. INTRODUCTION

It has been a major trend during the last decade that people in formal semantics turned to modeling semantic relations for which sentence boundaries are penetrable. An example for such a relation can be the next intersentential anaphor:

(1) A GOOD-LOOKING YOUNG MAN ENTERED THE ROOM.

(2) EVERYBODY WAS SURPRISED WHY HE WAS WALKING ON HIS HANDS.

It is quite natural to consider the sentence (2) as if it contained additional information about the GOOD-LOOKING YOUNG MAN mentioned in the sentence (1).

There is no *trivial* possibility for the extension of the findings of classical model-theoretical semantics to such relations. (But anyhow there are some possibilities.)

The most natural way of the extension, and I think, the only *natural* one, is the following: handle the whole text - the discourse - as if it were only one sentence, as the conjunction of the sentences occurring in the discourse. Unfortunately, this way proves to be a dead end. As we know well from elementary logic courses, conjunction is associative and commutative, and these two properties are sufficient to show that a given discourse is logically equivalent to any permutation of its sentences, and this is obviously false.

(3) THERE WAS NOBODY IN THE ROOM.

(4) A MAN ENTERED THE ROOM.

(5) HE WAS WAITING FOR HIS WIFE.

If we mix these three sentences, the meaning of the whole text changes and some permutation of the sentences seem to result in meaningless discourse. We can decipher essential semantic relations from the order of sentences (see, e.g., KAMP 1983, PARTEE 1984).

DRT is a framework which handles intersentential anaphoric chains nicely. Its handling is based on the presupposition that the connexion between anaphoric expressions and their antecedents is their coreferentiality. This coreferentiality is a local relation: it is supported only by a given context. In order to matching coreferential expressions DRT needs a way of constructing formal models of contexts. And context models seem to be very useful for handling several problems of context dependency.

2. DRT: A POSSIBLE APPLICATION

Some thirty years ago QUINE realized (QUINE 1960) that such expressions as WATER or INFORMATION form a strange class of common nouns. Sometimes they behave as predicates:

(6) THE FLUID IN THAT BOTTLE IS WATER.

(7) WHAT THIS NEWSPAPER CERTAINLY DOES NOT CONTAIN IS INFORMATION.

In some other cases they look like names:

(8) WATER IS WET.

(9) INFORMATION IS VALUABLE.

Another observation was that WATER or INFORMATION resist pluralization. This shows that in the case of this class of nouns the usual explanation of such problems does not work. The "general explanation" goes as follows:

(10) MAN IS MORTAL.

(11) SOCRATES IS A MAN.

In (10), MAN is a predicate, as well as in (11), because (10) can be translated into (10'):

(10') THE SET OF MEN IS A SUBSET OF MORTALS.

Furthermore, (11) actually means that

(11') SOCRATES IS A MEMBER OF THE SET MEN.

One cannot even name set of "WATERS" without pluralization, and it is far from being obvious what constitutes the members of such a set. They need a different semantic representation.

These problems led to the principled distinction of count/mass nouns. (See, e.g., PARSONS 1968.) This distinction is intuitively clear but not sharp. PELLETIER had shown in 1975 that substantially any noun can be used as a mass noun. (The basic idea of his argumentation, the idea of a *universal grinder* goes back to David Lewis.) And vice versa, mass nouns as WINE can be used as count ones:

(12) THERE ARE FOUR DIFFERENT WINES IN MY CELLAR.

Now, if mass nouns are to be modeled semantically in a different way that is different from modeling count nouns then we should decide before semantic modeling, if the noun in questions belongs to mass or count nouns. The lexicon does not give sufficient information in this respect. Sometimes there are syntactic markers to qualify a given occurrence of a certain common noun. A noun preceded by a numeral or an indefinite article is recognised as count noun as well as a plural noun, while a measure expression may indicate a mass noun, but it does not guarantee that it actually *is* a mass noun.

Some problematic sentences:

(13) I WOULD LIKE TO GET FOUR POUNDS OF APPLES;

(14) I FOUND A ONE-INCH DWARF ON MY TABLE.

In (13), the expression APPLES is a mass noun perhaps, but in plural, while in (14), DWARF is not.

H. C. BUNT (BUNT 1978) mentioned several syntactic criteria mass nouns. But even he knows well that his criteria do not qualify all possible cases. BUNT has a proposal to solve the problematic cases: If there is no syntactic criterion to determine whether *n* is a mass noun (in a given context) or not, let us change the context. If the new context contains syntactic markers sufficient to classify the noun *n* as a mass noun, and *n* has the same meaning in the two contexts then *n* was a mass noun in the original context as well.

A minor problem with this solution is its circularity: The aim of the classification was that it could serve as a basis of choosing the right semantic representation, but the classification is based on the notion of identity of meanings, which depends on the semantic representation.

DRT is an available framework to solve the problem, at least partially. Sometimes a given noun cannot be classified in a sentence, because the sufficient syntactic markers are missing, but these markers classify an anaphoric expression coreferential with the given noun. This is evidently sufficient to classify the noun itself. See for example the following discourse:

(15) THE GOLD DISSOLVED IN THE LIQUID.

(16) ONE PINT OF IT WAS ENOUGH TO DISSOLVE THE WHOLE QUANTITY.

Here IT and THE LIQUID are coreferential, and the measure expression ONE PINT OF classifies IT and through this THE LIQUID mass nouns. If we have a systematic method to handle anaphoric chains, this can help to solve the qualification problem as well. (Unfortunately, there remain unclassified cases; we did only one step.)

3. SOME CHARACTERISTICS OF DRT

The first published version of DRT (that of KAMP's 1981) handles anaphora problems in the following way. First KAMP defines a limited fragment of English, and then he separates a class of nominal expressions. The members of this class are characterised by the following property: at their first occurrence these expressions introduce a new "object" into the discourse. [In HEIM's version (HEIM 1983) we collect informations about an object on a file card. The mentioned expressions open (at their first occurrence) a new file card.] KAMP adds a set of expressions (substantially variables) to the fragment of language. These are reference markers, and any member of the mentioned class of nominal expressions introduce a new reference marker into the discourse. These reference markers serve to connect anaphors to their antecedents. Each anaphoric expression selects a reference marker as its antecedent. This choice is restricted by different principles: a masculine anaphoric pronoun needs a masculine antecedent etc.

If we are interested in the linguistic status of DRT, it is not very difficult to point out that it is on the borderline of syntax and semantics. Discourse representation structures (DRSes) are basically syntactic entities, but anyhow these structures are – self-referential – partial models of the fragment of English as well. And so it is quite natural to built up a semantic representation of discourses by embedding these partial models into a usual model of our language.

In KAMP 1981 the attention was restricted to nominal anaphors, and so, the embedding system could be very simple: a fragment of a model of classical first-order languages. If we want to extend the original framework to the field of sentential anaphors, we need a more complicated embedding system. I see basically two possibilities:

(A) We can use an intensional framework (as VAN EIJCK or ASHER), or

(B) we can embed DRSes into a version of Situation Semantics (as VAN EIJCK).

I'll argue for the second possibility, but the first one has its advantages as well.

4. DRT AND SEMANTIC TYPES

(17) SUSAN MISUNDERSTOOD DRT.

As it is well known, (17) can be neither true nor false in a classical framework, since there is not a unique theory called DRT. There are different branches of DRT. Some of its branches (JAN VAN EIJCK's, NICOLAS ASHER's, or BARBARA PARTEE's for example) made efforts in order to handle (among others) sentential, temporal or VP anaphors. They all introduced different kinds of discourse referents (i.e., reference markers) for individuals, for events, for time data etc. Sometimes it was explicit and sometimes it remained in silent. But anyhow, there are widespread conventions in notation; *e* very often refers to a reference marker introduced by a sentence, whereas *a*, *u* or *v* denote reference markers referring to an individual. And *t* generally means a reference marker somehow connected to time.

The main reason why these conventions are used, I think, is that none of the authors believe that a *sober* DR-theorist would define an embedding function that maps event-type reference markers to time-streches. But after all, not all the DR theorists are always sober and one who is not may try it.

It does not cause any dangerous problem if we are interested in *truth* only. But sometimes falsity can be interesting as well.

The question of falsity arises when we extend the fragment introduced by KAMP 1981 in order to handle definite descriptions as well:

(18) THE PRESENT CHAIRMAN

From the point of view of the fragment of language, THE PRESENT CHAIRMAN and THE PRESENT KING OF FRANCE walk together, and that means that all problems of non-referring definite descriptions appear in DRT.

To handle these problems it is the best to introduce a partial embedding system. For example, embed our DRSes into an interpretation of an intensional logic with semantic value gaps. (See that of RUZSA 1986.) It introduces zero entities to "fill" semantic value gaps. Non-referring definite descriptions "refer" to the zero entity, and some built-in-mechanisms guarantee that in extensionel contexts the value gap is inherited, that is

(19) THE UNICORN THAT PEDRO OWNS IS BEATEN

proves to be neither true nor false. This partial embedding structure makes sense of type restrictions in embeddings as well. A "bad" embedding never gives truth value to a discourse, a "good" one sometimes does.

If we follow this way, naturally arises the question: how many different kind of reference markers we actually need? We can only estimate it: One for individuals, one for time streches, one for places, another one for events (or event types). And we have seen that continuous concepts (so called mass terms) need a semantic representation different from that of individual expressions. This suggests that we need another kind of reference markers. (I circumscribed this problem in a paper read at the previous Debrecen meeting on LOGIC AND LANGUAGE in 1986.) So the number of kinds of reference markers is at least five and the shadow of VP ellipsis indicates that the whole (perhaps) intensional type theory is needed, together with some additional kinds. But this is very bad news indeed, since the whole type theory arises the objections against it. (See, e.g., MENZEL 1986.) The universal properties are excluded as well as self-exemplifying properties. We can tolerate these objections if we actually need the type theory, but I think, we don't.

5. A TYPE-FREE VERSION OF DRT

The semantic framework what I am going to show is free of MENZEL's objections because it is free of types, but it can handle bad embeddings as well as non-referring definite descriptions and other sources of "truth value gaps".

I think that not all the anaphora problems can (and need) be handled in DRT, only those which are basically coreference problems. For example, VP ellipsis is different, it can be handled by copying. (See ROBERTS 1986.) Referring expressions are basically those which were mentioned above. G. LINK had shown in his well-known paper (LINK 1987) that individuals, time streches and places are semi-lattices, and I don't want to say anything new about them. But I go further and say that mass terms introduce a new semi-lattice as well as event types do. These semi-lattices are the basis of my type-free representation.

6. THE SEMI-LATTICE OF THE REFERENTS OF MASS TERMS

The most important problems connected to mass terms are the following ones:

1. How can we support semantically inferences as the next?

(20) WHAT GLISTENS IN MY GLASS IS WATER.

(21) WATER IS WET.

---

(22) WHAT GLISTENS IN MY GLASS IS WET.

2. How to handle descriptions containing mass terms?

(23) WATER IN MY GLASS

It naturally differs from other definite descriptions since if there is a quantity of water satisfying the conditions of the description, then there are infinitely many other quantities as well.

3. What is the connexion between objects and pieces of matter constituting them?

In the following I'll sketch a semantic idea which answers these questions.

- The referents of mass nouns are members of a semi-lattice.

- The partial ordering relation in the semi-lattice models the BEING OF relation.

- The unit of the semi-lattice (if there is any) can be the referent of MATTER.

- This semi-lattice can be dense (in order to model continuity).

- Inherited properties as being WET can be modeled by sub-semi-lattices.

- Other properties are to be modeled by a subset of the elements of the semi-lattice. Such non-inheriting property can be, e.g., the WEIGHS 500 TONES.

(24) THE WATER IN THE POOL WEIGHS 500 TONES.

(25) THE WATER UNDER MY SWIMMING CAP IS A PART OF THE WATER IN THE POOL.

---

(26) THE WATER UNDER MY SWIMMING CAP WEIGHS 500 TONES.

- Definite descriptions containing mass terms are similar to class-descriptions: They denote the MAXIMAL (and not the unique) element in the semi-lattice satisfying the conditions.

- In order to handle descriptions such as

(27) THE UNICORN OF PEDRO

it seems to be useful to built in a zero entity. This makes the empty piece of matter superfluous.

- The connexion between objects and pieces of matter constituting them is a *partial* function from the semi-lattice of individuals to the semi-lattice of pieces of matter. (This function must be partial in order to get a realistic picture: it is not necessary to speak about the material of the number 7 even if it is a member of the universe.)

> I first discussed these ideas in my paper 'Mass term revolt' in 1985 which simply duplicated what had already been said in G. LINK's famous paper (LINK 1983), a publication which was not known to me at the time.

7. SENTENCES IN REFERRING FUNCTION: THE SEMI-LATTICE OF EVENT TYPES

By an Austinian view of truth, sentences are referring expressions: they refer to event types; and we say that the sentence S is true in a given situation iff the given situation is the type of [S]. - The following questions are to be answered:

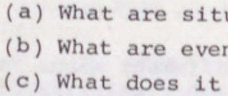
(a) What are situations?

(b) What are event types?

(c) What does it mean for a situation S being of the type e ?

Some months ago, Barwise and Cooper sketched the following idea:

If $a_1, \ldots, a_n$ are type-free entities (since "everything is a first class citizen") and r is a relation (which is characterized, among others, by its set of argument places), and $p \in 2$ then



is a fact.

Situations are characterized by the set of facts supported by them. Two situations are equal iff they support the same set of facts.

In my view, an event type is a set of events. An event type can be characterized (and modeled) by a set of facts, and is identical to the set of those situations which all support each member of the given set of facts. And S is of type e iff $S \in e$.

Let S_1 and S_2 be sentences and

$$e_1 = [\![S_1]\!], \qquad e_2 = [\![S_2]\!].$$

In order to model the referent of the conjunction of S_1 and S_2 we need a new event type: the minimal event type which is greater than e_1 and e_2. To have such an element we need a semi-lattice. The ordering in this semi-lattice is *not* the set-theoretical SUBSET OF relation. However, we cannot say that the set of event types is to be a lattice even if modeling disjunction seemingly needs it so. Only the conjunction has to be modeled by such an event type. I think that disjunction is to be modeled by a "lifted" event type: If $e_1 = [\![S_1]\!]$, $e_2 = [\![S_2]\!]$ then

$$[\![S_1 \text{ or } S_2]\!] = \{S \mid S \models \langle^{or}\ e_1\ e_2\ _1\rangle\}$$

And similarly for negation:

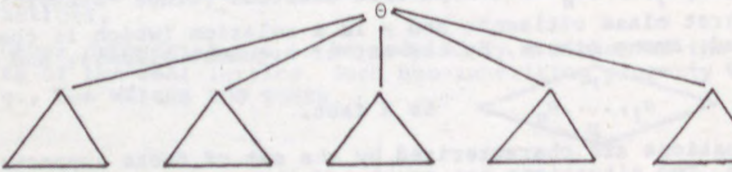
$$[\![\text{it is not the case that } S]\!] = \{S \mid S \models \langle^{not}\ [\![S]\!]\ _1\rangle\}$$

This actually corresponds to VAN EIJCK's handling of such sentences in DRT. The construction rule corresponding to "and" is the *only one* that does not introduce new DRSes, while that of "neg" and "or" lead to new embedded$_2$ DRSes.

I think there is a source of misunderstanding between Barwise and Cooper. Barwise overgeneralized the "construction rule" for conjunction, and Cooper did so, e.g., with disjunction. Neither of them is right: These two things are essentially different.

8. CONCLUSIONS

If I am right, all the referents (of different kinds of referring expressions) are members of certain - pairwise different - semi-lattices. We can get a new semi-lattice by introducing a single new entity Θ greater than any elements of any of our semilattices. This new element will serve as the zero entity.



We can handle the type restriction (if we need any) by the partiality of the fact forming operation:

$$\sigma = \left\langle \begin{array}{c} r \\ a_1, \ldots, a_n \\ 1 \end{array} \right\rangle$$

is a fact only if $a_1, \ldots, a_n$ are available arguments for the given argument places of r.

Let $a_1, \ldots, a_n$ be elements of $A \cup B \cup C \cup D \cup E \cup \{\theta\}$. If for some i $(1 \leq i \leq n)$, $a_i = \theta$, r is undefined, hence $\sigma = \theta$.

Any situation either supports a given fact or not, but θ is neither supported nor not-supported by any situation.

If x is a reference marker and f is an embedding function, then $f(x) \in A \cup B \cup C \cup D \cup E$.

REFERENCES

N. ASHER 1986: Belief sentences in DRT. *Journal of Philosophical Logic.*

H. BUNT 1978: Mass terms and ensembles. Reprinted in PELLETIER 1980.

H. BUNT 1986: *Mass Terms in Model-Theoretic Semantics.* Academic Press.

J. VAN EIJCK 1985: *Quantification in Natural Language.* Ph. D. dissertation, Groningen.

I. HEIM 1982: *The Semantics of Definite and Indefinite NP's.* Ph.D. dissertation, Amherst.

H. KAMP 1981: A theory of truth and semantic representation. In: Groenendijk et al., *Formal Methods in Study of Language.* Amsterdam.

H. KAMP 1983: Tense in text. In: Bäuerle et al., *Meaning, Use, and Interpretation.* De Gruyter.

G. LINK 1983: The logical analysis of mass terms and plurals: A lattice-theoretic approach. In: Bäuerle et al., *Meaning, Use, and Interpretation.* De Gruyter.

G. LINK 1987: Structured domains for events. In: *Proceedings of the Amsterdam Colloquium '87,* forthcoming.

C. MENZEL 1986: Completeness for a type-free second-order logic. *CSLI Report.*

T. PARSONS 1980: Mass terms and amount terms. In: PELLETIER 1980.

B. PARTEE 1984: Nominal and temporal anaphora. *Linguistics and Philosophy.*

J. PELLETIER 1975: Non-singular reference. In: PELLETIER 1980.

J. PELLETIER 1980: (ed.) *Mass Terms. Some Philosophical Problems*. D. Reidel.

L. PÓLOS 1986: Mass terms revolt. (In Hungarian.) *Tertium non datur 3.*

W.O. QUINE 1960: *Word and Object*. M.I.T. Press.

I. RUZSA 1986: Intensional logic and semantic value gaps. *Logique et Analyse 29* (114), 187-206.

DEPARTMENT OF SYMBOLIC LOGIC AND METHODOLOGY OF SCIENCE
L. EÖTVÖS UNIVERSITY, BUDAPEST, HUNGARY

VEIKKO RANTALA

INTERPRETING NARRATIVES: SOME LOGICAL QUESTIONS

It is peculiar about literary texts, narratives in particular, that a reader of a text has to move back and forth in several directions: between the text and the possible events which he is going to associate with the consecutive sentences of the text, between such events, and, finally, between the sentences themselves. Any logical reconstruction of interpretation has to explicitly consider this semantic feature, in one way or another. This feature is ultimately a consequence of the fact that an interpretation of a given sentence occurring in such a text - since the text is a sequence, or, rather, a sequence of nested sequences of sentences - is *constrained* by the linguistic meanings of other sentences of the text, and by their actual interpretations, as well. A narrative text is intensional, even though it consisted of simple extensional sentences. The meaning of the whole text is not a 'function' of the meanings of its constituent parts, considered in isolation, that is, the Principle of Compositionality does not hold in any straightforward sense.

These features are, of course, obvious and familiar to us, but it is less obvious how they should be taken into account in formal semantics. One consequence is clear, however. If we employ possible worlds semantics, which seems natural here and which has in fact been employed, in one form or another, by several literary theorists and philosophers,[1] we have to consider, intuitively speaking, courses of events, that is, sequences of events or situations, or worlds - however one wants to call them - rather than mere sets of them. Narratives describe histories of their characters, not mutually independent states of affairs. There may exist several ways to do it formally, but it is obvious that the most straightforward way is to expand familiar formal languages by introducing a concatenation operator for sentences and to consider sequences of possible worlds with respect to which such concatenations can be true or false. This kind of semantics is, of course, immediately suggested by the very nature of narratives, and it has been employed, more or less informally, by some semioticians,[2] but, as far as I know, no straightforward extension of 'ordinary' formal propositional or predicate logic exists so far.[3]

In this paper, I shall define basic notions of such an extension; it will be done in Section 2, below. My aims in this essay are rather restricted, however. For instance, I shall simplify discussion by only considering (modal) propositional logic, that is, a logical surface structure of narratives and their reception. Furthermore, no really important results about logic or interpretation will be derived here; it is not even known at present whether in this logic or in its quantificational extensions such results are obtainable at all. On the other hand, the kind of logical treatment discussed in this paper may suggest some new ways to approach literary criticism.

In Section 1, I attempt to find out some basic features of the logic to be defined in Section 2 by considering constraints on which actual readers are dependent when they approach narrative texts. Section 1 is thus supposed to motivate the formal definitions of Section 2.

1. CONSTRAINTS

In order to gain preliminary insight into the contextual features of the interpretation of narrative texts, in particular, into the role of constraints, let us consider the following short passage from 'Death and the Compass' by Jorge Borges:[4]

> The train stopped at a silent loading station. Lönnrot got off. It was one of those deserted afternoons that seem like dawns. The air of the turbid, puddled plain was damp and cold. Lönnrot began walking along the countryside.

Since the passage is a part of a longer text, it should not be detached from the whole story. We may, however, consider certain aspects of the passage, in particular, mutual relations of its sentences and the question of how consecutive events, 'possible worlds' should be chosen in order to interpret the passage.

Let the five sentences of the passage be denoted by the letters *p,q,r,s,t,* respectively, and the passage by the concatenation *pqrst*. In what ways they constrain each other is rather evident. For instance, how *q* is to be interpreted is constrained by the fact that it comes after *p*: Lönnrot got off the specific train that stopped at a silent loading station. Similarly *r* and *s* state some restrictions for the interpretation of *t*, i.e., conditions under which Lönnrot started walking along the countryside; *p* affects the interpretation of *s* since it implies that the air around the loading station was damp and cold, etc.

But constraints may also be directed from right to left, that is, a sentence may constrain sentences which are preceding it. Thus *q* implies that Lönnrot was in the train that stopped at a silent loading station,[5] *r* and *s* state the same restrictions for the interpretations of *p* and *q* as for that of *t*, and so forth.

Constraints of the kind discussed above, i.e., constraints which are due to the fact that when interpreting sentences of a text we have to consider its other sentences, will be called *textual constraints*. If we think that, ideally, each sentence of the text has a *general* linguistic meaning, that is, meaning which it has when it is considered in isolation - when it is considered as a type rather than as a token in the text - then the presence of other sentences around it makes its meaning more specific by cutting off possibilities which its general meaning admits of.

Let us illustrate the role which 'possible worlds' play in the forthcoming formal semantics by thinking of them as imaginary events or situations, or small worlds, which a reader is 'picking' or 'choosing' one after another when interpreting the consecutive sentences of a narrative. We may even think of them as mental images or mental scenes of the reader. Though our actual images may not be very definite, we can use the notion of mental image as a heuristic device which is relatively free from ontological problems and which, on the other hand, accords with our disposition to visualize the events a story is describing. When using the metaphor of picking or choosing events, I am ignoring the obvious fact that one usually reads a text too fast to be able to consider sentences, thus to visualize events they describe, one by one. Let us assume, however, that this can be done - for example, by reading the text very slowly.

Now, when a reader picks a world *v* which he thinks to correspond to the sentence *q* occurring in the text *pqrst*, that is, at which it would be true, then, according to what we said above about textual constraints concerning *q*, *v* must be such that in it Lönnrot got off the very same train that belongs to the situation, say *u*, which the reader chose in order to interpret the sentence *p*. In *v*, on the other hand, there must be a deserted afternoon that seems like a dawn and the air must be damp and cold. If the reader picks the world v for q so that this is not the case, then it does not provide a correct interpretation for the occurrence of q in the text *pqrst*. Therefore, after he has read r (and s), he has to go back, as it were, and choose a more appropriate world for q, that is, a world which satisfies the conditions stated by r (and s).

It is clear that such back-and-forth adjustments concerning interpretations of individual sentences are actually done, and can be done, only locally and that not all adjustments are important. Incoherent and incomplete, and wrong moves from one image to another are often made. This is why incorrect interpretations are possible. But an 'ideal reader' would always proceed in the manner described above, and not only locally - and an ordinary reader may gradually approach it by reading a text over and over again.[6]

Our heuristic discussion about textual constraints seems to confirm what was said earlier about the failure of the Principle of Composionality. This feature must be taken into account in the formal semantics of narratives. Therefore, I shall assume, at least tentatively, that valuation functions are initially defined on sequences of sentences. The value of a sequence is not, in general, a function of the values of its component parts; how the latter are related to the former depends e.g. on methodological situations and on the nature of the texts to be reconstructed.

Textual constraints derive from what is explicitly said in the text. Common knowledge and beliefs and social rules concerning language are, of course, influential here. A reader of *pqrst* is not able to conclude that Lönnrot got off the train that stopped at a silent loading station if, he does not understand how the language is generally used in texts or if he does not know the function of trains. So, textual constraints, as I want to define them, have to do with readers' abilities at some common and general social and linguistic levels. The forthcoming formal semantics is based on the idealizing assumption that a text possesses a conventional meaning at a general linguistic level; a valuation function will formally correspond to such a meaning. However, such a meaning may not be unique, and some actual texts may not even possess independent meaning at all. Consider, for example, ambiguous sentences, metaphors, complicated and theoretical expressions, or works of art. Such an expression need not even mean the same thing for two readers belonging to the same linguistic community and being equally competent language users. Furthermore, two readers may have different aesthetic and cognitive views, abilities and skills, different experiences in literary criticism, or different intentions concerning how the text should be read. Even the same reader may approach a given text differently on distinct 'literary occasions'.

Therefore, it seems appropriate to say that there also exist more personal constraints on which interpretations depend. These are constraints for a reader which pertain to more personal levels of the reader's background and which are less dependent on any specific text. They may, of course, provide different frameworks for different texts - one's background may be differently oriented for different purposes, intentionally or nonintentionally - but they are not primarily due to the syntactic structure of any given text. I shall call them *pragmatic constraints*, for obvious reasons. They may also be called *structural constraints*, since they are independent of semantic relations.

There is a countless number of personal factors, some of which are even contingent, and principles which have effect on *local* features of an interpretation. Such local features are not always cognitively or aesthetically significant, as, for instance, details of the station at which Lönnrot got off the train. On the other hand, theoretical pragmatic constraints - various theoretical principles and attitudes of the reader - are often *global* covering the whole interpretation. For example, if a reader assumes that environmental conditions have causal effects on how people feel, or if he assumes that the author whose text he is reading thinks so, then he cannot think, in order to be coherent in his interpretation, that Lönnrot felt very happy as he got off the train and started walking along the countryside. In terms of possible worlds, this would mean that since Lönnrot did not feel very happy in the world *x* in which the air of the turbid, puddled plain was damp and cold and which was chosen by the reader to interpret the sentence *s* then for the sentences *q* and *t* he has to choose worlds in which Lönnrot was not very happy, either. The same conclusion follows, of course, if the reader does not assume any general causal connection between environment and mood but only that e.g. in 'Death and the Compass' there is a similar connection between environment and Lönnrot's mood.

From these examples we can see, furthermore, that pragmatic constraints can be influential backwards as well as forwards, in like manner as textual constraints.

When a reader is picking new events while reading, he has to do it so that it is in conformity with the pragmatic constraints in the first place. In many cases it may not even be possible for him to choose otherwise; for instance, he cannot read the text so that it is beyond his abilities or knowledge, even though the general linguistic meaning of the text left room for such readings. Pragmatic and textual constraints are, of course, interwoven in actual interpretations, but in a pragmatic sense the former precede the latter, though logically and linguistically the latter are primary. Roughly speaking, the reader *has* to proceed so that relevant pragmatic constraints will be satisfied, but he *attempts* to do it so that it corresponds to the text. Pragmatic constraints indicate how he is permitted to go on when picking new worlds for consecutive sentences of the text. As we saw above, what he is permitted to do at each step may depend on his earlier choices, and sometimes on his choices to come. This feature in fact suggests that my *urn model* semantics is applicable, in some form or another, to the semantics of narratives. On the other hand, it also suggests that urn model semantics could be applied by using the notions of *semantical game*, in Hintikka's sense, and *interpretative strategy*.[7]

In a sense, then, pragmatic constraints indicate the worlds which are *accessible*, independently of any specific text, at each stage of any interpretation which assumes these constraints. It follows that they may also have a predictive function in the sense which I shall discuss now. Consider again the text *pqrst*. Having reached a world *x* for the sentence *s* such that in *x* Lönnrot is at the loading station in the middle of the damp and cold air of the turbid, puddled plain, a reader may have *expectations* concerning the next step of the story. These expectations are dependent on his interpretative assumptions and on the worlds he chose before *x*. Thus, for instance, on the assumption concerning the relation of environmental conditions and Lönnrot's mood the reader certainly expects that Lönnrot will not feel very happy at the next stage of the story, if he is still around. In terms of possible worlds, this amounts to the fact that *all* the worlds which are accessible from *x* are such that in them Lönnrot does not feel very happy, if he still is somewhere in the neighborhood of the loading station. On the other hand, it may not be expected by the reader that Lönnrot began walking along the countryside.

Let us assume now, that later in the text the author indicates - this he does not actually do, however - that Lõnnrot felt happy, after all, as he started walking along the countryside. Then the reader's expectation will be falsified.[8] Then, on our assumptions concerning *x*, the reader cannot proceed from *x* consistently with the text, whence - if we assume ideal reading - he has to return to some earlier world *w* for the sentence *r* in order to reach another world *x'* for *s* which is accessible from *w* and such that there exists a world for *t* which is accessible from *x'* and in which Lõnnrot felt happy.[9] If such an *x'* cannot be found - such would be the case, for instance, if the reader's assumption concerning the relation of environmental conditions and mood is global and thus intended to cover the whole of the interpretation - pragmatic constraints are incompatible with the general linguistic meaning of the text, that is, with textual constraints, and must be changed.

Another modality which is suggested by the notion of pragmatic constraint is, of course, possibility. As we have seen, expectations amount to a kind of necessities for readers which may emerge at some stages of reading. Hence, possibility and expectation, related to the same pragmatic constraints, are dual notions. Reading a text in the framework which is provided by an appropriate system of pragmatic constraints may give rise to expectations concerning forthcoming events and, on the other hand, indicate possibilities the reader can speculate about. Both notions have cognitive, aesthetic, and conceptual connotations, since the notion of pragmatic constraint does, but, as it seems to me, it is the *relation* between what is predictable and what is possible in the story which is of greater cognitive and aesthetic significance than these modalities considered separately. If too much can be successfully predicted about the story, the text is conceptually narrow in the sense that it does not leave very much room for the reader to approach it creatively. If, on the other hand, the text admits too many possibilities, it is conceptually loose, lacking cognitive and emotive intensity. Only if these dual features are in harmony - whatever that means, exactly - the story can exhibit cognitive and aesthetic tension.

2. FORMAL SEMANTICS

In this section, I shall formally define what it means to say that a narrative text is considered true on a given interpretation and that an expectation prompted by a text is satisfied. As I have suggested above, at least the following notions are needed in such a semantics: (i) a space of events - i.e. a set of 'small' worlds or situations; (ii) courses of events - sequences of events from the space of events; (iii) pragmatic constraints - indicating which courses of events are considered 'admissible' for a reader; (iv) textual constraints - indicating the general or conventional linguistic meanings of sentences and their concatenations.

Let us consider (i)-(iv) formally. This will be done only very briefly. The definitions to follow are motivated by our considerations in the previous section. Let U be a nonempty set and let

(1) $C_n \subseteq U^n \quad (n = 1,2,...)$;
$\mathbf{C} = \langle C_1, C_2, ... \rangle$.

The set U represents a space of events and the infinite (countable) sequence **C** the pragmatic constraints effective on a given literary occasion. **C** represents the possible ways in which a reader who is committed to the constraints in questions can choose consecutive small worlds. If U and **C** are as indicated, the pair

(2) $\mathbf{F} = \langle U, \mathbf{C} \rangle$

will be called a *frame*, and it represents the literary occasion in question; it is a framework relative to which texts can be interpreted.

Any finite sequence $\mathbf{u} = \langle u_1, ..., u_k \rangle$ from U^k, for any k, is called a *course of events*, and it is called *admissible* relative to **F** if it belongs to C_k. In what follows, we shall more briefly write

(3) $\mathbf{u} = u_1 ... u_k$.

Since a narrative text is composed of sequences of sentences, describing courses of events, we need in our formal language L a concatenation operator in addition to ordinary operators. As noted before, I restrict discussion to propositional logic. Textual constraints will be represented by a valuation function on the concatenations of atomic formulas and by reducing the truth conditions for arbitrary concatenations to this function.

The *language L* contains propositional variables and a propositional constant T - which together are the atomic formulas - connectives $\wedge, \neg, \circ$.

and a modal operator E. The ordinary notion of formula is extended by adding formulas of the forms $\circ(\theta_1,...,\theta_n)$ (for all n = 1,2,...) and Eθ. The formula $\circ(\theta_1)$ can be identified with θ_1. It is assumed here, for simplicity, that each θ_i and θ are formulas not containing the operator $\circ$ and that only the last formula θ_n can be of the form Eθ. These assumptions are only made for expositional simplicity and can be dispensed with. The former formula is called a *concatenation*, and it will be written more briefly as '$\theta_1...\theta_n$'. Its *length* is n and the θ_i (i=1,...,n) are its *components*. The notion of length is generalized for an arbitrary formula inductively in the obvious way.

The latter formula represents *expectation;* and a concatenation $\theta_1...\theta_n$Eθ can be read as 'it is expected after $\theta_1...\theta_n$ that θ'. The notion of *possibility* is defined as usual:

$$P\theta =_{def} \neg E \neg \theta$$

and read: 'it is possible after $\theta_1...\theta_n$ that θ'.

Given a frame **F**, as in (2), a *valuation* V *on* **F** is a function assigning a subset of U^n to each concatenation $p_1...p_n$ of atomic formulas:

(4) $V(p_1...p_n) \subseteq U^n$ (n =1,2,...);

and

(5) **M** = ‹**F**, V›

is a *model* for L. If **u**, as in (3), is a course of events, then

(6) **I** = ‹**M**,**u**›

is an *interpretation* for L. If **u** is admissible relative to **F**, then **I** is an *admissible interpretation* for L.

The valuation V represents the textual constraints for the language, and the condition (4) reflects the fact, discussed in the previous section, that they have to do with the general linguistic meanings of texts.

As observed above, it is natural to assume that sentences occurring in a text constrain the meaning of another sentence of the text by restricting the possibilities the latter would individually admit of. Propositional constant T, which we only introduced for certain technical purposes, plays a role of logical truth, whence we assume that its presence in a concatenation (text) is not restrictive in this way and, on the other hand, that its meaning is not restricted by others. To express these requirements formally, we need some notation. Consider a concatenation $\sigma = p_1...p_n$ where $p_1,...,p_n$ are atomic formulas. Define:

(7) $V_i(\sigma) = \{ u\colon \exists u_1 \ldots \exists u_{i-1} \exists u_{i+1} \ldots \exists u_n (u_1 \ldots u_{i-1} u u_{i+1} \ldots u_n \in V(\sigma))\}$,

($i = 1,\ldots,n$, where it is, of course, assumed that $u_1 \ldots u_{i-1}$ is empty when $i=1$ and that $u_{i+1} \ldots u_n$ is empty when $i=n$; and similarly below);

(8) $\sigma(-i) = p_1 \ldots p_{i-1} p_{i+1} \ldots p_n$.

Consider now any component p_i of σ ($i=1,\ldots,n$). The valuation V is assumed to satisfy the following conditions; they are our formal representations of the informal conditions just mentioned:

(9) $V_i(\sigma) \subseteq V(p_i)$;

$V_i(\sigma) = U$ if $p_i = \top$.

(10) If $p_i = \top$, then

$V_k(\sigma(-i)) = V_k(\sigma)$, when $k=1,\ldots,i-1$,

$V_k(\sigma(-i)) = V_{k+1}(\sigma)$, when $k=i,\ldots,n-1$.

As we observed in Section 1, textual constraints must be relativized to pragmatic ones: when a reader is interpreting a text, he commits himself to pragmatic constraints. This is why **I** is called an admissible interpretation only if the course of events **u** is admissible. However, general truth conditions will be stated relative to any course of events. Given any interpretation (6) for L, we shall denote '$\mathbf{I} \models \varphi$' or '$\mathbf{u} \models \varphi$', if φ is *true* at **u** in **M**. There hardly exists any unique intuition concerning truth conditions for interpretation - different kinds of text may require different conditions - but the definition (11) below is a straightforward generalization of the usual (classical) ones.

The following conventions will be used. If $\mathbf{u} = u_1 \ldots u_k$ and $\mathbf{v} = v_1 \ldots v_m$, then $\mathbf{uv} = u_1 \ldots u_k v_1 \ldots v_m$; if **u** is the empty sequence, then $\mathbf{uv} = \mathbf{v}$. Furthermore, $\mathbf{u}|n = u_1 \ldots u_n$ if $1 \leq n < k$; $\mathbf{u}|0$ is empty; and $\mathbf{u}|n = \mathbf{u}$ if $n \geq k$. If $\sigma = \theta_1 \ldots \theta_n$ and θ is a formula not containing the concatenation operator, then we write '$\sigma\theta$' for '$\theta_1 \ldots \theta_n \theta$'. Furthermore, a given occurrence of a formula θ as a component of a concatenation σ will be displayed by simply writing '$\sigma(\theta)$' if no confusion will arise. If this occurrence is replaced by a similar formula δ, we shall denote '$\sigma(\delta)$'

In what follows, it is assumed that $p_1,\ldots,p_n$ are atomic formulas; θ,δ formulas not containing the concatenation operator; φ,ψ any formulas; σ a concatenation of length n (for any $n = 1,2,\ldots$); and σ' as σ or, alternatively, nothing at all, in which case $n = 0$. Let **I** be as in (6), above, whence **u** is of length k. The truth conditions are the following:

(11) (i) $\mathbf{u} \models p_1 \ldots p_n$ iff $\mathbf{u}|n \in V(p_1 \ldots p_n)$;

(ii) $\mathbf{u} \models \neg\varphi$ iff $\mathbf{u} \not\models \varphi$;

(iii) $\mathbf{u} \models \varphi \wedge \psi$ iff $\mathbf{u} \models \varphi$ and $\mathbf{u} \models \psi$;

(iv) $\mathbf{u} \models \sigma(\neg\theta)$ iff $\mathbf{u} \not\models \sigma(\theta)$ and $\mathbf{u} \models \sigma(\mathsf{T})$;

(v) $\mathbf{u} \models \sigma(\theta \wedge \delta)$ iff $\mathbf{u} \models \sigma(\theta)$ and $\mathbf{u} \models \sigma(\delta)$;

(vi) $\mathbf{u} \models \sigma'\mathsf{E}\theta$ iff for all $\mathbf{v} \in U^{\max(n-k,0)+1}$ such that $(\mathbf{u}|n)\mathbf{v} \in C_{n+1}$, $(\mathbf{u}|n)\mathbf{v} \models \sigma'\theta$.

The semantic properties of T which were indicated in (10), above, explain the intuitive meaning of the condition (iv). It seems appropriate to define validity (relative to appropriate conditions for **C** and V) as follows. A formula of L is *logically valid* if it is true at every course of events which is of the same length as the formula, and *pragmatically valid* if it is true at every such admissible course of events, in every model for L. Valid formulas or axiomatizations will not be studied here, however, but it is clear that if relevant, additional conditions are stated for **C** and V, different sets of conditions may yield different valid formulas (both concatenations and formulas containing E) in both senses of validity.[10]

When interpreting a text, the reader is not only constructing a *single* course of events, or a single world, for the text. In fact, if interpreting, in a proper sense of the word, is to determine the *meaning* of the text, then to call a structure of the form (6) - which only corresponds to a single act of reading - 'interpretation' is to speak about interpretation in some narrow sense. If we apply and modify the well-known heuristic idea of current philosophy according to which (part of) the meaning of a sentence can be characterized as the *proposition* it determines, i.e., as the collection of possible worlds at which it is true, we may tentatively say that (part of) the *interpretation* (in a given model **M**) of a 'text' σ of length n consists of the set of all admissible courses of events of this length at which σ is true, that is, $V(\sigma) \cap C_n$, where $V(\sigma) = \{\mathbf{u} \in U^n : \mathbf{u} \models \sigma\}$. In heuristic terms, it consists of the proposition which is determined by the text in the framework which the adopted textual and pragmatic constraints provide.

NOTES

[1] See, e.g., Eco [2], Lewis [5], Maitre [6], Pavel [7], Petőfi [8].
[2] See, e.g., Eco [2], Petőfi [8], Ryan [9].

[3] A very preliminary and tentative sketch is contained in my [12]. A similar approach can, of course, be applied to the semantics of musical texts, as well; see Kurkela [4], Rantala [11].

[4] In Borges [1], pp. 76-87.

[5] This is also implied by the text preceding *p*, but we ignore it here.

[6] Hence an ideal reader could be considered as a 'hermeneutic limit' of an ordinary reader.

[7] See [10] for urn models; and Hintikka [3] and Saarinen [13] for game- theoretical semantics and the notion of strategy. Since it is not quite clear to me at the moment what advantages game-theoretical semantics would yield in this connection and how it would exactly look here, I shall just give a simple set-theoretical construal of this feature. But the general pattern of urn models will be used.

[8] There are of course many other ways to falsify expectations, but their consequences are similar.

[9] So, here we have another example of the back-and-forth character of reading.

[10] As I have already remarked, it is not known to me at present whether anything interesting in a purely logical sense can arise in this framework, but it seems to provide us with some new logical insights concerning interpretation.

REFERENCES

[1] J.Borges, *Labyrinths*, A New Directions Book, New York, 1962.

[2] U.Eco, *The Role of the Reader: Explorations in the Semiotics of Texts*. Indiana University Press, Bloomington, 1979.

[3] J.Hintikka, 'Quantifiers in Logic and Quantifiers in Natural Languages', in S.Korner (ed.), *Philosophy of Logic*, Basil Blackwell, Oxford, 1976, pp. 208-232; reprinted in [13], pp. 27-47.

[4] K.Kurkela, *Note and Tone. A Semantic Analysis of Conventional Music Notation*, Acta Musicologica Fennica 15, Helsinki, 1986.

[5] D.Lewis, 'Truth in Fiction', *American Philosophical Quarterly* 15 (1978), pp. 37-46.

[6] D.Maitre, *Literature and Possible Worlds*, Middlesex Polytechnic Press, London, 1983.

[7] T.G.Pavel, '"Possible Worlds" in Literary Semantics', *The Journal of Aesthetics and Art Criticism* 34 (1975), pp. 165-176.

[8] J.Petőfi, 'Meaning, Text Interpretation, Pragmatic-Semantic Text Classes', *Poetics* 11 (1982), pp. 453-491.

[9] M-L.Ryan, 'On the Window Structure of Narrative Discourse', *Semiotica* 64 (1987), pp. 59-81.

[10] V.Rantala, 'Urn Models', *Journal of Philosophical Logic* 4 (1975), pp. 455-474; reprinted in [13], pp. 347-366.

[11] V.Rantala, 'Zenemű es lehetséges események', *Tertiur non datur* 3 (1986), pp. 99-111; reprinted as 'Musical Work and Possible Events', in V.Rantala, L.Rowell, and E.Tarasti (eds.), *Essays in the Philosophy of Music*, Acta Philosophica Fennica, Helsinki, 1987.

[12] V.Rantala, 'Possible Courses of Events', in M.Kusch and H.Schröder (eds.), *Text, Interpretation, Argumentation*, Buske, Hamburg, 1988.

[13] E.Saarinen (ed.), *Game-Theoretical Semantics*, D.Reidel, Dordrecht, 1979.

UNIVERSITY OF TAMPERE, FINLAND

CLAUDIA CASADIO

EXTENDING CATEGORIAL GRAMMAR

(AN ANALYSIS OF WORD ORDER AND CLITICIZATION IN ITALIAN)

0. Summary

The aim of the present work is to apply an extended version of a categorial grammar to a set of Italian sentential patterns involving dislocation, subject-inversion and a variety of cliticization processes. The basic model is the algebraic grammar defined in Lambek (1958) and reproposed in Lambek (1985), in which directional category symbols and a wide set of categorial rules are employed. The characterization, within this framework, of complex category symbols is explored and an extension of the set of categorial rules is considered, in connection with the occurrence of "displaced" constituents.

1. Introduction

Categorial Grammar represents an interesting model for the analysis of natural language, both for its elegance and simplicity, and for the straightforward relation it establishes between a *syntactic structure*, simply consisting of the morpho-phonological representation, and a *semantics*, which may be directly introduced in a *compositional* format. Classical categorial grammars, in the Ajdukiewicz/Bar-Hillel tradition, are special versions of the (CF)PSG model, where the information involved in the PS-rules is directly expressed by the relations between basic and functorial category-symbols listed in the dictionary. Applying to adjacent constituents only, these models present the same inadequacies as IC systems in the analysis of many relevant linguistic facts such as long distance dependencies, generalized conjunction, elliptical occurrences, etc. (see the equivalence results of CGs and PSGs in Bar-Hillel, Gaifman and Shamir 1960 and the discussion in Buszkowski 1987).

Two kinds of extensions have recently been put forward: (a) the definition of the category symbols as complex entities consisting, at least, of a set of features and of level indices (the framework of the X-bar theory is of relevance here; see e.g. Bach 1983); (b) the implementation of the set of categorial rules that in the standard model consists of the sole rule of *functional application* of a functor category to its argument(s)(possibly specified with respect to direction; see Bar-Hillel 1960). A set of combination rules has been proposed in Ades & Steedman(1982), with extensions e.g. in Steedman (1985), allowing the analysis of a wide range of unbounded dependencies and conjunction constructions. However, a *flexible* model of categorial grammar was already available in Bar-Hillel's times, defining a set of categorial rules within an elegant algebraic theory. This is the *Calculus of Syntactic Types* proposed by Joachim Lambek.

2. Lambek's categorial grammar

In his 1985 paper, Lambek gives the following definition of a categorial grammar:

"A *categorial grammar* of a language may be viewed as consisting of the syntactic calculus freely generated from a finite set {S,N,....} of basic types together with a dictionary which assigns to each word of the language a finite set of types composed from the basic types and I by the three binary operations." (1985, p.10)

The *syntactic calculus* is a deductive system L defining a set of types (or categories) closed under the three binary operations occurring in (1)(to be read as: A times B, C over B, A under C, where M is the associated multiplicative system) and including identity[1]:

(1) a. $A.B = \{x.y \in M \mid x \in A \wedge y \in B\}$ (for every $A,B,C \subseteq M$)
b. $C/B = \{x \in M \mid \forall y \in B,\ x.y \in C\}$
c. $A\backslash C = \{y \in M \mid \forall x \in A,\ x.y \in C\}$

In L theorems such as those occurring in (4) are derivable on the basis of the axioms in (3) and of the inference rules in (2) (for every $A,B,C \subseteq M$):

(2) a. A.B -> C iff A -> C/B
b. A.B -> C iff B -> A\C
c. if A->B and B->C then A->C

(3) a. X -> X
b. (A.B).C = A.(B.C)

(4) a. (A/B).B -> A
b. A.(A\B) -> B
c. B -> (A/B)\A or A/(B\A)
d. (A\B)/C <-> A\(B/C)
e. (A/B)/C <-> A/(C.B)
f. (A/B).(B/C) -> (A/C)
g. A/B -> (A/C)/(B/C)

where (3a) is an axiom scheme, for X of any category, and (3b) is the associative law, which plays a crucial role with respect to the generative capacity of the system. In fact, the multiplicative system M corresponding to the syntactic calculus is not necessarily associative (i.e. a semigroup); if the associative law is assumed, a more powerful system results, allowing unbracketed strings and characterizing a strong notion of constituency (this system is analyzed in Lambek 1958); if associativity is dropped, then only bracketed strings are allowed corresponding to standard grammatical constituents, but the types must increase in number and complexity. For the purpose of the present analysis the associative law will be assumed to hold since, on this basis, the combination rules given by theorems (4d)-(4g) are admitted. The resulting system allows the generation of several constituents which may look strange or unnatural, at a first glance, but of which the semantic status is fully justified, if we consider them as the result of processes of lambda abstraction[2].

The theorems (4a)-(4b) correspond to the well-known *cancellation* rules of the Ajdukiewicz/Bar-Hillel grammar, but further operations are possible such as that in (4f) which allows two adjacent functorial categories to combine. The content of this rule, known as *functional composition* on the analogy of the mathematical operation, is that a function from A to B combines with a function from B to C, to give a function from A to C. The theorems (4c) and (4g) represent rules to expand types into more complex ones, and the *interchange* rule (4e) is derivable on the basis of (4d), which states that, according to associativity, parentheses may be omitted (the same holds for (4f) and (4g))[3].

We may think of the syntactic calculus as a universal system of rules in which a language-specific dictionary is embedded. For instance, taking the set B of *basic categories* given in (5a), the derivations of complex types for the English expressions in (5b)-(5d) are obtainable on the basis of the inference rules given in (2):

(5) a. B = {N, NP, S}

b. if *Mary* -> NP and *Mary works* -> S,
then *works* -> NP\S [2b];

c. if *apple* -> N and *an apple* -> NP,
then *an* -> NP/N (Det) [2a];

d. if *Mary ate an apple* -> S, *Mary* -> NP,
ate an apple -> NP\S, where NP\S = VP,
and *an apple* -> NP, then *ate* -> VP/NP [2a]

The syntactic calculus will assign the type S (sentence) to a given string of words if and only if the dictionary assigns a type B_i to each word and the derivation $B_1, \ldots, B_n$ -> S is a theorem of the syntactic calculus (i.e. the combination of the assigned n types ends up in the single type S, on the basis of the axioms and the rules of the system).

3. Categories as complex symbols

The first generalization we are going to consider concerns the definition of the category symbols. In a Lambek grammar, such as in an Ajdukiewicz/Bar-Hillel grammar, categories (or types) are simple unities without internal structure. However, the developments of the generative analysis have shown the advantages of defining complex category symbols specified with respect to both sets of features and levels of expansions. This is pointed out by Emmon Bach[4] when he says that the internal structure of categories must pay attention not only to the relative ordering of their arguments, but also to such matters as case, number, person, i.e. to that set of features which are relevant to determine lexical, phonological, semantical, etc. properties.

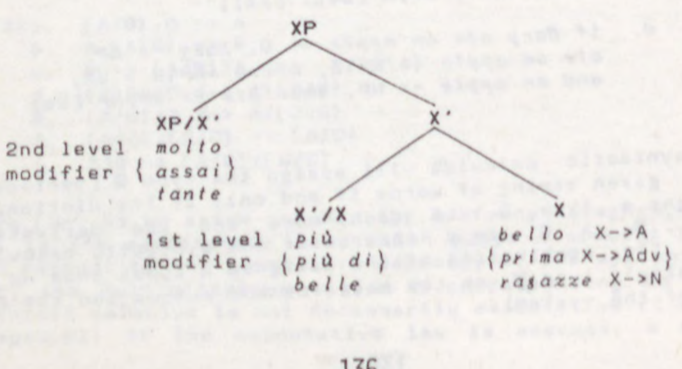
In this perspective, we will assume category symbols specified with respect to: (i) a lexical head, (ii) appropriate sets of features, (iii) the levels in which the category expands (here, two levels will be sufficient for the relatively simple contexts considered). Compare the representations given in (6) which refer to the formal theory of features introduced in Gazdar et al.(1985):

(6) X = [head:W, bar:i] W = [+V, +N]
NP = [head:-V,+N, bar:2, gen:ms, num:sg, case:ACC]
VP = [head:+V,-N, bar:2, per:3, VForm:fin,...]

The generalized category symbol X is specified with respect to a *head*, consisting of a set of features W, including at least the standard X-bar specifications, and a level index(bar) whose range is: $0 \leq i \leq 2$. Allowing more feature specifications, we obtain complex category symbols such as NP, masculine, singular, accusative, or VP, 3rd person, finite form[5]. The use of variable symbols X allows the statement of interesting generalizations among categories or rules, such as the familiar representation in Figure 1, under the convention on indices in (7):

(7) X^{i+1}/X^{i} $X^{i} \rightarrow X^{i+1}$ $(i=0, X^{i+1}=X'; i=1, X^{i+1}=XP)$

- Figure 1 -



The sole innovation here is the functorial notation of the category symbols which induces a direct mapping of the head on its first and second level modifiers. These modifiers take two basic general forms: X'/X, XP/X', to which a variety of right-oriented endocentric modifiers may be assigned (e.g. A'/A, NP/N', N'/N, etc., cf. Bach 1983). Parallel generalized symbols may be introduced for left-oriented modifiers. Moreover, the head itself may be specified with respect to a set of subcategorized items, its complement. This directly follows from the functorial notation, depending on the dictionary-assignments[8]. For example, under the feature specifications of the head given in Figure 1, the structure automatically produces the analysis of the strings in (8) (for simplicity, further feature specifications of the modifier system are omitted):

(8) a. molto più bello
much more beautiful

b. assai più di prima
much more than before

c. tante belle ragazze
many beautiful girls

4. The set of categorial rules

The second generalization that will be explored concerns the set of categorial rules. An illustration of the way the L system works, with respect to a set of Italian simple clauses, is given in the examples below.

4.1 Rules of functional application

The basic theorems (4a) and (4b), directly derived by the axiom (3a) and the inference rules (2a) and (2b), are applied in (9) and (10) producing the constituent analyses in (9b), (10b). The derivation is presented in the format of Ades and Steedman(1982), where a shortened name of the rule which is employed is written on the right side of each horizontal line:

(9) a. Maria mette una tovaglia nuova sul tavolo.
Mary is putting a new table-cloth on the table.

```
b.  mette      una tovaglia nuova     sul tavolo
    VP/PP/NP          NP                  PP
    ____________________________A(r)
            VP/PP
            ______________________________________A(r)
                            VP
```

(10) a. Gianni corre velocemente sotto la pioggia.
Gianni is running quickly under the rain.

```
b.  corre   velocemente   sotto la pioggia
    VP      VP\VP         VP\VP
    ______________A(l)
          VP
          _______________________A(l)
                     VP
```

In both cases, a rule of functional application is employed, combining a functor with an argument on the right (9b) or on the left (10b). Since left and right sides are relevant within Italian word order, the two rules will be defined as *Application* rules specified with respect to the direction of application, i.e. *right* (r) or *left* (l); see (11) below:

(11) a. X/Y Y -> X **right application**: A(r)

b. X X\Y -> Y **left application**: A(l)

4.2 Rules of type expansion

An important set of rules is represented by theorems (4c) and (4g), which are rules to expand types into more complex ones and are known as *raising* or *lifting* rules. The effect of their application is that an argument-type becomes a functor-type able to apply to what previously was the functor. See the examples in (12)-(14):

(12) a. Maria telefona domani.
Mary (will) telephone tomorrow.

```
b.  Maria      telefona
    ___NP___L  NP\S        VP = NP\S
    S/(NP\S)
    ___________________A(r)
             S
```

Within the word order that we assume as basic, the subject NP is the left-argument of the predicate VP which, through the directional assignment NP\S, is allowed to combine with it by functional application, *via* A(l). On the other hand, the quantificational reading of the subject NP may be obtained by raising the subject type by means of the rule corresponding to theorem (4c): this produces the analysis in (12b), where the predicate is the argument of the lifted subject[7].

The example in (14) concerns the derivation of expanded types for categories with indices X/X. Here the category of those adverbials is considered which, in Italian, may occur both at the leftmost side of the sentence and within the VP. We may assume to assign the basic type of sentential modifiers to such words, as shown in (13):

(13) a. Certamente Maria telefonerà.
Certainly Mary will telephone.

b.
```
Certamente   Maria   telefonerà
   S/S        NP       NP\S
             ____________________A(l)
                       S
  ________________________A(r)
           S
```

To combine with a VP, as it happens in (14b), the adverbial type has to be expanded by the rule given in theorem (4g), corresponding to (14c):

(14) a. Maria certamente telefonerà.
Mary certainly will telephone.

b.
```
Maria   certamente      telefonerà
 NP     ____S/S____D      NP\S
        (NP\S)/(NP\S)
        __________________________A(r)
                 NP\S
  _______________________A(l)
             S
```

c. S/S -> (NP\S)/(NP\S) = VP/VP

The definition of the two expansion rule schemas occurs in (15) again with the associated shorthands referring to the more standard usages of these rules:

(15) a. X -> (X/Y)\X or X/(Y\X) **type lifting: L**

b. X/Y -> (X/Z)/(Y/Z) **division: D**
(*ibid.* with \)

4.3 Rules of functional composition
A set of rules of particular interest, which has been deeply investigated in Ades and Steedman(1982) and subsequent literature, is that of *composition* rules. Compare (16a):

(16) a.
```
corre   velocemente   sotto_la_pioggia
 VP       VP\VP           VP\VP
         ________________________C
                 VP\VP
```

where a new constituent is made up of two constituents both assigned to a functorial category. In this case, we have an alternative analysis of (10b) in which a complex modifier VP\VP is created to combine with the head VP in one step[8]. This kind of rule may be employed in several different situations; see e.g. (17):

```
(17)    su___ il___ tavolo      su+il = sul (on the)
        PP/NP NP/N'   N'
        __________C
           PP/N'
```

All the compounds preposition-plus-article (articulate prepositions), which are typical components of the Italian determiner system, may be easily analysed in this way, as the result of the composition of the respective functions. The rule schema of *composition* is given in (18):

(18) X/Y Y/Z -> X/Z **functional composition: C**
(ibid. with \)

4.4 A readjustment rule

As we have seen, the Lambek grammar defines a relatively powerful system of rules, but several questions are still open with particular reference to the treatment of long distance dependencies and constituent conjunctions considered e.g. in Ades and Steedman(1982) and Steedman(1985). These phenomena, in fact, involve an inversion or *permutation* (in the sense of van Benthem 1986, 1987) of the functor-argument order, a fact that does not follow directly from the L system and has to be postulated as an additional axiom or rule. As an example relative to a simple clause, consider (19) where the *locative* PP occurs between the subject and the predicate:

(19)a. Maria sul tavolo mette una tovaglia nuova.
Mary on the table is putting a new table-cloth

```
b.  sul_tavolo  mette___   una_tovaglia_nuova
        PP      VP/PP/NP          NP
                __________________________A(r)
                          VP/PP
    ________________________________?
                   VP
```

At the stage of the derivation in which the argument PP must combine with its functor, we are faced with two options, both implying an extension of the Lambek system[9]:

option (i): if PP VP/PP -> VP,
then PP -> VP/(VP/PP) (by [2a]);

option (ii): PP VP/PP -> VP ([4a] under permutation)

In the first case, a complex type may be derived for the basic type of PPs which, inverting the functor-argument order, allows the combination with the predicate. Anyway, this strategy is costly since it has as a consequence the indefinite increasing of the raised types associated to the initial categories. Moreover, an important relation is lost, that holding, within the set of basic categories, between a functorial head(e.g. (VP/PP)/NP) and its subcategorized arguments. A second solution consists in allowing some kind of permutation rules within an **L** grammar. Since this move increases the generative power of the grammar (see van Benthem 1987, Steedman 1985), such rules need to be restricted, e.g. to specific structural positions. However, for the analysis of cliticization below, this second option is preferable, with the permutation rule schema presented in (20)[10]:

(20) Y X/Y -> Y **backward combination: B**

We may think of the rule above as a kind of *readjustment* operation putting distant or "dislocated" things in the right place with respect to the functor-argument order (reflecting the basic word order).

4.5 Unbounded dependencies and subject inversion

Unbounded dependencies will not be considered in detail here, but a brief mention is useful to better understand the role assigned to the rule above, within the present framework. Consider the examples in (21) the analyses of which are given in Figure 2 and Figure 3:

(21)a. Gianni ha detto che Maria ha perso il treno
Gianni said that Mary had missed the train

b. Che cosa hai detto che Maria ha perso?
What did you say that Mary had missed?

The direct object of the embedded clause in (21a) occurs in first position in (21b) as an effect of WH-movement in standard transformational grammar terms. From the point of view of a categorial grammar, the straightforward combination displayed in Figure 2, by means of the sole rule of (left and right) *application*, is no more possible, since the last functor of the sequence has missed its argument. As Figure 3 shows, the repeated use of *composition* generates the complex constituent (with a lacking NP-argument) adequate to combine with the leftmost NP. (This has to be distinguished from the possibly null pronoun which is the main subject.) Such a combination is performed, here, by the **B** rule, but it could be performed by some kind of type expansion rules as well[11].

- Figure 2 -

```
Gianni ha detto che  Maria ha perso il treno
NP     VP/S     S/S  NP    VP/NP    NP A(r)
                           ____________________
                                NP\S=VP A(1)
                           ____________________
                                  S A(r)
                     __________________________
                           S A(r)
            ___________________________________
                  NP\S=VP A(1)
_______________________________________________
                S
```

- Figure 3 -

```
Che cosa (tu) hai detto che Maria ha perso?
NP        NP  VP/S      S/S NP  (NP\S)/NP A(1),[4d]
                                ___________________
                                     S/NP C
                            _______________________
                                   S/NP C
                      _____________________________
                           NP\S/NP A(1),[4d]
                  _________________________________
                        S/NP B
_____________________________________________
                S
```

Finally, the B rule may be employed to "close" the derivation of sentences such as that in (22):

```
(22)  domani telefona  Maria
        NP\S            NP     L
                      _________
                      S/(NP\S)
      _________________________ B
                S
```

where a case of *subject inversion* is given, with respect to the example in (12): the subject occurs on the right of the predicate, contrasting the basic order predicted by the category assignment. This word order is very common in Italian and almost every simple sentence has an "inverted" counterpart (for comments and exceptions see Burzio 1986). Since the functor cannot correctly apply to its argument, a solution consists in raising the subject-type by means of the rule L and in "reconstructing" the appropriate functor-argument order by means of the rule B.

5. The analysis of cliticization

The readjustment rule B appears particularly useful in the analysis of cliticization contexts. Three main groups will be distinguished here with respect to their occurrence within the context: (i) object clitics NP, PP; (ii) object clitics N',N; (iii) subject clitics. The last two both concern *ne* cliticization, but exhibit relevant differences, as the categorial analysis will show.

5.1 Object clitics: NP, PP

The first cliticization pattern is presented in (23)-(24). It involves cliticization from a direct or indirect object position:

(23) a. Gianni ha dato un libro a Maria.
Gianni has given a book to Mary.

b. lo ha dato a Maria
(he) it has given to Mary

c. le ha dato un libro
(he) to her has given a book

d. glielo ha dato
(he) to-her-it has given

```
(24)a. lo ha_dato a_Maria      b. le ha_dato un_libro
       NP VP/PP/NP  PP            PP VP/PP/NP   NP
       __________B                   ____________A(r)
         VP/PP                          VP/PP
           __________A(r)         ______________B
              VP                        VP

    c. te  lo  do           cf. *lo te do
       PP  NP  VP/PP/NP
           ___________B
             VP/PP
       _________B
          VP

    d. te  lo  do
       PP  NP  (VP/PP)/NP  (by [4e])
       (NP.PP) VP/(NP.PP)
       _________________B
              VP

    e. glielo  ha dato
       (NP.PP) VP/(NP.PP)
       _________________B
              VP
```

The correct derivation follows quite easily by means of rule B, where, in each case, the category of the clitic is predicted by the functorial index of the verbal head: e.g. it is a NP in (24a) and a PP in (24b); in (24c) both the NP-object and the PP-object are cliticized, so that rule B applies twice. However, in this last case, we would like also to obtain a different derivation in which the constituent represented by the verb combines directly with a constituent made up of the set of clitics. In Italian, in fact, clitic items frequently form clusters such as *glielo* of (24e). The alternative derivation is given in (24d), where the two clitics are assumed to form a unique constituent[12] and the VP receives a category assignment which is derived by the rule given in theorem (4e). A parallel assignment occurs in (24e).

5.2 Object clitics: N′,N

The second pattern concerns cliticization from a NP in object position; compare (25):

(25) a. Ho visto due ragazze.
(I) have seen two girls.

b. Ne ho viste due.
(I) of them have seen two.

c. Ho letto un libro interessante.
I have read an interesting book.

d. Ne ho letto uno interessante.
I have read one interesting.

In both cases the clitic *ne* makes reference to the head of the NP which is the direct object of the verb. We will, therefore, assume that it is assigned to the category N^i ($0 \leq i \leq 1$), i.e. the category of the nominal head or of its immediate projection. The clitic being sited in the pre-verbal position, an effect of *determiner stranding* is produced, since the determiner occurs in isolation at the end of the sentence. In (26b), the application of *functional composition* allows the generation of a VP-Det constituent which is combined by rule **B** to the clitic *ne*:

```
(26)a.  Ho visto   due    ragazze
        VP/NP      NP/N'  N'
                   ____________A(r)
                        NP
        ________________________A(r)
                  VP

    b.  ne   ho viste   due
        N'   VP/NP      NP/N'
             _______________C
                  VP/N'
        ____________________B
                VP
```

In (27) we have a complex NP object so that, under cliticization, we have a string of three functor categories: VP/NP, NP/N′, N\N′. Composition should apply twice to produce the expected complex constituent, but at the first step functors contrasting in directionality are met:

```
(27)a.  Ho letto   un   libro   interessante
        VP/NP      NP/N' N      N\N'
                         ____________________A(l)
                                N'
                   ________________________A(r)
                           NP
        ______________________________A(r)
                   VP

    b.  ne   ho letto   uno   interessante
        N    VP/NP      NP/N'  N\N'
                        _______________C*
                            NP/N
             __________________________C
                        VP/N
        _____________________B
                 VP
```

To obtain the correct result we have to allow an extension of the *composition* rule to the effect that two functor categories contrasting in directionality may combine, under the assumption that directionality is imposed by one of the two (generally the main functor)[13]. The rule is instantiated in (28), where the definition C* (*strong composition*) is due to Moortgat(1986):

```
(28)    C*: N->N' . N'->NP = N->NP
        [directionality: main functor]
```

5.3 Subject clitics

The last pattern involves *ne* cliticization again, but this time the NP-domain which the clitic refers to is the sentential subject. Compare the examples in (29):

(29)a. Sono ritornate due ragazze.
Two girls came back.
b. Ne sono ritornate due.
(of) them came back two.

Here we have again a case of determiner stranding, but with different consequences with respect to the preceding examples:

```
(30)a.  Sono ritornate   due ragazze
        NP\S=VP          NP   L
                         S/(NP\S)  B
        ___________________________
                    S

    b.  Ne   sono ritornate   due
        N'   NP\S             NP/N'C*
                  _______________
                     N'\S      A(l)
        ____________________________
                    S
```

(31) C*: N'->NP . NP->S = N'->S
[directionality: main functor]

As before, the predicate combines with the determiner by *strong composition* (see (31)) but, assuming that directionality is given by the main functor, the resulting category is N'\S which directly applies to the clitic category. In this case, therefore, the readjustment rule B is not needed at all. Subject clitics, in fact, exhibit particular properties with respect to object clitics, which have been discussed in Belletti and Rizzi(1981) and, more recently, in Burzio(1986). These works show that a special set of predicates, which have been called *ergative*, allows *ne* cliticization with reference to a subject. This is excluded with *non-ergative* verbs such as *telefonare* (cf. **ne telefoneranno molte*) and in cases such as (32), where an intervening object occurs between the predicate and the inverted subject[14]:

(32)a. *ne hanno mangiato una mela due
(of)them have eaten an apple two

b. N' VP/NP NP NP/N'

The line of a simple explanation (only mentioned here) is suggested by the categorial analysis:a NP category occurs between the two functors blocking the application of the *composition* rule. Assembling the verb with the object NP, we obtain a complex predicate which is different in depth with respect to the (lexical) predicate associated to an *ergative* verb (cf.(30b)). As a further argument, consider the pattern in (33) again from Belletti and Rizzi (1981):

(33)a. Quante ne sono rientrate?
How many came back.

b. Due, ne sono rientrate.

c. *Due ne sono rientrate

While (33a) and (33b) are considered grammatical and accepted by the speakers, (33c), without any intonation pause, is rejected: as pointed out in Belletti and Rizzi (1981), (33b) is a case of topicalization of a postverbal determiner which, on the contrary, is adjacent to the clitic in (33c). We may so represent the situation:

(34) * (Det . cl)

where the creation of a constituent made up of a determiner and a clitic is not permitted. Now, how may we deal with the question within a categorial framework? Consider the example in (35):

(35) Quante ne sono rientrate
NP/N' * N' NP\S
____________?

Assuming the restriction in (34), we are not allowed to combine the determiner with the clitic. Anyway, since (35) is a grammatical string, at least one possible derivation has to be found. An immediate solution is given by theorem (4e), re-proposed in (36) and exemplified in (37b) with respect to the derivation in (37a):

(36) (X/Y)/Z <-> X/(Z . Y) **Interchange: I** [4e]

(37) a. NP-> (Det . N′)

b. (Det . N′)\S <-> N′\(Det\S)

On this basis, the derivation in (38) follows, where a complex constituent lacking a Det-argument is made up of the clitic and the predicate (this also represents a different solution to the analysis of subject-clitic constructions given in (30)):

```
(38)   Quante   ne   sono rientrate
       Det      N′   ____NP\S____ I           Det = NP/N′
                      N′\(Det\S)
                ________________ A(1)
                      Det\S
       _______________________ A(1)
                 S
```

Concluding remarks

The above analysis has shown that a categorial framework can be successfully applied to a wide fragment of a natural language including several word order phenomena. Specifically, the bidirectional symbolism defined within the Lambek model allows a characterization of *clitic* constructions distinguishing *object clitics* (within the VP-domain) from *subject clitics* (within the S-domain). The relation between the predicate and the determiner appears to be of relevance here, with particular reference to the determiner stranding effect obtained by means of the rule of functional composition

However, we have to face several open questions. In the first place, the status of the categorial rules has to be considered in more detail, in connection with the new extensions introduced to handle with discontinuous constituency. The unconstrained application of the rule B, for example, produces the undesirable strings in (40b), (40c) together with the well formed sentences in (39):

(39) a. Ho visto due ragazze.
(I) have seen two girls.

b. Ne ho visto due.
(I) of them have seen two.

(40) a. #ho visto due ne

b. #ragazze ho visto due

c. #libro ho letto uno interessante

and we also need to exclude (40a) which, as a matter of fact, is admitted by the analysis of (26b). Moreover, the domain of application of rules such as *strong composition* must be strictly determined, to avoid virtually annulling the flexibility of the bidirectional system.

These observations suggest that a restriction of categorial rules is needed with particular reference to those rules, such as rule B, which significantly increase the generative power of the system. In fact, restrictions on the application of backwards rules are stated in Ades and Steedman(1982) and in Steedman(1985). We think that the formal theory of features sketched above may be of valuable help in this respect: *constraints* on rule application may be stated as *feature co-occurrence restrictions* (see Gazdar *et.al.* 1985) to the effect that e.g. a functorial category which is assigned to a *determiner* does not apply to *unstressed* arguments like the clitic items; this would capture the "filter" given in (34), not permitting (under percolation of the relevant features) the derivation of structures such as: [Det cl] or [[VP Det] cl] (which yield to (33c) and (40a)).

In conclusion, a strategy deserves to be mentioned, which is alternative to the present approach based on the application of rule B and which proposes a different characterization of clitics. The main point is to consider clitic items as functors which apply to the verbal head, where the appropriate functorial categories are derived by means of *type raising* rules (see *option* (i) above). Requiring that clitics apply to lexical verbs only, overgeneration results, such as those in (40), are avoided by means of the different conditions holding for clitic items (functors) and *nouns* (arguments). This strategy, which has been suggested in Zwarts(1985), is coherent with the researches in progress in the field of *combinatory categorial grammar* (see Steedman 1987, Szabolcsi 1985, 1986).

NOTES

* This work is part of a research program on the Syntax and Semantics of Natural Language supported by the Philosophy Department of the University of Bologna and by the Humanities Faculty of the University of Chieti.The author is particularly grateful to Prof.G.Sandri for having stimulated her interest in the subject. She also wishes to thank Richard Oehrle and Anna Szabolcsi for their precious comments and suggestions on preceding formulations of the paper, although assuming full responsibility for the content. A preliminary version of this paper was presented at the annual conference of Italian Generative Grammar (Pisa, Spring 1985).

[1] Identity is not a necessary feature of the system: if it is included among the basic operations, the semigroup M become a *monoid* and allows the generation of the empty string. This may be of interest for further examination, but will not be considered here.

[2] For this point see Van Benthem(1986), (1987).

[3] Different denominations may be found in the literature for these rules, e.g. the term *forward partial combination* employed in Ades and Steedman (1982) for the rule in (4f); the term *type raising* or *type lifting* is frequently used for the rule in (4c), while the rule in (4g) is known as *division* or *Geach's rule*, due to the fact that Geach independently justified and introduced similar rule schemas (see Geach 1972).

[4] Bach (1983, p.103).

[5] See Gazdar et al.(1985, pp.17-41).

[6] For an analysis of head categories as functor categories, with respect to the *projection principle*, see Szabolcsi(1985).

[7] See Steedman (1985, pp.533 ff.), for the application of type-raising rules to subject-NPs and leftward-extracted items (topicalized NPs, *wh*-phrases, etc.); see also Dowty(1987), Zwarts(1985).

[8] The possibility of different, but equivalent, analyses of the same string is a main property of the Lambek model which is *structural complete* in this respect, i.e. associates a family of related types to an expression; see Moorgat(1986) for a discussion.

[9] In fact inverted structures of the kind Y X/Y don't follow directly from the axioms and the inference rules which allow the derivation of (left and right) oriented structures,i.e. X/Y Y and X X\Y; to obtain the required inversion it is necessary to assume some kind of *permutation* in the input condition of the inference rules (option (i)) or to postulate a rule the effect of which is to perform the functor-argument application in the correct way (option (ii)).

[10] The rule name derives from Ades and Steedman(1982) where a similar rule is defined, although within a different approach employing right oriented symbols.

[11] This is the strategy followed e.g. in Steedman(1985) to account for several kinds of unbounded dependencies in Dutch and English.

[12] This assignment follows from the assumption that the clitic-cluster is similar to a conjunction whose members reflect the order of occurrence of the predicate arguments (e.g., the 1st member is the object NP, etc.). A more natural assignment, corresponding to the order of the clitic-sequence, would require a rule such as: (X/Y)/Z -> (X/Z)/Y (which is stronger than (4e)) the effect of which is to change the argument order of the predicate.

[13] Instances of mixed composition are employed in Szabolcsi (1985) and the rule is discussed, with reference to the Lambek grammar, in Moortgat(1986).As pointed out in this latter work, the rule, defined as *strong composition*, is not derivable within the L system and its inclusion, as a new axiom, has the effect of collapsing L into its permutation closed version L*; however, we may assume a restricted version holding within local domains only (such as the N' domain in (25d),(26b)), e.g. stating appropriate lexical constraints for the words which may have access to the rule (see Moorgat 1986, p.8).

[14] We will not discuss the *ergative hypothesis* here, but an interesting solution (suggested by Anna Szabolcsi, pers. comm.) consists in assigning the category S/NP to *ergative* verbs like *arrivare* with respect to the basic assignment NP\S to *non-ergative* verbs.

REFERENCES

Ades,A.E. and M.J.Steedman:1982, 'On the Order of Words', *Linguistics and Philosophy*, 4, pp. 517-558.

Bach,E.:1983, 'Generalized Categorial Grammars and the English Auxiliary', in F.Heny and B.Richards(eds.), *Linguistic Categories: Auxiliaries and Related Puzzles*, Reidel, Dordrecht, 1983, vol.II, pp.101-120.

Bar-Hillel,Y.:1960, 'Some Linguistic Obstacle to Machine Translation', *Advances in Computers*, 1, rep. in Bar-Hillel(ed.)(1964), *Language and Information*, pp. 75-86.

Bar-Hillel,Y., C.Gaifman, and E.Shamir:1960, 'On Categorial and Phrase Structure Grammars', *The Bull. of the Research Council of Israel, pp.1-16*, rep. in Bar-Hillel(ed.)(1964), *Language and Information, pp. 99-115.*

Belletti,A. and G.Rizzi:1981, 'The Syntax of *ne*: some Theoretical Implications', *The Linguistic Review*, 1, pp. 117-154.

Benthem,J.van:1986, 'Categorial Grammar and Lambda Calculus', in D. Skordev(ed.), *Druzhba Summer School in Applied Logic*, Plenum Press, New York.

Benthem,J.van:1987,'The Lambek Calculus'in Oehrle *et al.*(eds.)(1987)

Burzio,L.:1986, *Italian Syntax. A Government-Binding Approach*, Reidel, Dordrecht.

Buszkowski,W.:1987, 'Power of Categorial Grammars', in Oehrle *et al.* (eds.)(1987).

Dowty,D.:1987, 'Type Raising, Functional Composition, and Non-Constituent Conjunction',in Oehrle *et al.*(eds.)(1987).

Gazdar,G., E.Klein, G.K.Pullum and I.A.Sag:1985, *Generalized Phrase Structure Grammar*, Oxford, Blackwell.

Lambek,J.:1958, 'The Mathematics of Sentence Structure', *American Math. Monthly*, 65, pp. 154-170.

Lambek,J.:1985, 'Categorial and Categorical Grammars',in Oehrle *et al.*(eds.)(1987).

Moortgat,M.:1986,'Lambek Categorial Grammar and the Autonomy Thesis' presented at ZWO Symposium "Morphology and Modularity",Utrecht.

Oehrle,R.T., E.Bach and D.Wheeler(eds):1987, *Categorial Grammars and Natural Language Structures*, Reidel, Dordrecht (forthcoming).

Steedman,M.J.:1985, 'Dependency and Coordination in the Grammar of Dutch and English', *Language*, 61, pp.523-568.

Steedman,M.J.:1987, 'Combinators and Grammars', in Oehrle *et al.* (eds.)(1987).

Szabolcsi,A.:1985, 'Combinators and the Projection Principle', paper presented at the conference: *Categorial Grammars and Natural Language Structures*(Tucson 1985).

Szabolcsi,A.:1986, 'Filters versus Combinators', in Bodnár et al. (eds.), *Intensional Logic, Semantics and Philosophy of Science*, Festschrift for I.Ruzsa, (forthcoming).

Zwarts, F.:1985, 'Composition Rules in Categorial Grammar', paper presented at the conference: *Categorial Grammars and Natural Language Structures*(Tucson 1985).

INSTITUTE OF PHILOSOPHY
UNIVERSITY OF CHIETI, ITALY

ANNA SZABOLCSI

ON COMBINATORY CATEGORIAL GRAMMAR*

0. Objective

Traditionally, the main attraction of categorial grammar, initiated in Ajdukiewicz (1935), lies in the parallel treatment of syntactic and semantic compositionality. In recent years there has been a revival of interest in categorial syntax on its own right, resulting in various extensions with linguistic and/or logical motivation. One such line of research has led to the proposal that the interpretation of the concatenation of expressions should go beyond functional application to include a number of operations identifiable with certain combinators in Curry and Feys (1958). As part of this program, the present paper will briefly address the following three issues:[1]

(1) General linguistic motivation in terms of the Projection Principle
(2) Specific linguistic motivation for certain combinators
(3) Questions for further research

1. The Projection Principle

Current theories of grammar share the natural assumption that syntactic representations must be compatible with the relevant properties of the lexical items they contain. In Chomsky (1981, 29) this is expressed as follows:

(4) Projection Principle
Representations at each syntactic level [...] are projected from the lexicon, in that they observe the subcategorization properties of lexical items.

Although technically speaking (4) is but one specific necessary condition on well-formed syntactic representations, "projection from the lexicon" has a conceptually distinguished status. It is agreed that any other proposed principle (condition) may be called into question, but not (4). Moreover, any other proposed principle (condition) tends to have the best chance to be accepted if it can be derived from (4) in some sense.

If this distinction is justified, it is important to observe that not all principles that have been found necessary to characterize the crucial aspects of syntactic wellformedness appear to be derivable from (4), or from lexical properties in general. Even the syntactic "tools" that are assumed to enable lexical items to satisfy their needs have to be constrained in a way that does not follow from lexical properties. We may now ask whether these constraints reveal something deep about language. If they do, then the "laws of syntax" are to a significant extent independent of the properties of its "lexical substance". Alternatively, those constraints may simply indicate that we are not using the right "tools". It may be the case that making more explicit assumptions about the nature of lexical items we can design "tools" that capture the effect of those constraints in a constructive way. In other words, those constraints may reveal something deep but not qua constraints; rather, as observations helping to identify the "real tools". If so, the Projection Principle might be a sufficient condition, i.e., the constitutive principle of grammar.

The program that emerges from these considerations involves the following tasks:

(5) a. Model lexical items in a way that determines what kind of "tools" may be operative in grammar.
b. Find a specific set of such tools that produce all and only grammatical results.
c. Ideally, show that this set is coherent in some relevant sense; i.e., that it has some property that makes its appearance in grammar more likely than that of any arbitrary set in the domain implied by (a).

The proposal to be made is based on a minimal but essential departure from commonplace assumptions. It is agreed that to say that a verb subcategorizes for two complements is basically the same as to say that it is a function/functor with two argument(place)s. Does it make a difference, then, whether we talk about lexical items with subcategorization properties or about lexical items as functions? It does because we have independent knowledge of the behavior of functions but not of the behavior of items with subcategorization properties. Whatever comes as a theorem on the former view will appear as a contingent empirical matter on the latter: it needs to be discovered with hard work, and even if it is successfully discovered, it can only be built into the theory in the form of a new axiom.

It seems useful, then, to take lexical items to be functions/functors. This leads to the same view that categorial grammars have always adopted, although not necessarily exploiting all its consequences.

The minimal implication this way of looking at lexical items has for syntax is that the concatenation of expressions may be interpreted as functional application. Now, if the categories of expressions are kept constant, and if application is the only interpretation for concatenation,[2] then we predict that lexical functors must be adjacent to their arguments. One set of data that may constitute a glaring counterexample to this includes sentences informally describable as containing a "gap" left by a leftward or rightward "extracted" constituent:

(6) a. I think that Mary saw Bill.
    b. WHO do you think that Mary saw --?
(7) a. I put the cup on the table.
    b. I put -- on the table A CUP OF TEA AND A JAR OF MILK.

It may be claimed that the counterexamples are only apparent and the difficulty can be avoided, for instance, as follows. The gap is immediately filled by some designated item interpreted as a variable; the extracted constituent is prefixed or suffixed to the sentence much in the same way as abstraction operators are introduced in the lambda calculus; and the properly bound variable item is deleted or ignored at the end of the derivation. This is essentially the solution followed by various works in both the Chomskyan and the Montagovian schools.

This solution makes one expect that the possibilities for gaps and extracted constituents to occur in natural language are essentially the same as those for operators and variables in lambda calculus syntax. This expectation is not borne out. A few paradigmatic points of divergence are as follows:

(8) a. Free variables: fx
       * -- saw Bill.
    b. Vacuous operators: $\lambda x[a]$
       * What did Mary see Bill?
    c. Crossed binding: $\lambda x \lambda y[fx(gy)]$
       * What1 do you wonder who2 to talk about --1 to --2?
    d. Binding over arbitrary domains: $\lambda x[\ldots x \ldots]$
       * Who1 did you go home before --1 fell asleep?

This leaves us with the need for constraints that was claimed to be problematic above.[3]

The advantage of the present approach is that we need not (implicitly or explicitly) model our syntax after the lambda calculus. There are a variety of operations that can be performed on typed functions, and therefore the way we modelled lexical items implies that in principle any of these can serve as the interpretation of concatenation, cf. (5a). The question is whether we can indeed find a specific set of such operations that yield all and only the correct results, cf. (5b).

2. Specific combinators

There have in fact been various proposals to extend the apparatus of categorial grammar with operations on functions. The prime example is the use of functional composition by Ades and Steedman, Lambek, Moortgat, and others. (9) and (10) are based on Steedman (1987). (The formulation of concatenation as application is quoted in order to make clear the use of directional categories.)

(9) Concatenation interpreted as application:
a. α of X/Y, β of Y -> $\alpha\beta$ of X, $\alpha'(\beta')$
b. β of Y, α of X\Y -> $\beta\alpha$ of X, $\alpha'(\beta')$.

(10) Concatenation interpreted as composition:
a. α of X/Y, β of Y/Z -> $\alpha\beta$ of X/Z, $\lambda x[\alpha'(\beta' x)]$
b. β of Y\Z, α of X\Y -> $\beta\alpha$ of X\Z, $\lambda x[\alpha'(\beta' x)]$
c. β of Y/Z, α of X\Y -> $\beta\alpha$ of X\Z, $\lambda x[\alpha'(\beta' x)]$.

Given this extension, and some auxiliary assumptions to be made explicit below, the whole string DO YOU THINK THAT MARY SAW in (6) can be composed into one big functor with the same right-NP domain as SAW. Given (9), it cannot apply to the intended argument WHAT; if however we assume that the ability of WHAT to appear on the left is grounded in its lexical category, namely, S/(S/NP), then it may apply to DO YOU THINK THAT MARY SAW, yielding the correct interpretation.

It is even more interesting to observe that the use of composition does not only enable us to handle sentences containing "extracted constituents" and "gaps" without using placeholder variables; at least a significant portion of the sentences that need to be filtered out in grammars without composition are also straightforwardly accounted for. SAW BILL is not a sentence with an illicit free variable: it is simply a functor S\NP. WHAT DID MARY SEE BILL cannot arise because DID MARY SEE BILL is not a functor with domain NP. The string WHAT DO YOU WONDER WHO TO TALK ABOUT TO can only be assembled to carry the unreal but nested reading that you want to talk to a thing about a person. As the categories below indicate, TO has no hope to combine prior to WHO getting combined and interpreted as the object of ABOUT:

```
(11) ... who  to talk     about  to
              (S/PP)/PP   PP/NP  PP/NP
              ---------------
                 (S/PP)/NP
```

The string WHO DID YOU GO HOME BEFORE FELL ASLEEP cannot be assembled at all: concatenation as composition is not defined for BEFORE of S/S and FELL ASLEEP of S\NP. If this sample of examples is representative, we can say that composition actually explains why "extraction" is possible at all, and under what circumstances it is possible.

Promising as this line of research seems to be, it is immediately clear that there are sentences whose existence is easily understood under the variable introduction approach but cannot be handled by the grammar extended with just composition:

(12) One operator binding two variables: $\lambda x[...x....x...]$
What1 did you refuse --1 without looking at --1?

Generative grammarians call the second "gap" a parasitic one to express the fact that it is illicit unless there is another "gap" corresponding to the same "extracted" constituent. The relative positions of the two gaps also matter, so constraints need to be invoked again, cf.: [4]

(13)* Who1 do you think --1 refused the offer without
John threatening --1?

These constraints are suggestive of a further specific operation on functions being at work here, however. In Szabolcsi (1983) it was introduced under the name connection (Steedman 1987 renamed it as substitution); I present only one directional version below: [5]

(14) Concatenation interpreted as connection:
α of (X\Y)/Z, β of Y/Z -> $\beta\alpha$ of X/Z,
with the interpretation $\lambda x[\alpha' x(\beta' x)]$.

Connection straightforwardly handles (12) but not (13):

```
(12) ... refuse   without looking at
         VP/NP      (VP\VP)/NP
         -------------------------
                VP/NP

(13)* ... refused the offer   without John threatening
              VP (=S\NP)             (VP\VP)/NP
```

Finally, let us spell out the auxiliary assumption made in the composition examples. One rather innocent change in categories can be effected by directionality preserving type-raising in syntax. Apart from other uses discussed in Dowty (1985), this option is necessary for the subjects in (6) to acquire the category S/(S\NP):

(15) Type-raising:
α of X -> α of Y/(Y\X), Y\(Y/X), $\lambda f[f(\alpha')]$.

(6) Who do you think that Mary saw

```
(6)  Who do you think that Mary   saw
     .......................  NP    (S\NP)/NP
                             --------
                             S/(S\NP)
                             -------------------
                             S/NP
```

Steedman (1985) raises the question of what inventory of operations composition, connection and type-raising are drawn from, and observes that they are identical to the combinators B, S and C*(=CI) of Curry and Feys (1958). To quote their definitions:

(16) B = $\lambda fgx.f(gx)$
S = $\lambda fgx.fx(gx)$
C = $\lambda fxy.fyx$
I = $\lambda x.x$

As Steedman points out, this correspondence is interesting for the following reasons. First, one of the fundamental results of combinatory logic is that the expressive power of the lambda calculus can be obtained without using abstraction and bound variables. Second, while the use of variables is a major source of computational overheads in practical programming languages, the use of the particular combinators in (16) has been shown to be very efficient. Therefore, the linguistically motivated suggestion that the operations introduced above are (among) the "right tools" for grammar seems to receive coherent mathematical and psychological support.

3. Questions for further research

The above argument for the combinatory nature of grammar was based on a specific kind of examples. Let us assume that their proposed treatment is in fact empirically adequate and explanatory. We may ask now whether combinatory logic can serve as a guideline for further research, i.e., whether it is suggestive of how other phenomena should be treated, and whether it can possibly help us to detect some coherence in the set of actual grammatical operations (cf. 5c). In what follows I will make some suggestions and also mention some apparent problems.

Let us start with that property of combinatory logic that was found relevant in the previous section, namely, that it has no bound variables, and note that the sentences discussed there actually did not contain any overt lexical item that could be interpreted as a bound variable (the alleged variables were "gaps"). Nevertheless, there are lexical items that can be suspected to have exactly this interpretation. Consider the personal and reflexive pronouns below:

(17) a. Everyone thinks that Mary loves him.
∀x[x thinks that Mary loves x]
b. Everyone loves himself.
∀x[x loves x]

Although the existence of such items may seem like a problem, I wish to argue that the combinatory approach is actually also revealing here.

The fact that reflexives must be bound and personal pronouns may be bound is usually taken to be a primitive property of these items; furthermore, constraints are used to make sure that they are bound correctly, i.e., by a more prominent argument inside/outside some specified domain.[6] However, there is a natural alternative to this treatment; namely, to assume that the definitive properties of these items are inherent in their lexical meaning, which can be made explicit in combinatory terms.

Consider the following derivation for (17b); further details are discussed at length in Szabolcsi (1987):

```
(18) Everyone      loves             himself
     S/(S\NP)      (S\NP)/NP         (S\NP)\((S\NP)/NP)
     λf∀x[fx]      λzλy[LOVEzy]      λgλu[guu]
                   --------------------------------
                            S\NP
                            λu[LOVEuu]
     ----------------------------------------------
                   S
                   ∀x[LOVExx]
```

The starting point in devising this proposal is that the lexical interpretation of HIMSELF is that of the duplicator. In order for HIMSELF to have such an interpretation, it must have a matching category. Such a matching category can be obtained by type-raising, cf. (15), with Y = Z|NP. If we assume that the combinator <u>C</u> in itself is not operative in our grammar, i.e., that the argument order of functions is fixed in the lexicon, then the duplicator interpretation will automatically capture the most salient constraint on the binder as well, namely, that it must be a more prominent argument.[7]

The duplicator is known as combinator <u>W</u> in Curry and Feys (1958). Reference to it offers an occasion to contemplate about what/how many operations may be expected to be at work in grammar. If we identify the set of possible grammatical operations with the set of possible combinators, we are faced with an undesirable embarras de richesse. By using composite combinators we can perform indefinitely complex operations in one swoop. It is obvious that at most a very small subset of such operations are plausible in the grammar of natural languages. Now, one possibility is to wait and see what we will need in practice. (In fact, this kind of wait-and-see strategy is followed by all current theories of grammar, whatever tools they use.) Another possibility is to try to make predictions on the basis of distinctions that can be meaningfully made among combinators. One such distinction can be made in terms of interdefinability. It is interesting to observe that the combinators in (16), i.e., B, S, C, and I constitute a primitive set; if we maintain that we actually need B, S and C*, that is another set of primitives, with less expressive power. We may take this to indicate that natural languages restrict the set of operations freely available in syntax to such primitive sets. A number of combinators definable in terms of those primitives may still be available, but only if they are embodied in the categories or interpretations of specific lexical items, whose properties need to be learned separately in any case. Now, how should the status of W be judged? In typefree terms it is definable as follows:

(19) a. W = $\lambda fx.fxx$
b. S(C*)= $\lambda ghx.gx(hx)(\lambda yf.fy) = \lambda fx.fxx$

If we have a reason to assume that our combinators are essentially typefree, only their "implementation" is typed, then the fact the apparently lexical duplicator is definable in terms of combinators that we claim to be free in syntax can be taken to support this reasoning.

It is time now to turn to problems. I will restrict my attention to two issues that are easy to expose in the present context.

First, notice that reflexives and bound pronouns differ with respect to the locality of their possible binders:

(20) a.* Everyone1 thinks that Mary loves himself1.
b.* Everyone1 loves him1.

In the paper referred to above I suggested that these properties can in fact be captured in English (i) by requiring that reflexives only apply to lexical verbs, and (ii) by assigning a B(BW)B kind of interpretation to bound pronouns in the lexicon (which is otherwise motivated by cases known as pied piping). Nevertheless, locality problems can be replicated with "gaps" as well. While none of the sentences in (21) are very good, their grammaticality surely decreases as the two gaps get "closer" to each other:

(21) a.? Who1 did you warn --1 that they were going to arrest --1?

b.?? Who1 did you give a picture of --1 to --1?
c.* Who1 did you introduce --1 to --1?

Our syntax is insensitive to such distinctions, however, and there being no lexical item in the place of gaps, the kind of lexically grounded solution for (20) does not carry over to (21), either.

Secondly, consider the fact that it is sometimes possible to replace illicit gaps by so-called resumptive pronouns. (22b) is marginal in English but its Swedish equivalent (23b) is standard:

(22) a.* Which word1 did no one know how --1 is spelled?
b.? Which word1 did no one know how it1 is spelled?
(23) a.* Vilket ord1 visste ingen hur --1 stavas?
b. Vilket ord1 visste ingen hur det1 stavas?

Operator-bound resumptive pronouns obviously fall outside the scope of combinatory grammar. Their existence is especially disturbing given the possibility of the following conjunction (all these examples are taken from Engdahl):

(24) Det finns viss ord1 (som) jag ofta träffar på --1 men aldrig minns hur de1 stavas.
'There are certain words (that) I often come across but never remember how they are spelled'

It was observed at the outset that a syntax "modelled after" the lambda calculus will necessarily invoke axiomatic kind of constraints. The question was raised whether such constraints reveal something deep about the nature of grammar, or they reveal the inadequacy of the tools that need to be constrained. It appeared that the latter can be argued to be the case: an interesting constructive treatment of many crucial cases could be developed by exploiting the possibility of introducing specific combinators into the grammar. The examples just mentioned indicate, however, that a pure combinatory grammar is guilty of both overgeneration and undergeneration. Even more importantly, it turned out that natural language makes use of variable binding strategies that are inevitably suggestive of the syncategorematic nature of the binder. These facts indicate that no homogeneous combinatory account can be given for the range of phenomena considered.

There seem to be two ways to interpret this situation. One is to say that these phenomena come under the same heading and therefore they must be handled with homogeneous tools. In this case the problematic cases are to be the paradigmatic ones. This means that the grammars criticized at the outset are in fact on the right track: the constraints they use do indeed reflect the (funny) nature of grammar. The considerations put forth in connection with combinators may then have a metatheoretical kind of status in that they explain something about the nature of those constraints.

It may be the case, of course, that the homogeneity requirement can be eliminated in the light of some independently justifiable division of labor between combinatory and non-combinatory tools in grammar. For that reason it is not clear yet whether we have a draw, a stalemate, or a winner. In any case, it seems to (have) be(en) an instructive game.

NOTES

* This paper is a revised version of Szabolcsi (1985, 1986). I am particularly grateful to J. Higginbotham and L. Kálmán for recent discussions.

1. For historical surveys, see Casadio (1986, 1987).
2. See Engdahl (1986), Hausser (1984), Montague (1974) and Partee (1979).
3. See Gazdar et al. (1985) and van Riemsdijk and Williams (1985).
4. See Chomsky (1982), Engdahl (1983) and Kayne (1983).
5. This operation does not cover parasitic gap sentences with subjacency violations, e.g., HE IS A MAN WHO1 EVERYONE WHO KNOWS --1 ADMIRES --1. The essence of my proposal for these lies in assigning category ((S/NP)/((S\NP)/NP))/(R/NP) to EVERYONE.
6. See Chierchia (1985), Chomsky (1981), Cooper (1983).
7. This account, like that of Reinhart (1983), assimilates bound pronouns to reflexives and free pronouns to names.

REFERENCES

Ades, A. — M. Steedman (1982), On the order of words. LINGUISTICS AND PHILOSOPHY 4.

Bach, E. — R. Oehrle — D. Wheeler, eds. (1985), CATEGORIAL GRAMMARS AND NATURAL LANGUAGE STRUCTURES. Reidel (in press).

Casadio, C. (1985), Semantic categories and the development of categorial grammars. In Bach et al. (1985).

Casadio, C. (1987), Extending categorial grammars. This volume.

Chierchia, G. (1985), Aspects of a categorial theory of binding. In Bach et al. (1985).

Chomsky, N. (1981), LECTURES ON GOVERNMENT AND BINDING. Dordrecht: Foris.

Chomsky, N. (1982), SOME CONCEPTS AND CONSEQUENCES OF THE THEORY OF GOVERNMENT AND BINDING. Cambridge: MIT Press.
Cooper, R. (1983), QUANTIFICATION AND SYNTACTIC THEORY. Dordrecht: Reidel.
Curry, H. — R. Feys (1958), COMBINATORY LOGIC I. Amsterdam: North Holland.
Dowty, D. (1985), Type raising, functional composition and non-constituent conjunction. In Bach et al. (1985).
Engdahl, E. (1983), Parasitic gaps. LINGUISTICS AND PHILOSOPHY 6.
Engdahl, E. (1986), CONSTITUENT QUESTIONS. Dordrecht: Reidel.
Gazdar, G. — E. Klein — I. Sag — G. Pullum (1985), GENERALIZED PHRASE STRUCTURE GRAMMAR. Oxford: Blackwell.
Hausser, R. (1984), SURFACE COMPOSITIONAL GRAMMAR. München: Wilhelm Fink.
Kayne, R. (1983), Connectedness. LINGUISTIC INQUIRY 14.
Lambek, J. (1958), The mathematics of sentence structure. AMERICAN MATHEMATICAL MONTHLY 65.
Montague, R. (1974), FORMAL PHILOSOPHY. New Haven: Yale.
Moortgat, M. (1985), Mixed composition and discontinuous dependencies. In Bach et al. (1985).
Partee, B. (1979), Montague grammar and the well-formedness constraint. In Heny — Schnelle, eds., SYNTAX AND SEMANTICS 10. New York: Academic.
Reinhart, T. (1983), ANAPHORA AND SEMANTIC INTERPRETATION. London: Croom Helm.
van Riemsdijk, H. — E. Williams (1985), INTRODUCTION TO THE THEORY OF GRAMMAR. Cambridge: MIT Press.
Steedman, M. (1985), Combinators and categorial grammars. In Bach et al. (1985).
Steedman, M. (1987), Combinatory grammars and parasitic gaps. NATURAL LANGUAGE AND LINGUISTIC THEORY (in press).
Szabolcsi, A. (1983), ECP in categorial grammar. Ms. Max-Planck-Institut für Psycholinguistik.
Szabolcsi, A. (1985), Combinators and the projection principle. Paper to the conference on Categorial grammars and natural language structures, Tucson.
Szabolcsi, A. (1986), Filters versus combinators. In Bodnár — Máté — Pólos, eds., INTENSIONAL LOGIC, SEMANTICS, AND THE PHILOSOPHY OF SCIENCE. Budapest (in press).
Szabolcsi, A. (1987), Bound variables in syntax (Are there any?). In Groenendijk — Stokhof — Veltman, eds., PROCEEDINGS OF THE 6TH AMSTERDAM COLLOQUIUM. Amsterdam: ITLI.

INSTITUTE OF LINGUISTICS
HUNGARIAN ACADEMY OF SCIENCES

Chomsky, N. (1982). SOME CONCEPTS AND CONSEQUENCES OF THE THEORY OF GOVERNMENT AND BINDING. Cambridge, Mass.: MIT Press.
Cooper, R. (1983). QUANTIFICATION AND SYNTACTIC THEORY. Dordrecht: Reidel.
Curry, H. — R. Feys (1958). COMBINATORY LOGIC. Amsterdam: North Holland.
Dowty, D. (1988). Type raising, functional composition, and nonconstituent conjunction. In Oehrle et al. (1988).
Engdahl, E. (1983). Parasitic gaps. LINGUISTICS AND PHILOSOPHY 6.
Engdahl, E. (1986). CONSTITUENT QUESTIONS. Dordrecht: Reidel.
Gazdar, G. — E. Klein — I. Sag — G. Pullum (1985). GENERALIZED PHRASE STRUCTURE GRAMMAR. Oxford: Blackwell.
Hausser, R. (1984). SURFACE COMPOSITIONAL GRAMMAR. München: Wilhelm Fink.
Kayne, R. (1983). Connectedness. LINGUISTIC INQUIRY 14.
Lambek, J. (1958). The mathematics of sentence structure. AMERICAN MATHEMATICAL MONTHLY 65.
Montague, R. (1974). FORMAL PHILOSOPHY. New Haven: Yale.
Moortgat, M. (1988). Mixed composition and discontinuous dependencies. In Oehrle et al. (1988).
Partee, B. (1979). Montague grammar and the well-formedness constraint. In Heny — Schnelle (eds), SYNTAX AND SEMANTICS 10. New York: Academic Press.
Reinhart, T. (1983). ANAPHORA AND SEMANTIC INTERPRETATION. London: Croom Helm.
van Riemsdijk, H. — E. Williams (1986). INTRODUCTION TO THE THEORY OF GRAMMAR. Cambridge, Mass.: MIT Press.
Steedman, M. (1988). Combinators and grammars. In Oehrle et al. (1988).
Szabolcsi, A. (1987). [illegible] NATURAL LANGUAGE AND LINGUISTIC THEORY [illegible]
Szabolcsi, A. (1989). ECP in categorial grammar. Ms., Max-Planck-Institut für Psycholinguistik.
Szabolcsi, A. (1989). Combinators and the projection principle. Paper at the Conference on Categorial Grammar and natural language [illegible]
Szabolcsi, A. (1989). [illegible] In [illegible]
Szabolcsi, A. (1983). The possessor that ran away from home. THE LINGUISTIC REVIEW 3.

INSTITUTE OF LINGUISTICS
HUNGARIAN ACADEMY OF SCIENCES

ANATOLI STRIGIN

LOGIC FOR SYNTAX

Transformational grammar can be viewed as a theory of internal syntactic representations that are computed, for instance during parsing. The minimal hypothesis for a theory of this structural level (computational level theory) is that it gives only a limited characterization of computational processes constructing the representations. A fuller characterization is the proper subject of a theory of processes (algorithmic level theory of Marr (1982)). This seems appropriate for two reasons. A grammar as a formal system is neither bound to any particular process philosophy like distributed or centralized processing, nor does it make any claims about particular algorithms computing the representations. Both aspects belong to the algorithmic level. Introducing the two theoretical levels raises the question of correspondences between them, though. Such correspondences are not obvious. The situation gets all the more complicated if one relies on some particular programming language or on some particular algebraic model like a pushdown automaton. It would be more profitable to have some noncommittal way of discussing algorithms, if only for the sake of establishing which of the syntactic relations described by the grammar are more important conceptually, or to have some leeway with respect to theories of implementation level, e.g. connectionist models, etc. Here I want to discuss a rather speculative line of thought bearing on the issue, in the hope that it is of interest both to logicians and to linguists.

These speculations were suggested to the author by the work of Szabolcsi (forthcoming) in combinatory categorial grammar (CCG), and by the work of Martin-Löf in the intuitionist theory of types (Martin-Löf (1975),(1982)), neither of whom could be held responsible for possible misinterpretations of their ideas. The aim of CCG grammar is to replace the static description of the syntactic structures in the "government and binding" type of theory (Chomsky (1981), GB for short) by some sort of functional description. In CCG syntactic categories are interpreted as function types. The basic operation is application. It is hoped that such a description provides a more adequate explanation of constraints on such operations as wh-movement in GB.

The syntactic structure arises in observance of the fact that the functions are typed: every lexical item is assigned an allowed type (or more), there are typed combinators (interpreted as rules), and typed application. Hence, theoretical analysis of a sentence looks like an attempt to assign it a designated type S, all types for elementary units (e.g. words) being assumed as given. The functions of the types given by syntactic categories are semantic representations, i.e. they are semantic by nature. However, a functional description does not impose this interpretation, and it is legitimate to view these representations as some kind of 'syntactic' functions. Although the interpretation of those functions as semantic representations is natural, it skips the level I am interested in.

To produce a syntactic analysis of a sentence we have to proceed according to certain rules by which types may be assigned to an untyped expression. These rules (called <u>type</u> <u>assignment</u> <u>schemes</u> or TAS, for short) look like this in the case of pure combinatory logic:

(1) Assume we have expressions built up from the two combinators $\underline{S}$ and $\underline{K}$ and some ground stock of type constants and type variables. A type assignment formula is $X \in \alpha$, where X is untyped, and α is a type scheme defined by

- (i) each type-constant or type-variable is a type scheme;
- (ii) where α, β are type schemes, $(\alpha\text{->}\beta)$ is a type scheme;

(2) a TAS for combinatory logic is a deductive system consisting of the two axiom schemes

$$(\text{->}\underline{K}) \quad \underline{K} \in (\alpha\text{->}(\beta\text{->}\alpha))$$
$$(\text{->}\underline{S}) \quad \underline{S} \in ((\alpha\text{->}(\beta\text{->}\gamma))\text{->}((\alpha\text{->}\beta)\text{->}(\alpha\text{->}\gamma)))$$

assigning type schemes to $\underline{K}$ and $\underline{S}$, which are defined by

$$\underline{K}xy=x$$

and

$$\underline{S}fgx=(fx)(gx)$$

and a rule of inference describing the behaviour of the types under application

$$(\text{->}e) \quad \frac{X \in (\alpha\text{->}\beta) \qquad Y \in \alpha}{(XY) \in \beta}$$

This definition is an abbreviation of the one given in Hindley (1985).

If we try now to assign a type scheme to, say, $\underline{SKK}$, we obtain the following:

$$\cfrac{\cfrac{\underline{S} \in (\alpha\text{->}\beta\text{->}\alpha)\text{->}(\alpha\text{->}\beta)\text{->}\alpha\text{->}\alpha \qquad \underline{K} \in \alpha\text{->}\beta\text{->}\alpha}{(\underline{SK}) \in (\alpha\text{->}\beta)\text{->}\alpha\text{->}\alpha} \qquad \underline{K} \in \alpha\text{->}\alpha\text{->}\alpha}{\underline{SKK} \in \alpha\text{->}\alpha}$$

where β is $(\alpha\text{->}\alpha)$, and $\alpha\text{->}\beta\text{->}\gamma$ is $(\alpha\text{->}(\beta\text{->}\gamma))$, etc.

Consider now the two parts of this derivation separately. If the type schemes are suppressed, we have the construction of SKK. On the other side, the axiom schemes (->K) and (->S) and the rule (->e) constitute an axiomatization of the intuitionisticimplication calculus, if the names of the combinators are dropped, and -> is interpreted as implication, with (->e) being the rule of modus ponens. This calculus is intuitionistic, because although the two axiom schemes occur in classical calculus as well, the positive implicational formula $((\alpha\rightarrow\beta)\rightarrow\alpha)\rightarrow\alpha$ is a theorem in the classical system, but not in the system without negation.

This correspondence between the types, and the intuitionist implicational logic was described in Curry & Feys (1958). As Martin-Löf argues, it is basic to the whole intuitionistic philosophy. Essentially, the following principles seem to be responsible (this is not a rendering of Martin-Löf's ideas, for which his work should be consulted):

- every constructive object comes equipped with a type;
- every type is defined by constructions, producing objects of this type ;
- propositions are types, namely the types of the proofs of the propositions in questions. Hence, a combinator expression is a proof of the corresponding theorem.

Propositions must be distinguished from judgements that propositions hold.

Now what is interesting about this philosophy from the point of view of syntax is the way it regards the types: types are derivative of the notion of construction.

Given that syntactic categories are typed functions, there is an evident analogy: a category can be considered a notion derivative from the notion of a construction for this category. Type assigning in CCG is a stipulation of which kinds of programs must be run to produce an element of a category. That is essentially the procedural interpretation of types given in Martin-Löf (1982). In other words, possible success of a description of natural language grammar in the framework of CCG paired with the *intuitionistic* (or should it better be called *constructivistic ?*) philosophy suggests regarding the syntactic categories as being defined by the programs computing objects of these categories, i.e. by the programs for parsing, and for generation. The syntactic structure of a sentence in CCG amounts then to a hypothesis about the kinds of programs used. The details of the programs should be supplied by the algorithmic level theory.

The description obtained in this way may be crude, but it seems to be useful. Consider an example taken from Szabolcsi (forthcoming). A reflexive pronoun like 'himself' must be manifestly nominal, and at the same time perform an operation of identifying two arguments of a transitive verb (this is only an example, mind you). In terms of combinators it should be a type-raised noun, and the duplicator at once. The type-raising combinator $\underline{C*}$ is $\lambda xy.yx$ in λ-calculus notation. To determine its type scheme two rules for type assignment in the natural deduction style will be used (Hindley (1985)):

$$(\to e)\ \frac{M \in (\alpha\to\beta) \qquad N \in \alpha}{(MN) \in \beta}$$

$$\begin{array}{c}[x \in \alpha]\\ \cdot\\ \cdot\\ (\to i)\ \dfrac{M \in \beta}{(\lambda x.M) \in (\alpha\to\beta)}\end{array}$$

Rule (->i) means that if we already have a deduction of a formula $M \in \beta$ from $x \in \alpha$, and, perhaps some set of assumptions $\underline{B}$, then we can deduce, from $\underline{B}$ alone the formula $(\lambda x.M) \in (\alpha\to\beta)$.

The computation goes like this:

$$\dfrac{\dfrac{\dfrac{[y \in (\alpha\to\beta)] \qquad [x \in \alpha]}{(yx) \in \beta}\ (\to e)}{\lambda y.yx \in ((\alpha\to\beta)\to\beta)}\ (\to i)}{\lambda xy.yx \in (\alpha\to((\alpha\to\beta)\to\beta))}\ (\to i)$$

The assumptions are cancelled here. Now the type of the duplicator $\underline{W}$, i.e. of $\lambda fx.fxx$, is $((\alpha\to(\alpha\to\beta))\to(\alpha\to\beta))$. Given the application of $\underline{C*}$ to an argument $\underline{a}$ of category α, i.e. $\underline{C*}a$, the types of this phrase and of the duplicator coincide if β of the raised phrase is actually $(\alpha\to\beta)$. If, for instance, α is an NP, then it should be raised over a type like that of a transitive verb to produce the type-scheme of the duplicator $\underline{W}$. Given that $\underline{C*}$ is used in its general form in the grammar, it makes sense to divorce $\underline{W}$ from it. The type raising combinator is like a program module which can get additional specifications from lexical items. The duplicator is the semantic operation associated with reflexive pronouns, and its type is plugged into the raising module to produce the necessary syntactic category. Of course, this is only an example.

Given this analogy between type assignment, programs and intuitionistic logic, we can move on to the full intuitionistic type theory. Then types like disjoint union A+B, cartesian product A×B, recursion operator, etc. can be introduced. Some of these types seem plausible, some not. It might just turn out that the brain does not use all kinds of functions. Two reasons may immediately come to mind:

(a) This is due to the neuronal structures. In this case there is a special 'subintuitionistic' logic of the brain's syntax module, and studying this logic could shed some light on the mechanism.

(b) The brain does have access to the whole range of function types, the only limitation being some kind of 'memory overload'. This might be the case if, say, some combinators are not 'directly implemented', but are simulated by others.

The two alternatives merit some discussion, but the decision criteria are not very clear. Thus, it can be argued that the combinator K defined by $\lambda xy.x$ is not used by the brain because it destroys structure (this argument is attributed to Mark Steedman in Szabolcsi (forthcoming)). What is actually involved is whether a function should strictly depend on all its arguments. This is not so for K. Consider the way its type scheme is deduced:

$$\dfrac{\dfrac{[x \in \alpha]}{\lambda y.x \in (\beta\text{->}\alpha)}\ (\text{->}i)}{\lambda xy.x \in (\alpha\text{->}(\beta\text{->}\alpha))}\ (\text{->}i)$$

At the first deduction step the assumption $y \in \beta$ is vacuously cancelled, since it has not been used. This is permitted in the typed λ-K calculus, and has as a consequence that a function does not necessarily depend on all its arguments. So saying that it destroys structure is not quite accurate. In fact, the application of $\lambda xy.x$ to some argument a produces $\lambda y.a$, which is a nice description of 'lexical access function': if some 'activation' y is supplied, the result is a. So, in a way, K seems to be useful.

This was a suggestion that CCG or something similar can be used in the algorithmic level theories. Does it solve the correspondence problem? Yes, to a degree. If both the GB kind and the CCG kind of descriptions are more or less accurate, the information contained in some X-bar category of GB must be consistently reflected in some syntactic category of CCG. The converse is even more interesting, since it allows to put forward hypotheses of how much structure is due to the procedural aspects of syntactic knowledge.

REFERENCES

Chomsky, N. (1981) *Lectures on Government and Binding.* Foris, Dordrecht.

Curry, H. and Feys, R. (1958) *Combinatory Logic I.* North Holland, Amsterdam.

Hindley, J.R. (1985) Combinators and Lambda-Calculus, A short Outline, in G. Cousineau, P.L. Curien, B. Robinet, eds., *Combinators and Functional Programming Languages, Proceedings of the 13th Spring School of the LITP, Val d'Ajol.* Springer, Berlin.

Marr, D. (1982) *Vision.* Freeman, San Francisco.

Martin-Löf, P. (1975) An Intuitionistic Theory of Types: Predicative Part. In S. Rose and A. Sheperdson, eds., *Logic Colloquium '73.* North Holland, Amsterdam.

Martin-Löf, P. (1982) Constructive Mathematics and Computer Programming. In *Logic, Methodology, and Philosophy of Science VI.* North Holland, Amsterdam.

Szabolcsi, A. (forthcoming) Bound Variables in Syntax (Are there any?). In Groenendijk et al., eds., *Proceedings of the 6th Amsterdam Colloquium. Instituut voor Taal, Logica en Informatie, Universiteit van Amsterdam.*

ZENTRALINSTITUT FÜR SPRACHWISSENSCHAFT

BERLIN, GDR.

AARNE RANTA

PROSPECTS FOR TYPE-THEORETICAL SEMANTICS OF NATURAL LANGUAGE

1. Sentences, propositions, types

What is to be demanded of an "intelligent" computer is that it recognize and produce formal linguistic acts. To demand that it recognize linguistic acts proper would be to assume that a linguistic act can be identified by formal means; and to demand that it actually perform them would be to assume that a computer can be held responsible for its outputs.

My criterion of correctness for a theory of language is that the theory may be used in programming a computer in the way presented above. I think only one point is controversial here, namely the reference to computers. A theory of language is, quite uncontroversially, a system of rules - phonetic, syntactic, semantic, etc - ultimately purported to provide a systematic representation of the linguistic means by which certain acts can be performed. What the reference to computers amounts to is the demand that all these rules be executable.

I am indeed glad if someone realizes that this demand begs the question of choice between classical and constructive mathematical frameworks: the will to make linguistics suitable to computer implementations is indeed one of the strongest "eclectic" arguments in favor of a constructive system (cf. Beeson 1986, Martin-Löf 1982, Nordström 1986). Independent philosophical motivations, which I believe there are, fall outside the scope of this essay.

The concept form of a linguistic act is presumably an innovation of Martin-Löf's type theory. Assumptions and theorems in the theory are expressions which have one of the four forms of judgement (see Martin-Löf 1982, p.161, 1984, p.5):

A set

A=B

a:A

a=b:A

A judgement, then, is an act - an utterance or writing down, or whatever (when it is performed in order to communicate something to others it is called an assertion) of an expression of one of these forms; it is not a mere form but has as its essential components a person, a time, felicity conditions (conditions for the utterance to really count as a serious judgement). Once an act has been identified as a real act of one of the forms, we can check whether it is correct and what are the conventional consequences of performing it to the speaker and to the audience.

Judgements are typically explained by laying down their conditions of correctness. Conventional consequences are, ideally, determined by these conditions: what falls upon you when making a judgement does not exceed what you can support, warranted as you are to make it.

To make a warranted judgement 'A set' you must know what it is to be an element of the set A, and what it is for two presented elements of the set A to be equal. That is, you must know what set A is - sometimes by knowing a definition of it but not, of course, generally.

In most approaches to semantics, e.g. in Montague grammar, a proposition is explained by telling what its truth-condition is. This is, in fact, at the same time an explanation of the form of its most direct proofs. If proofs are considered not as processes but objects, i.e. constructions, a truth-condition can be considered as an explanation determining a set. Consequently, a proposition may be considered as a set, and its elements are its proofs.

What the truth-condition prescribes is the form of canonical proofs of the propositions. A construction of another form counts as a proof of that very proposition only if its existence guarantees the existence of a canonical proof. Then it can be treated as a non-canonical proof of that proposition. In a constructive approach like ours, that guarantee can only be that the suggested non-canonical proof is effectively calculable to canonical form.

The same distinction concerns all sets. A canonical element of the set of cats, or a canonical cat, is presumably just such an animal that we are accustomed to recognize immediately, and we are warranted in taking something else, say the name 'Garfield', as a non-canonical cat if and only if we are able to present a canonical cat corresponding to it.

(Digression:) Cat might be a good example of an undefinable set: we are warranted to take certain objects as canonical cats without reason. Thomas Aquinas pointed out that substantival terms are not definable as conjunctions of adjectival terms only, hence there must be irreducible substantival terms; and we, in order to observe anything at all, must observe some substances and not only qualities (see Geach & Anscombe, 1961, pp. 86-87). What we can teach a computer to do is thus always limited. Its "recognition" of substances consists in its having basic data types programmed in its memory - and this should be considered as good as possible a simulation of human capacities. What a theory of semantics can accomplish is hardly more than this: there will always be expressions whose meanings are left unexplained.

If you know that a is a canonical object of the set A or possess a method of computing it into one, you are warranted to make the judgement 'a:A'. If you consider A as a proposition, you might merely want to say

A true,

without mentioning what you know makes A true. This is, indeed, the type-theoretical explanation of assertive sentences or declarative or indicative (which word would be better not used, as it may cause confusion to a morphologically defined verb form).

An assertive sentence is thus explained as consisting of a proposition and a form of linguistic act. That form makes it an assertive sentence, and the proposition makes it that assertive sentence which it is. Another form would make it an interrogative sentence with different conditions of correct uttering, etc. But the conditions corresponding to each form can be treated as functions which operate on propositions, i.e. known truth conditions. This "double articulation" makes the theory simpler, in addition to being philosophically well-motivated. And in practice, to communicate with a machine, you must indicate what it should do with the content you express to it - do you expect it to adopt a new piece of data, to answer a question, or to cause a new state of affairs to obtain.

There being, presumably, a strongly limited amount of forms of linguistic act, the most extensive part of type theoretical semantics consists of explanations of propositions, whose number is unlimited. For a theory to be possible at all, it must explain propositions in terms of a finite class of primitive expressions, which it must, in the end, assume to be understood. The theory of propositions tells what expressions may combine to form other ones, and how the truth-conditions of propositions depend on expressions out of which they are formed.

Primitive and complex expressions, which are either propositions or can be used as parts of propositions, are each of some <u>type</u> or <u>category</u>. The category of an element of a set is just that set. We might say that the <u>meaning</u> of a proper name consists in its evaluating to a definite canonical element of a definite set, but this notion of meaning is not in any use in our theory: we do not have to make explicit appeal to it in explaining propositions.

The category of sets, or propositions, will be denoted by the letter <u>p</u>. It cannot be explained once and for all and is thus not a set, but each semantic theory in fact fixes a domain of propositions, which from the outside can be considered as a set. (This is not so deep as it may seem: a well-defined theory of course only concerns a well-defined domain, and what will ever be expressible in a natural language is not in this sense well-defined - it should not be assumed to be, anyway.) By analogy, we may thus replace the form of judgement '<u>A</u> set' by

A:p

where A, as a set-according-to-the-theory, is indeed, when considered from outside the theory, an element of the set p.

If an expression e can be combined with any expression x of the type t to yield an expression e_x of type s(x), we take e to be of the type (x:t)s. This explains the substitution scheme

$$\frac{a:t \qquad e\_x:(x:t)s(x)}{e\_a:s(a)}$$

We take _ to mean syntactic combination, which is most typically concatenation, but general account of which is a non-trivial matter. Its theoretical counterpart is functional application, simple like in 'f(x)' or by a special ap-operation, see Martin-Löf 1984, p.29. What is the main point here is that the categorial grammar is not based on a small finite amount of basic categories, like e.g. in David Lewis 1972. Where he has the category of truth-values containing just two objects we must have the infinitely large p, as we have no warrant that each proposition can be effectively evaluated to yield one of the two truth-values. And where he takes all functions to be defined for all objects, for us "the category of all objects" makes no sense at all. That is why, to give an example, the universal quantifier is not of some type o→p but of the type (x:p)((y:(z:x)p)p).

In addition to substitution, we shall only need two ways of construing complex propositions here:

(i) The Cartesian product of a family of types B(x) indexed on a type A, symbolized (Πx:A)B(x). Its canonical objects are functions (λx)b(x) where b(a):B(a) for each a:A. (To be exact, these are not functions proper but value-ranges of functions: the variable x in them is bound, and they can therefore be applied only by means of the application operator.)

(ii) The disjoint union of a family of types B(x) indexed on a type A, symbolized (Σx:A)B(x). Its canonical objects are pairs (a,b) where a:A and b:B(a). We shall also employ the projection operations here: p((a,b))=a and q((a,b))=b.

We shall not need the equality definitions in this essay. For the case (ii), to give an example, it would run as follows: (b,c)=(d,e):(Σx:A)B(x) if and only if b=d:A and c=e:B(b).

Constructions of these forms are transparent in the sense that logical inferences can be explained and performed by straightforward appeal to their forms. To make outputs in plain English possible, a generative syntax operating on these constructions is needed. One possibility is to build a generalized phrase structure grammar resembling that of Gazdar & al. 1985. One feature in each category would then indicate the semantic type of expressions in that category. (This might be of some help in treating agreement in Gazdar & al.'s way, as they take what agrees to what to be determined by semantic interpretations of expressions: for example, if something is interpreted as a functor, its agreement features are determined by those of the expression that is interpreted as its argument; see op. cit. p.84.)

That syntax is not independent of semantics, because it can by definition only generate meaningful sentences. The system of categories, which determines which combinations are grammatical, is motivated (although not, in the final theory, fully determined) semantically. There will hardly be any room for such sentences as 'Colorless green ideas sleep furiously'. This is due to what Chomsky 1965 would call subcategorization - the difference to Chomsky being that "subcategories" are given prior to larger categories.

Another sense in which syntax and semantics, or form and content, are here kept together is that "semantic interpretations" do not form a realm distinct from expressions: some expressions, namely primitive canonical ones, are assigned themselves as interpretations, whereas meanings of non-canonical and definable expressions are given in terms of irreducible expressions.

In the following section I shall illustrate how syntax and semantics are indeed intertwined in a way which seems to be totally absent in Chomsky but not foreign in Frege, and clearly present in Hintikka's game theoretical semantics.

2. Anaphora and Quantifiers

I shall in this section concentrate on interpretive semantics, that is, on how a computer can read plain English sentences and extract transparent propositions from them. The other part of interpretation would consist in extracting forms of linguistic act from English sentences - let us now assume it can be done in some way or another.

The set p of propositions and the set of interpretable English sentences share a common corpus of basic non-complex expressions. English is thus interpreted into English-with-transparent-modes-of-combination. I shall assume the following non-complex canonical expressions:

```
man:p
donkey:p
Bill:man
(x)(y)(x_owns_y) : (x:man)((y:donkey)p)
(x)(y)(x_beats_y) : (x:man)((y:donkey)p)
```

Following Hintikka & Kulas 1985 I use the device of contexts in dealing with anaphora. For certain kinds of steps in the analysis of a sentence - eventually a group of sentences combined to make one formal linguistic act - it is specified what individuals, together with their types, of course, are to be considered as given, or belonging to the context. If b:B is given, so are the following (given with b:B):

(i) Objects b':C which can be obtained from b:B, eventually in combination with other given judgements, by an effective operation'. For instance, if $a:(\Sigma x:A)B(x)$ is given, so are the projections $p(a):A$ and $q(a):B(p(a))$.

(ii) All objects c:C whose name c is used in B. For instance, if 'b:Bill owns a donkey.' is given, so is 'Bill:man.'

The first of our rules, and the only example concerning anaphora, is the rule (T.it) which is used in assigning truth conditions to sentences where the pronoun it occurs:

(T.it) X_it_Y

is to be analyzed as

X_b_Y

where X_c_Y is a proposition provided that c:A, and the context contains an up to equality-in-A unique object b:A.

Notice that the anaphoric relation is not treated as a syntactic one between words in the sentence, as is quite usual. It would be impossible to do so in the general case, as there need not be any noun for a pronoun to substitute. Another reason is illustrated by the example pair of sentences:

Take a gun and a knife but don't use it.
Take a gun or a knife but don't use it.

Only the latter one is grammatical, as only in it the context unambiguously fixes the reference of 'it'. The difference between 'and' and 'or' which causes this can hardly be uncircularly stated in purely syntactic terms. Consequently, there can be no correct semantics-independent syntax.

The treatment of quantifiers in Martin-Löf's Type Theory, and to some extent in most constructivist accounts, comes very close to their treatment in game-theoretical semantics. In particular, the conception of propositions as sets and the possibility to form the type of functions from any given type to another make it possible to interpret branching quantifier sentences (see Hintikka 1979, p.92, Ranta 1987).

(T.every) To analyze either of
(T.a)

(i) X_every A_Y
(ii) X_a A_Y

where we must have

A:p

and a uniform operation ' on A such that

X_b'_Y:p provided that b:A

first analyze A, i.e. check what it is to be an object b:A in terms of basic expressions. Second, given b:A, check what it is to be a proof c(b):X_b'_Y (where b' is, by definition, given with b:A). Canonical objects of

(i) are functions $(\lambda x)c(x)$ such that c(b):X_b'_Y provided that b:A;

(ii) are pairs (b,c) where b:A and c:X_b'_Y.

Hintikka deals with restricting relative clauses as parts of quantifier phrases, but I have chosen to take them separately, as constructive subsets (cf. Martin-Löf 1984, p.53):

(T.who) If

A:p

and

b'_B:p provided that b:A

then

A_who_B:p

and the objects of 'A who B' are pairs (b,c) where b:A and c:b'_B.

This rule has to be given priority over other ones in the order of application: the rightmost 'who' of the input sentence must first be completed in an interpretable way and then interpreted. No other ordering principles are needed here. Note that this rule is a source of multiple readings.

And then, at last, a traditional touchstone example:

(1) Every man who owns a donkey beats it.

The treatment of this example in Martin-Löf's type theory was originally suggested by Göran Sundholm in Sundholm 1986, p.502. What is new here is the introduction of explicit rules to yield the analysis. I hope it will be immediately seen how a whole lump of problems can be solved by the same theory.

By the priority of (T.who) we first pick up

(2) man who owns a donkey.

Following (T.every), we must then check what it is to be an object of (2). By (T.who) and (T.a) it is to be a pair (b,(d,c)) where b:man, d:donkey and c:b_owns_d. In other words, (2) is to be analyzed as

(3) (Σx:man)((Σy:donkey)x_owns_y)

which is assumed to be understood, the basic vocabulary being what it is.

Given b:(3), p(b):man and p(q(b)):donkey are given as well. The only possible way to complete b'_beats_it into a proposition, given b:(3), is to take it to be

(4) p(b)_beats_it.

By (T.it), (4) is the proposition

(5) p(b)_beats_p(q(b)).

Now we have analyzed (1) as the type of functions which to each b:(3) assign a proof of (5); i.e. the type

(6) (Πz:(Σx:man)((Σy:donkey)x_owns_y))p(z)_beats_p(q(z)).

3. Force

In a T-rule analysis no force component will be found. But given the notion of formal force (in the beginning of this essay) we may want the computer to take it into account. It might be marked simple-mindedly by a specific symbol like . or ? or !, or by the mood of the verb. What is essential is its being on a level other than propositions: force is not the outermost propositional form. (If it were so, we could form the implication of a command and a question, and so on, iterating as long as we please.)

In Martin-Löf's type theory there are forms of judgement only, and I have only considered two of them. In the first section I stated the correctness conditions governing these formal forces. The pre-theoretical intuition behind the rule for a:A, or A true, can be stated as follows:

(F.ass) P utters A true correctly only if P knows a proof of A.

Let us now consider the following forms of question:

? A

Wh(x:A)B(x)

A question is naturally explained by telling what counts as an answer. That is, the governing rules concern not the utterer but the audience:

(F.sentence-quest) Reply to ?A by uttering either A true or not-A true.

(F.wh-quest) Reply to $Wh(\underline{x}:\underline{A})\underline{B}(\underline{x})$ by uttering $\underline{B}(\underline{a})$ true for some $\underline{a}:\underline{A}$.

We must require a harmony between conditions of correct utterance and consequences of utterance (cf. Prawitz 1977, p.9). At least we must have some conception of harmony in order to make the machine work and spare it from infinite searching. The following conditions of correct utterance seem to be reasonable fair play rules:

(F'.sentence-quest) Utter $?\underline{A}$ only if you may utter '$\underline{A}$-or-not-$\underline{A}$ true'.

(F'.wh-quest) Utter $Wh(\underline{x}:\underline{A})\underline{B}(\underline{x})$ only if you may utter '$(\Sigma\underline{x}:\underline{A})\underline{B}(\underline{x})$ true'.

In our constructive system these formulations are strikingly simple. If we had a classical framework, the condition in (F'.sentence-quest) that $\underline{A}$ be decidable could not be expressed in the "object language". As for the second case, a constructive proof of $(\Sigma\underline{x}:\underline{A})\underline{B}(\underline{x})$ guarantees that only finite search for a right instance is needed (cf. Beeson 1986, Nordström 1986).

As for assertions, there is the consequence that anyone may believe $\underline{A}$ true, once someone has correctly asserted it. A computer uses its database, which contains expressions of the form $\underline{a}:\underline{A}$, as premisses in inferences and eventually as "given", or "frame of reference", in case there is no context in the sense of sec. 2 available.

ACKNOWLEDGEMENTS

to Jan von Plato and Per Martin-Löf for reading the manuscript and making valuable comments, and to Academy of Finland for financial support.

REFERENCES

BEESON Michael, Proving Programs and Programming Proofs, pp.51-82 in Barcan Marcus & al. (eds.), Logic, Methodology and Philosophy of Science VII, Elsevier Science Publishers, Amsterdam, 1986.

CHOMSKY Noam, Aspects of the Theory of Syntax, The M.I.T. Press, Cambridge (Ma.), 1965.

GAZDAR Gerald & KLEIN Ewan & PULLUM Geoffrey & SAG Ivan, Generalized Phrase Structure Grammar, Blackwell, Oxford, 1985.

GEACH Peter & ANSCOMBE Elizabeth, Three Philosophers, Basil Blackwell, Oxford, 1961.

HINTIKKA Jaakko, Quantifiers in Natural Languages: Some Logical Problems, pp. 81-117 in Saarinen (ed.), Game-Theoretical Semantics, Reidel, Dordrecht, 1979.

HINTIKKA Jaakko & KULAS Jack, Anaphora and Definite Descriptions, Reidel, Dordrecht, 1985.

LEWIS David, General Semantics, pp. 162-218 in Davidson & Harman (eds.), Semantics of Natural Language, Reidel, Dordrecht, 1972.

MARTIN-LÖF Per, Constructive Mathematics and Computer Programming, pp.153-175 in Cohen & Los & Pfeiffer & Podewski, Logic, Methodology and Philosophy of Science, North-Holland, Amsterdam, 1982.

MARTIN-LÖF Per, Intuitionistic Type Theory, Notes by Giovanni Sambin, Bibliopolis, Napoli, 1984.

NORDSTRÖM Bengt, Martin-Löf's Type Theory as a Programming Logic, Report 27 of Programming Methodology Group, University of Göteborg and Chalmers University of Technology, Göteborg, 1986.

PRAWITZ Dag, Meaning and Proofs: on the Conflict between Classical and Intuitionistic Logic, pp.2-40 in Theoria 48, 1977.

RANTA Aarne, Propositions as Games as Types, manuscript (submitted to Synthese), University of Stockholm, Stockholm, 1987.

SUNDHOLM Göran, Proof Theory and Meaning, pp.471-206 in Gabbay & Guenther (eds.), Handbook of Philosophical Logic, Vol. III, Reidel, Dordrecht, 1986.

DEPARTMENT OF PHILOSOPHY, UNIVERSITY OF STOCKHOLM
STOCKHOLM, SWEDEN

BEESON Michael, Proving Programs and Programming Proofs, pp. 51-82 in Barcan Marcus & al. (eds.), Logic, Methodology and Philosophy of Science VII, Elsevier Science Publishers, Amsterdam, 1986.

CHOMSKY Noam, Aspects of the Theory of Syntax, The M.I.T. Press, Cambridge (Ma.), 1965.

GAZDAR Gerald & KLEIN Ewan & PULLUM Geoffrey & SAG Ivan, Generalized Phrase Structure Grammar, Blackwell, Oxford, 1985.

GEACH Peter & ANSCOMBE Elizabeth, Three Philosophers, Basil Blackwell, Oxford, 1961.

HINTIKKA Jaakko, Quantifiers in Natural Languages: Some Logical Problems, pp. 81-117 in Saarinen (ed.), Game-Theoretical Semantics, Reidel, Dordrecht, 1979.

HINTIKKA Jaakko & KULAS Jack, Anaphora and Definite Descriptions, Reidel, Dordrecht, 1985.

LEWIS David, General Semantics, pp. 169-218 in Davidson & Harman (eds.), Semantics of Natural Language, Reidel, Dordrecht, 1972.

MARTIN-LÖF Per, Constructive Mathematics and Computer Programming, pp. 153-175 in Cohen & Los & Pfeiffer & Podewski (eds.), Logic, Methodology and Philosophy of Science VI, North-Holland, Amsterdam, 1982.

MARTIN-LÖF Per, Intuitionistic Type Theory, Notes by Giovanni Sambin, Bibliopolis, Napoli, 1984.

NORDSTRÖM Bengt, Martin-Löf's Type Theory as a Programming Logic, Report 27, of Programming Methodology Group, University of Göteborg and Chalmers University of Technology, Göteborg, 1984.

PRAWITZ Dag, Meaning and Proofs: on the Conflict between Classical and Intuitionistic Logic, pp. 2-40 in Theoria 43, 1977.

RANTA Aarne, Propositions as Games as Types, manuscript (submitted to Synthese), University of Stockholm, Stockholm, 1987.

SUNDHOLM Göran, Proof Theory and Meaning, pp. 471-506 in Gabbay & Guenthner (eds.), Handbook of Philosophical Logic, vol. III, Reidel, Dordrecht, 1986.

DEPARTMENT OF PHILOSOPHY, UNIVERSITY OF STOCKHOLM
STOCKHOLM, SWEDEN

JAROSLAV PEREGRIN

INTERSUBSTITUTIVITY SCALES AND INTENSION

1. Semantics in Logic and Semantics in Linguistics

Semantics is a branch of science which has obtained a rather peculiar position within the system of human science. Many people claim that they are doing semantics, but, curious enough, they are often people working in very different fields of theory and they also often do very different things under the heading of semantic theory. Even if we leave aside such branches of science as psychology and cognitive theory, which also more and more inquire into purely semantic problems, we have left two of the most important, namely logic and linguistics. Both of them study the semantic aspect of language; while linguistic semantics consists usually in comparison of expressions with respect to their meanings, logic focusses on the theory of truth and entailment.

However, we do not believe that what logic deals with is a part of semantics different from that which is investigated in frames of linguistic semantics. As we shall argue below, meaning and truth are two inseparably interlocked concepts and a theory of one of them is eo ipso a theory of the other. The difference between semantics as a part of linguistics and that which stems from logic is thus rather a matter of approach.

Metaphorically, if we compare the objective of the semantic theory with the enterprise of making a map of an unknown landscape, we can picture both these branches of science in their basic intentions as follows. The approach of a linguist resembles that of a man who walks across the land and pictures carefully various interesting details of the surrounding land onto a writing pad. He thus recognizes details of the character of the inquired land, can partially generalize them, but to assemble his pieces of knowledge into a real map might cause him problems. The logician, on the other side, resembles a man who has gone to inspect the land in an airplane. He is able to draw a generally appropriate, but maybe a rather rough and simplified map.

Hardly anyone doubts that a cooperation of the two mappers, i.e. a joining of the forces of logic and linguistics in the field of semantics, is desirable. What present an obstacle for such a cooperation are usually the different conceptual frameworks both branches of science have built. The aim of this paper is to contribute to the mutual understanding by a clarification of the concept which is crucial with respect to the connection of the theory of meaning with the theory of truth, namely the concept of synonymy.

2. Is the Concept of Synonymy Explicable?

The concept of synonymy plays undoubtedly a central role in the realm of semantics. On the other hand, it is far from being clear and it is under extensive discussion. One of the most outstanding figures of the philosophy of language of the twentieth century, W.V.O.Quine, for example, strictly rejects the concept for its claimed essential unclarity and consequent uselessness for the sake of foundation of semantic theory. However, synonymy is in one or another way deeply rooted in our intuition and as the task of theory is to explicate such basic pre-analytic concepts, it seems to be worthwhile to try to inquire into it.

What is synonymy? Intuitively, we think of it as of the sameness of meaning. But to *explicate* the concept in this way we would have to have a clear explication of the concept of meaning (of course not based on the concept of synonymy), and this, as far as I can see, we have not. We have to explicate the concept of meaning via the concept of synonymy, and hence to explicate the concept of synonymy *without* the help of the concept of meaning. Is this possible?

A traditionally widely accepted explication of the concept is the identification of synonymy with the overall intersubstitutivity *salva veritate*; this means that two expressions are understood as synonymous just in case they are mutually interchangeable in any sentence without causing the sentence to change its truth value.[1] We agree with Quine that there is no other possibility than to explicate synonymy this way; however, we do not agree that even this explication is untenable.

Generally, there are two possible objections to this explication. The first of them, Quine's own, is that the notion of interchangeability *salva veritate* is dependent on the concept of analycity, which is not clear, and therefore the whole *explicatum* lacks the necessary degree of clarity. The second objection is that although such explication is possible, it leads to such a notion of synonymy which is in fact vacuous as for any two expressions there exists a sentence in which their interchangeability *salva veritate* is at least dubious. Let us consider the two objections.

According to Quine(1960) [2], the identification of synonymy with intersubstitutivity may be meaningful only thanks to the existence of intensional contexts in language. This is to say that the fact that, for example, *bachelor* and *unmarried man* are mutually interchangeable in extensional contexts, i.e. that (1) holds, does not assure the synonymy of the two terms (for if it did, also, for example, *American president* would be synonymous with *white man*). To guarantee the synonymy the interchangeability in intensional contexts, i.e. validity of (2), is needed.

> _Every bachelor is an unmarried man and every unmarried man is a bachelor_ (1)

> _Every bachelor is necessarily an unmarried man and every unmarried man is necessarily a bachelor_ (2)

However, Quine states, to be able to decide whether such a sentence as (2) is true, we would have to be able to decide whether (1) is analytic. And, as Quine shows persuasively, the concept of analyticity is not clear; he states that it is surely at least problematic to decide whether sentences of the kind of (3) are analytic.

> _Everything green is extended_ (3)

My opinion is that while the whole enterprise of semantic theory consists in fact in an explication of some basic intuitive notions, we must not ask for more than such an intuition can give us. This is to say if it is not clear if such a sentence as (4) is true, i.e. if such a decision lies outside of our intuition, then it is not reasonable to treat this as a failure of intuition, as a _hic sunt leones_ on a map of language, but we have rather to take it as a primary fact which takes part in the semantic specification of language.

> _Everything green is necessarily extended_ (4)

As the concept of synonymy exists (and seems to be basic) within our intuition, it should be explicable by means of what the intuition alone offers. To declare this intuition as imperfect means not to try to explicate it, but to try to squeeze it into an unwarrantedly preconceived mould. That is to say if it is not possible to decide if (4) is true, it is necessary to exclude the sentence from our considerations of synonymy: indeed, if the sentence lacks a truth value then it surely also cannot change a truth value and hence is trivially consistent with any notion of synonymy of our kind.

The other objection, namely that under the identification of synonymy and intersubstitutivity _salva veritate_ no two expressions will be synonymous, seems to be irrefutable. No two different expressions are generally interchangeable in quotational contexts; for example _unmarried man_ is surely not interchangeable with _bachelor_ in (5).

> _Bachelor has eight letters_ (5)

This fact has to be investigated in detail.

3. The idea of semantic space

Consider the word dog and the words animal and mammal. Which of the latter is semantically closer to the former? Intuitively, such question is meaningful and we can answer that it is mammal which is closer to dog. Moreover, we are able to make sense of such an answer even in frames of the accepted explication of synonymy: mammal is semantically closer to dog than animal as in any sentence in which we can substitute animal for dog we can also substitute mammal for it *salva veritate* (but not vice versa).[3]

This suggests the idea of a "semantic space", i.e. of a certain relation of semantic closeness between individual language expressions of the same syntactic category. The "distance" between two expressions seems to be in some way proportional to the number of contexts in which the two expressions are not interchangeable *salva veritate*. (It is not difficult to see that the "distance" conceived in such a way is really a distance, i.e. that it fulfils the characteristics of metrics as usually defined in mathematics.[4]) If this idea is right, then the semantics of natural language can be viewed upon as certain spatialization of expressions.

But even if we are able to speak about semantic closeness (or distance) in a clear sense, does it eliminate the objection that there is in fact no synonymy? Let us introduce the answer to this question by an illustrative example from everyday life. We often talk about sameness of height of two persons, but does such a sameness in fact exist? Surely the sameness is at most the indiscernibility regarding available measuring techniques. The sameness is in this case (and in any other case of this kind) a closeness to certain relevant degree. It seems not to be different with our concept of synonymy; it may be viewed as semantic closeness relative to some degree of "discernibility".

What we have said means that two expressions are considered as synonymous just in case their distance in the semantic space does not exceed a relevant limit of discernibility. But this immediately seems problematic, as under such definition synonymy is not necessarily transitive, i.e. if $expression_1$ is synonymous with $expression_2$ and $expression_2$ with $expression_3$, then $expression_1$ need not be synonymous with $expression_3$. This is the problem of all cases of sameness based on indiscernibility (as the sameness of height mentioned above is) and it shows that such a notion has only a limited field of application. Especially in our case of synonymy, as we would want to base the concept of meaning on it, we clearly need an equivalence, i.e. a relation which is transitive.

Let us thus return to the statement that under our definition the resulting synonymy relation need not be transitive. It need not, but is it really not? More properly formulated, does it hold for any given limit of discernibility that the resulting relation is not transitive? There, of course, exist two trivial cases, in which we gain a relation which *is* transitive, namely the limit being less than the minimal distance of any two expressions and it being greater than the distance of any two expressions. However, the trivial cases are not interesting; the former results in treating no two expressions as synonymous, the latter in treating any two expressions as synonymous. What interests us is whether there exist other, nontrivial cases.

What does it mean for such a nontrivial case to exist? Formally its existence means that there exists a limit d such that expressions can be divided into pairwise disjoint classes such that the distance of any two expressions of the same class is less than d, while the distance of any two expressions which are not of the same class is greater than d. Moreover, in order for the classification to be nontrivial, the resulting classes have to be at least two and at least one of them has to contain more than one member. This again in fact means the existence of "clusters" in the semantic space, i.e. it means that the space looks more like Figure 2 than like Figure 1. Now the question is: do such "clusters" really exist?

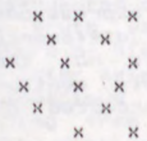


Figure 1

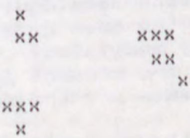


Figure 2

4. The Idea of an Intersubstitutivity Scale

Let us imagine that we manage to classify sentences of natural language into N pairwise disjoint groups with the property that, for $i=1,\dots,N-1$, if two expressions are interchangeable within all sentences of the ith group, then they are also interchangeable within the $(i+1)$th group. Then we can for any two expressions find the smallest i such that the two expressions are interchangeable within the $(i+1)$th group (if they are not interchangeable within any of the groups we can take $i=N$). We can then take such a value as the semantic distance of the two expressions in the sense of the previous paragraph. The relation between expressions "to be not farther from one another in the semantic space than i" is then for any $i=1,\ldots,N$ transitive and thus it is clearly an equivalence.

What I want to suggest is that an intersubstitutivity scale exists for natural language, that it gives rise to several levels of relative synonymy and that just this fact gives the concept of synonymy a real sense. The scale I am going to present consists of four levels and it is in fact the same scale which has been in a slightly different context recognized by Sgall et al. (1986, Chap.1.5). The first level of the scale gives rise to the absolute (and thus trivial) equivalence while the others lead to equivalences corresponding to a declining limit of discernibility. In the present contribution we do not enter a discussion of the problem whether this scale could be refined in a way.

We have claimed to base our concept of meaning on the concept of synonymy; hence by the acceptance of the various levels of synonymy we also necessarily accept various levels of meaning. A notion of synonymy simply implies a notion of meaning as synonymy-equivalence classes of expressions (if we are nominalists), or as some abstract objects each being common to one of the equivalence classes (if we are realists or in one or another way accept abstract objects of such kind). What is important to realize is that in the latter case there may exist quite different notions of meaning, which, however, correspond to the same equivalence classes of expressions and thus to the same notion of synonymy. A difference between such equivalent concepts of meaning is not the objective of the present paper; we shall simply mention the most common and the most traditionally accepted of them.

Now we shall present the scale we have mentioned; however, we of course cannot present a definite and exhaustive list of sentences (contexts) which constitute the individual classes of the scale. Fortunately, prototypical contexts of each of the classes can be picked up the intersubstitutivity in which indicates the overall intersubstitutivity in contexts of the whole class. We shall therefore specify the classes by means of such prototypical contexts.

We shall also restrict ourselves to the meaning and synonymy of sentences.

5. Coextensionality and Cointensionality

The weakest level of equivalence we introduce can be called *coextensionality*. It can be considered as the interchangeability *salva veritate* in the prototypical context

It is true that ... (8)

or, which is clearly the same, in the empty context. The interchangeability in the context (8) guarantees interchangeability in contexts of the kind of those which are schematicized by the logical operators of the first-order logic.

Coextensionality yields only two categories of sentences, namely true sentences and those which are not true. This also indicates that this relation can change very rapidly with time (e.g., such two sentences as *Now it is morning in Prague* and *Prague is the capital of Czechoslovakia* are coextensional in every morning and are not in any other time of the day). The meaning of a sentence based on the *synonymy=coextensionality* assumption can be identified with its truth value, and the consequence of the variability of coextensionality is, of course, the corresponding variability of meaning.

At this point someone can wonder why we think of the explication of synonymy as coextensionality at all, as it would be probably hard to find anyone who would be prepared to accept that synonymy could have such properties as coextensionality has (namely the rapid variability). The reason is a historical one. The identification of meaning of sentence with its truth value is inherent to the logical system of G.Frege, who can truthfully be considered as the founder of modern logic, so that his ideas have surely be influential. Let us, however, note that Frege was aware of the absurd consequences of identification of meaning of sentence with its truth value; it is more appropriate to see in this feature of his framework (as well as of the frameworks of the later extensionalists, such as Church or Quine) rather a conscious deviation of his system-internal concept of meaning from the intuitive one than a naive attempt to wholly explicate the intuitive concept in this way.[5]

The second level of equivalence, called *cointensionality* consists in the interchangeability in the contexts of the kind

Under such and such circumstances it would be true that ... (9)

This means that two sentences are cointensional if and only if one of them is true under some circumstances just in case the other is true. Traditionally the context created by means of the adverb *necessarily* is being mentioned in this regard, but the concept of necessity is ambiguous and thus we have substituted the validity-in-any-circumstances for it to achieve an unambiguous sense. It is clear that such two sentences as the two mentioned above, namely *Now it is morning in Prague* and *Prague is the capital of Czechoslovakia*, are certainly not cointensional.

The assumption *synonymy=cointensionality* leads to the notion of meaning that can be modelled as a function from some possible states-of-affairs, or possible worlds, to truth values. This approach introduced by R.Carnap underlies the formal semantic systems of Montague(1974), Cresswell(1973), Tichy(1978) and others. It has not such strikingly absurd consequences as the extensionalists' approach and it thus serves as a good basis for an analysis of various kinds of problems of natural language semantics. However, there are again examples of sentences which are cointensional, but not interchangeable in all possible contexts. Namely, for example, all truths of mathematics are cointensional (as they are true regarding any possible state-of-affairs), but they are surely not universally interchangeable and hence not absolutely synonymous.

6. Intentional Attitudes

The third level of our scale can be characterized by the interchangeability in contexts of the kind

Someone believes that ... (10)

There is a lot of discussion about these so-called belief-contexts in contemporary semantic literature. One of the central problems concerning this kind of sentences is whether a sentence as X believes that S is to be understood as X believes that the sentence 'S' is true. If the answer were positive, then belief-contexts would be a sort of quotational contexts and this level of equivalence would coincide with the fourth level, which means interchangeability in quotational contexts and hence universal interchangeability and absolute synonymy. But, together with Bigelow(1978a), we are convinced that the answer is not positive. For if it were, it would not be possible to speak in English about the beliefs of someone who does not know English, which surely is not the case. On the other side such sentences as mathematical laws are surely not equivalent on this level.[6]

In contrast to the previous levels, there is no common formulation of the notion of meaning which arises from the unification of synonymy with this kind of equivalence. The entire problem was recognized already by Carnap; he proposed that two sentences are interchangeable in belief-contexts just in case they have the same intensional structure, i.e. if they are constituted by the same number of elementary expressions and if their corresponding constituents are cointensional. This approach has been embodied in the logical theory of Lewis(1972) and later further elaborated by Bigelow(1978a;b). The resulting notion of meaning can be identified with a (finite) sequence of intensions. But is this notion of meaning really appropriate to this level of synonymy? It might seem that it has to be because in frames of intensional logic there seems to be no alternative. But this is, of course, a form of dogmatism and not an argument.

Let us consider (11) and (12).

<u>It is difficult to please John</u> (11)
<u>John is difficult to please</u> (12)

It seems more than clear that they are interchangeable in belief-contexts; nevertheless their intensional structures differ. It thus seems that the concept of intensional isomorphism relies too heavily on the surface structure of the sentence.

In the beginning of this paper we have mentioned the different ways logicians and linguists approach semantics of natural language. Then we have, in fact, followed the way of logicians; we have started from the truth value of the sentence and proceeded in the direction to its surface shape. The way of linguists have been left aside so far; and, fortunately, at this point, when the logicians' approach starts to be unsatisfactory, we can make use of the achievements of linguists, who make their way from the opposite side to meet us.

The problem linguistic semantics is usually interested in is just the fact that sentences of different surface shapes are semantically very close. Sgall et al. (1986), reassuming considerations of de Saussure and Hjelmslev, assign to any sentence a <u>tectogrammatical representation</u> which can be viewed upon as a representation of the, essential semantico-pragmatic structure of the sentence. It differs from the above mentioned intensional structure in that it does not take the surface shape of sentence at face value and tries to handle such aspects of the surface structure which do not constitute substantial semantic characteristics of the sentence.

Any tectogrammatical representation can be assigned a sentence which may be viewed upon as a "canonical expression" of the representation. However, not any sentence is such a canonical expression of its tectogrammatical representation; in the above example, (11) can be considered as the canonical expression of the tectogrammatical structure which belongs to both (11) and (12). The construction of tectogrammatical representation thus can be considered as a transfer of sentence to certain standard or normal form (i.e. as Quine's, 1960, regimentation); the standard forms may be identified with certain sentences of natural language. Thus we could, for the sake of mutual understanding between logicians and linguists, see the difference between Lewis' (1972) meanings and Sgall's et al. (1986) tectogrammatical representations not in their completely different nature, but in the fact that, while the former rest on the intensional structure of the sentence in question, the latter turns rather to the intensional structure of a standard form of the sentence, i.e. to the intensional structure which does not necessarily belong directly to the sentence in question, but to a sentence the synonymy of which with the sentence in question is established on empirical linguistic grounds. (11) and (12) thus are assigned the same intensional structure, namely that of (11).

Let us call the relation of sameness of tectogrammatical representation costructurality.

The remaining, fourth level of equivalence is the absolute synonymy implied by interchangeability within quotational contexts. There is no more to be said of this level besides that this kind of equivalence implies identity of the surface structure and is thus trivial.

7. Discussion

We have proposed four levels of equivalence of expressions each of which can be viewed upon as a kind of synonymy relative to certain level of "semantic discernibility". It might be still objected that such a relative synonymy is in fact no synonymy and any synonymy worth its name is the absolute one and that hence the objection that the concept of synonymy is empty is thus still in force. However, returning to our illustrative example of the notion of sameness of height of two persons we have to accept that if two persons differ as for their height only to the degree which is not detectable without some measuring apparatus, we consider them of the same height in the full-fledged sense of the word same. It thus seems that to be able to consider relative synonymy as a synonymy in the full-fledged sense of the term we have to show that the corresponding level of discernibility of meanings is in some sense distinguished. That is to say to show that to consider, for example, costructurality as synonymy we have to show that the distinction between interchangeability in quotational contexts and interchangeability in other contexts is in some sense substantial and that it is sometimes only the latter which is relevant for our semantic considerations.

However, such a distinction is recognized as substantial by nearly all logicians as well as linguists; the recognition has caused even the terminological differentiation of using a word (i.e.displaying it in non-quotational context) and mentioning that word (i.e.displaying it in quotational context).[8] Quotational contexts are needed only when we have to speak about language and thus in the majority of fields of human linguistic activity we do well without them and thus are very often interested only in non-quotational contexts.

The situation is similar with respect to the distinction between the contexts of intentional attitudes and other contexts which is relevant for the significance of cointensionality. The contexts of intentional attitudes are needed in psychology and in some other fields, but there are surely numerous ranges of problems which we may well treat without any need of them. To consider only contexts which are at most intensional is thus often sufficient; hence the level of discernibility corresponding to cointensionality is not insignificant.

With respect to the distinction between intensional and extensional contexts the situation is not so clear, althought many philosophers from Wittgenstein to Quine do consider this distinction as important.

The notion of synonymy as presented above thus seems to be plausible; indeed, many logical systems based on coextensionality (first-order logic), cointensionality (modal and intensional logics) and certain kinds of costructurality (various "hyperintensional" logics as those of Bigelow, 1978a,b) have proven themselves as of great practical value. Moreover, and this I consider as the decisive reason for the acceptance of the above explication of the concept of synonymy, this notion of sameness of meaning is in accordance with our intuitive usage of the concept. For if there were no synonymy except for the absolute and thus vacuous one, how it would be possible for the concept to get at the central place of our semantic conceptual framework?

8. Conclusions

We have tried to explicate the concept of synonymy, but from what has been said in the course of the explication some ideas can be drawn the import of which widely exceed the problem of synonymy. It was the relation of logic to linguistics that we questioned in the beginning of this paper and this is what we would now like to address.

The sense of formal logic depends on the capability of corresponding formal language to schematize natural language. However, it seems that it is possible to divide such a schematization into two steps: (i) the "regimentation" of the sentence in question, i.e. finding of its "normal form" in the sense of Section 6; and (ii) the logical analysis of the "normalized" sentence (which can be, of course, presented by means of a pre-logical schematization, such as tectogrammaticalrepresentation).

The suggested division is corroborated by various facts. Let us take, for example, the principle of compositionality, which appears to be the cornerstone of logic and formal semantics since Frege. Hausser(1984) claims that the principle can hardly be conceived as something other than <u>surface</u> compositionality; Seuren(1985), on the other side, tries to show that it is plausible to relate it to the level of deep structures. Now we are convinced that although Hausser is undoubtedly right, also Seuren's arguments are acceptable; namely, it seems to be more plausibe to restrict the direct application of the principle of compositionality only to the "normal" sentences, which is fully in accordance with the claim that only these sentences constitute the direct subiect of logical theory.

If this whole idea is right, then the division of semantic labour between logic and linguistics may be considered as coinciding with the division of semantic analysis into the above mentioned two steps: (i) linguists are to study the "normal" form of expressions (to avoid an oversimplified understanding of the task let us add that the frequent ambiguity of natural language expressions causes that an expression often corresponds to several distinct normal forms); and (ii) logicians are to analyze the normalized sentences.[9]

NOTES

1. This is what makes truth and meaning two sides of the same coin.
2. Cf. also Quine(1963) and Gibson(1982).
3. With some trivial exceptions such as *Dog begins with D or with A*.
4. To see this, it is necessary and sufficient to realize that (i) the distance of an expression from itself is zero; (ii) the distance of an expression *E1* from an expression *E2* equals the distance of *E2* from *E1*; and (iii) the distance of an expression *E1* from an expression *E2* plus the distance of *E2* from an expression *E3* is not smaller than the distance of *E1* from *E3*.
5. What Frege was interested in above all are mathematical judgements, and these are *standing* sentences (in Quine's terms), i.e. they have a definite truth value once for all. Moreover, the recognition of the fact that some coextensional sentences are clearly not intuitively synonymous had led him to the introduction of another level of meaning called *Sinn*, in contrast to *Bedeutung*, which, regarding sentences, coincides with their truth value.
6. Cf. also Peregrin(1987).
7. The authors are, of course, not the only linguists who try to formalize ideas of this kind. Various kinds of "deep structures" of Chomskyan linguistics clearly aim at a more or less similar goal.
8. See Sgall et al.(1986, Chap.1.1).
9. An attempt at an explicit although fragmentary linkage between tectogrammatical representations and formulas of intensional logic has been presented by Peregrin and Sgall(1987). For the logical analyses of linguistic problems connected with tectogrammatical representation see Materna and Sgall(1980, 1984), Materna et al.(1987) and Peregrin(prep.).

REFERENCES

Bigelow, J.C.(1978a): Semantics of Thinking, Speaking and Translating, in Meaning and Translation (ed. by F.Guenthner and M.Guenthner-Reutter), Duckworth, London.

Bigelow, J.C.(1978b): Believing in Semantics, Linguistics and Philosophy 2, 101-144.

Cresswell, M.J.(1973): Logic and Languages, Meuthen, London.

Gibson, R.F.,Jr.(1982): The Philosophy of W.V.Quine, Univ. Presses of Florida, Tampa/St.Petersburg/Sarasota.

Hausser, R.R.(1984): Surface Compositional Grammar, W.Fink Verlag, Munich.

Materna, P. and Sgall, P.(1980): Functional Sentence Perspective, the Question Test and Intensional Semantics, SMIL 1-2, 141-160.

Materna, P. and Sgall, P.(1984): Optional Participants in a Semantic Interpretation, in Contributions to Functional Syntax, Semantics and Language Comprehension (ed. by P.Sgall), Academia, Prague.

Materna, P., Hajicova, E. and Sgall, P.(1987): Redundant Answers and Topic-Focus Articulation, Linguistics and Philosophy 10, 111-114.

Montague, R.(1974): Formal Philosophy: Selected Papers of R.M. (ed. by R.Thomason), Yale Univ.Press.

Peregrin, J. and Sgall, P.(1987): An Attempt at a Framework for Semantic Interpretation of Natural Language, to appear in Theoretical Linguistics.

Peregrin, J.(1987): A Contribution to the Theory of Propositional Attitudes, Prague Bulletin of Mathematical Linguistics 48, 13-36.

Peregrin, J.(in prep.): Quantification and Reference in Natural Language, submitted to Studies in Logic, Grammar and Rhetorics.

Quine, W.V.O.(1960): Word and Object, MIT Press, Cambridge, Massachusetts.

Quine, W.V.O.(1963): Two Dogmas of Empiricism, in From a Logical Point of View, Harper and Row, New York.

Seuren, P.M.(1985): Discourse Semantics, Blackwell, Oxford.

Sgall, P., Hajicova, E., Panevova, J.(1986): The Meaning of the Sentence in its Semantic and Pragmatic Aspects, Academia, Prague.

Tichy, P.(1978): De Dicto and De Re, Philosophia 8, 1-16.

ORGATEX

HRADEC KRALOVE, CZECHOSLOVAKIA

REFERENCES

Bigelow, J.C.(19[illegible]): [illegible] Thinking, [illegible] Translation, [illegible] Duckworth, London.

Bigelow, J.C.[illegible]

Cresswell, M.J.[illegible]

Gibson, R.F., Jr.(1982): The Philosophy of W.V.Quine, University Presses of Florida, Tampa/St.Petersburg/Sarasota.

Hausser, R.R.(1984): Surface Compositional Grammar, W.Fink Verlag, Munich.

Materna, P. and Sgall, P.(1980): Functional Sentence Perspective, the Question Test and Intensional Semantics, SMIL 1-2, 141-160.

Materna, P. and Sgall, P.(1984): Optional Participants in a Semantic Interpretation [illegible] Functional [illegible] (ed. by P. Sgall), Academia, Prague.

Materna, P., [illegible] and Sgall, P.[illegible] Linguistics and Philosophy 10, [illegible]

Montague, R.(1974): [illegible]

Partee, B.H.[illegible] Semantic [illegible] Theoretical Linguistics.

Peregrin, J.[illegible] Theoretical Linguistics [illegible]

Peregrin, J.[illegible] Grammar and [illegible]

Quine, W.V.O.(1960): Word and Object, MIT Press, Cambridge (Mass.).

Quine, W.V.O.[illegible]

Quine, W.V.O.[illegible] Logical Point of View, Harper and Row, New York.

Searle, J.R.(1969): [illegible]

Sgall, P., Hajičová, E. and Panevová, J.(1986): The Meaning of the Sentence in its Semantic and Pragmatic Aspects, Academia, Prague.

Tichý, P.(1988): The Foundations of Frege's Logic, de Gruyter, Berlin.

ANNA MADARÁSZ

SEMANTIC GAMES AND SEMANTIC VALUE GAPS

It is an inveterated idea of mine that the semantical framework of classical logic, the so-called Frege–Tarski–Montague line of truth-condition semantics should be combined with new theories which use a broader concept of meaning and at the same time can provide logically correct methods to handle some interesting features of natural language which are inaccessible to a classical analysis. I sketched such an approach in two different ways:

(1) For the treatment of pragmatical problems I compared the standard Montagovian methods with those of Game Theoretical Semantics (GTS for short) in order to provide a more suitable semantical framework. (See e.g. [9].)

(2) My latest paper compares GTS with a semantics with semantic value gaps ('VG semantics' for short) and outlines a way to combine these two approaches. (See [10].)

Throughout these efforts I was led by the conviction that the standard logical semantics based on Fregean principles is on the one hand an indispensable basis for any analysis of a logical sort though on the other hand in itself it is insufficient for a proper analysis of the logical forms of natural language. This is why we need the new approaches, the rival theories, which, however, cannot be absolutely independent from classical, i.e. truth-conditional, semantics. My conviction, it other words, is that there is (or at least there can be) a coherent account of all the logical approaches.

In this paper I would like to illustrate some of my ideas how the basic principles of GTS and VG semantics can be brought into accordance and I might show some of the advantages such a combined system may have in modelling natural language. I will argue that in the rules of the dialogue games there should be a third possibility, when instead of winning or losing, a semantic value gap occurs, because dialogue games provide a new source for value gaps.

1

One of the motivations for the introduction of semantic value gaps is the fact that there are some unsolvable problems in the possible world semantics of modal logic. The variability of the domain of quantification from world to world has a very shocking effect in a Kripke-style system — the converse of the Barcan formula, the classic tautology "$\forall x.Fx \supset Fa$" will be refutable and descriptions will have to be expelled from the system. These problems just simply disappear when we introduce semantic value gaps. When a term t denotes an object u which is not a member of the domain of quantification of the world w, then we say that t has no denotation *in w.* This case is what we call a semantic value gap. This value gap may go over from the term to the formula containing this term, and a formula also may have no extension (i.e. it may be without a proper truth value). To generalize: we speak about a value gap when a well-formed extensional phrase which may belong to any syntactical category whatsoever, does not have an extension in a certain possible world.

There is no semantic value gap in the case of intensions. In possible world semantics, the intension of a phrase of type α is a function mapping possible worlds into the domain of extensions of type α. This function may be undefined for some worlds as we introduced semantic value gaps, and the function, giving the intension may be a partial one on the set W of possible worlds. All this means that value gap is not conserved in the case of modal formulas: even if A is without truth value in some worlds, "$\Diamond A$" is always true or false.

As in GTS, Hintikka's games are played mostly in a first-order context, it is feasible to show a first-order variant of VG semantics[*] — as elaborated in our Department in Budapest, by I. Ruzsa.

First-order logic with VG semantics. Let us assume a usual first-order language extended by two new logical symbols: 'I' (the descriptor) and 'NON' (a monadic sentence operator). The grammatical rules associated to the new symbols are as follows:

If x is a variable and A is a formula, "$\mathrm{I}x.A$" is a term.

If A is a formula, so is "NON(A)".

By an *interpretation* of the extended language let us mean a pair $\langle U,\rho\rangle$ where U is a nonvoid set of individuals and ρ is a function giving the factual value (extension) of the nonlogical constants of the language. In what follows, $|A|_v$ denotes the factual value of a phrase A, according to the assignment v of the variables. I shall use "$v[x{:}u]$" for denoting the assignment which is the same as v except (at most) that $v[x{:}u](x) = u$, where x is a variable, and $u \in U$.

If t is a term, the identity "$|t|_v = \Theta$" will mean that t is without any denotation, and if A is a formula, the identity "$|A|_v = 2$" will mean that A is without any truth value (according to the assignment v). Now the rules for calculating $|A|_v$ are as follows:

(i) If x is a variable, $|x|_v = v(x)$.

(ii) If c is a constant term, $|c|_v = \rho(c)$.

(iii) If p is a sentence parameter, $|p|_v = \rho(p) \in \{0,1,2\}$ where 0 and 1 stand for the usual truth values and 2 represents the truth value gap.

(iv) If P is an n-place predicate, $t_1, \ldots, t_n$ are terms $(n \geq 1)$, then

$$|P(t_1)\ldots(t_n)|_v = \begin{cases} 2 & \text{if for some } i,\ |t_i|_v = \Theta\ (1 \leq i \leq n), \\ 1 & \text{if } \langle |t_1|_v, \ldots, |t_n|_v \rangle \in \rho(P), \\ 0 & \text{otherwise.} \end{cases}$$

---

[*] This semantics is similar to that of Smiley. See [13].

(v) If t_1, t_2 are terms, then

$$|(t_1 = t_2)|_v = \begin{cases} 2 & \text{if one of } |t_1|_v, |t_2|_v \text{ is } \Theta, \\ 1 & \text{if } |t_1|_v = |t_2|_v \neq \Theta, \\ 0 & \text{otherwise.} \end{cases}$$

(vi) If A is a formula, x is a variable, and there is a unique $u_0 \in U$ such that $|A|_{v[x:u_0]} = 1$, then $|\text{I}x.A|_v = u_0$, in other cases $|\text{I}x.A|_v = \Theta$.

(vii) If A is a formula, $|{\sim}A|_v = \begin{cases} 2, & \text{if } |A|_v = 2, \\ 1-|A|_v & \text{otherwise.} \end{cases}$

(viii) If A, B are formulas, $|(A \supset B)|_v = \begin{cases} 2 & \text{if } |A|_v = 2 \text{ or } |B|_v = 2, \\ 0 & \text{if } |A|_v = 1 \text{ and } |B|_v = 0, \\ 1 & \text{otherwise.} \end{cases}$

(ix) If A is a formula and x is a variable, then

$$|\forall x.A|_v = \begin{cases} 2 & \text{if for all } u \in U, |A|_{v[x:u]} = 2, \\ 0 & \text{if for some } u \in U, |A|_{v[x:u]} = 0, \\ 1 & \text{otherwise.} \end{cases}$$

(x) If A is a formula, then

$$|\text{NON}(A)|_v = \begin{cases} 0 & \text{if } |A|_v = 1, \\ 1 & \text{otherwise (i.e., if } |A|_v \text{ is 0 or 2).} \end{cases}$$

Comments. As one sees, rules (iv) to (ix) guarantee that value gaps occur conservatively, i.e., if a component of a phrase does not have a semantic value then the complex will not have one either. The only exception is our new operator 'NON'. We need this new logical constant in order to be able to express statements of form "it is not the case that p". The symbol '$\sim$' is not suitable here, for, if p contains a non-referring term, "$\sim p$" will be without a truth value either, whereas "NON(p)" will be true. An example: Suppose Peter has not a wife. Then the sentence 'Peter's wife is blond' has no truth value, and hence,

$$|\sim \text{(Peter's wife is blond)}| = 2.$$

On the other hand,

$$|\text{NON(Peter's wife is blond)}| = 1,$$

and $$|\text{NON(Peter's wife is not blond)}| = 1.$$

We can introduce the constant 'VER' by the contextual definition

$$\text{VER}(p) =_{df} \sim\text{NON}(p).$$

That is, "VER(p)" represents the statement "it is true that p". Further,

$$\text{NON}(p) \ \& \ \text{NON}(\sim p)$$

means that p lacks a truth value, i.e., $|p| = 2$.

Let us note that in our first-order semantics we use two different value gaps — one for sentences and the other one for terms. We get a new source of value gaps by admitting *partial functions* as semantic values of predicates, that is, functions such that for some objects they yield the value 'true', for other ones the value 'false' but for some they may have no value at all. E.g., the sentences

Budapest plays the piano,

Budapest does not play the piano

are without any truth value, as the predicate 'plays the piano' is not defined for (inapplicable to) Budapest. As soon as we apply the functor 'NON' for these sentences, we shall get true sentences.

Concerning the central notions of logical semantics, let us note that in VG semantics, *irrefutability* and *validity* is to be distinguished. Here a sentence is said to be irrefutable if it is false in no interpretations and a sentence is valid if it is true in all interpretations. In classical case where there are no truth value gaps, these two notions coincide. In a similar way, weak and strong consequence is distiguishable.

The VG semantics is applied to systems of modal logic, resulting the *Prior-style modal semantics* (as I. Ruzsa calls it). Here I only note that A. N. Prior's strong necessity operator 'L' is definable as follows:

$$LA =_{df} \sim\Diamond \mathrm{NON}(A).$$

It is *intensional logic* where the advantages of a VG approach are most pregnant. Let us see a number of applications which might be mentioned only here. It is well known that one of the most vulnerable points of a possible world semantics is the fact that sentences which are logically equivalent have the very same intension. Here, however, not all irrefutable sentences will be synonymous. If there is a nonlogical constituent of B which does not occur in A, then there is an interpretation in which $|B| = 2$ whereas $|A| \neq 2$, and so in this case "A & $(B \vee \sim B)$" is not synonymous with A.

In a semantics where no value gaps occur, the intensions of the following two formulas are indistinguishable:

(a) $Fab \vee \sim Fab$, (b) $Fba \vee \sim Fba$.

In VG semantics, however, if the factual value of F is a partial function, then there can be an interpretation where $|F|$ is defined for the pair $\langle |a|, |b| \rangle$ whereas it is undefined for the pair $\langle |b|, |a| \rangle$, and in this case the second formula lacks a truth value, and so the intension of the two formulas will be different. A trivial example for such a difference is:

(a) Peter either likes or does not like Mozart's music.

(b) Mozart's music either likes or does not like Peter.

Another application: As a consequence of our truth definition for

$$\forall x[Fx \supset Gx],$$

this formula means nothing more than there is no object for which F is true and G is false. This is clearly different from subordination which may hold among F and G if the truth domain of the former is a part of the truth domain of the latter. Let us express this relation by "F SUB G". It can be defined as follows:

$$F \text{ SUB } G =_{df} \forall x[Fx \supset \mathrm{VER}(Gx)] \;\&\; \forall x[Fx = Fx].$$

Note that F and G cannot be contraposited here, for the falsity domain of G is not necessarily a part of the falsity domain of F. This can yield an elegant solution to the well-known paradox of confirmation: let us formalize the sentence 'Every raven is black' by 'Raven SUB Black', then the contraposition, '~Black SUB ~Raven' will not follow.

2

Let us extend our investigations with the considerations of *Game Theoretical Semantics*. I assume that the basic principles and objectives of GTS are known and merely let me mention some of its basic ideas. These ideas are explained most easily when they are used to provide a semantic for interpreted first-order languages. We assume that L is an interpreted language where a truth value is assigned to each atomic sentence of L in the usual way. The notions of truth and falsity are extended to complex sentences S of L by associating with each of them a two-person zero-sum game $G(S)$. The players are called by Hintikka 'Myself' and 'Nature'. The aim of the semantical games is to verify or falsify S depending on the role of the players. The definition of $G(S)$, e.g., for the universal quantification is as follows:

(G.∀) The game $G(\forall x.S(x))$ begins with a choice of a member from the set D of individuals by Nature. Let the name of the chosen individual be b, the rest of the game is $G(S(b))$. [Here b can be either in L or a new name given by the players.]

The games serve their truth-defining purpose:

(T) S is true iff there exists a winning strategy for Myself (the initial verifier) in $G(S)$. S is false iff there exists a winning strategy in $G(S)$ for Nature.

By a *strategy* (in the sense of general game theory) we mean a rule or function which tells the player which move to make in each possible situation that may come up in the course of the game. Once the strategies of both players have been determined, the entire course of the game, and hence also its outcome, is uniquely determined. A *winning strategy* for a given player is one which leads to a win no matter what strategy one's opponent is pursuing.

There is a close relationship between game-theoretical truth-definition and Tarski-type one. The main difference is in tactics; the game-theoretical truth-definition operates from outside in, whereas the Tarski - type truth-definition operates from inside out.

The GTS can be extended to the situations dealt with possible - world semantics. In this case a semantical game is played not on a single model, but on a space of models on which suitable alternativeness relations are defined. The game $G(S_0)$ connected with a sentence S_0 starts from the world w_0 in which S_0 is being evaluated. At each move, the players are given not only a sentence S, but a world w. (See e.g. [6] and [8].)

Now we can ask the following: how is a game-theoretical semantics supposed to work in the case of sentences which are neither true nor false? What should the game of seeking and finding mean in the case of such quantified sentences?

If we modify the semantical rules of our game to cover such simple sentences too which contain atomic sentences producing a value gap, then we have to reinterpret winning and losing in such a game. What should it mean after all?

If we allow sentences with truth-value gap, my proposal is as follows: Let us change the notion of interpreted language used in GTS so that the evaluation of nonlogical constants should follow the methods introduced in Prior-type VG semantics.

Let us introduce the concept of a *strong* and a *weak winning strategy*. The concept of strong winning strategy remains the former, original winning strategy. The weakening of winning strategy should mean that whatever Nature's moves are, Myself can produce by the end of the game a sentence which *is not false*, i.e., its value is either 1 or 2. This concept of a weak winning strategy permits the treatment of all those problems which had to be excluded by the true or false dichotomy.

Let us apply this, e.g., to the rule of $(G.\forall)$: Nature chooses an object, say b, and the game continues in the usual way. If it becomes evident that Nature cannot produce a false atomic sentence, Myself will win no matter whether the value of the final sentence is 2 or 1. Basically, it is the interpreted language which should be changed, and perhaps some rules should be modified too. To put it more precisely: provided Nature does not have a winning strategy, Myself will have one, but it no longer means that the game should end up in a true sentence. Consequently, Myself can win if no object can be chosen the name of which substituted in the sentence makes it false — it is enough that the sentence is true in some cases and without a truth value in the other cases.

3

As a final topic of this paper, I would like to speak about the advantages of a fusion of GTS and VG semantics in the treatment of dialogues and discourses.

If we consider which are the most important sources of a value gap in dialogue games it is obvious to suppose that they appear mostly (though not exclusively) in interrogative dialogue games.

When a participant of the dialogue does not know the true and informative answer in all worlds — in this case it is a gross overstatement to say that that particular participant says a false proposition.

To avoid such consequences we will have to apply VG-techniques. We will have to introduce the notion of an *interrogative model* in order to treat dialogues in a formally correct way, as it is defined in Hintikka's papers. (See [3], [4], [5], and [7].)

Let us admit, on Hintikka's argumentation, that interrogative models are needed, e.g.

(a) to distinguish between active and tacit knowledge,

(b) to avoid the paradox of logical omniscience,

(c) to shed light on the problematic of reflexive knowledge, of knowing that one knows.

By combining the interrogative model with a VG semantics we may reach the following results: Our framework will retain the advantages of possible world semantics, whereas its former inadequacy to the problems above will vanish. Let us call this new framework *VG interrogative model*, let it incorporate the treatment of questions, and so that of dialogues into the framework of a VG possible world semantics. This new framework is able to treat tacit knowledge and its activization. Finally, it is not of minor importance that we make explicit the partial feature of GTS.

How can such a combination of the two theories be developed? The major principles are as follows. (Here I follow Hintikka's method described in his paper [4].)

(1) The interrogative process is conducted in a fixed first-order modal language L.

(2) The interrogative process is relative to a model M of L.

(3) There are two players, the Inquirer (also called 'the Knower') and the 'Oracle'.

(4) There are two kinds of moves, deductive and interrogative moves. At each stage of the game, the Inquirer has a choice between these two types of moves. Each move is relative to a subtableau.

(5) The usual tableau concept is used. (The Gentzen tableau construction is the best one.)

(6) In a deductive move, the Inquirer carries out a step of tableau construction. The set of tableau rules is to be modified by the rule:

> A subtableau is closed if S and "NON(S)" (or S and "$|S| = 2$") occur in the same column in it.

(7) In the interrogative move, the Inquirer addresses a question to the Oracle, the Oracle's answer is to be added to the left column of the subtableau in question.

(8) There are two kinds of questions, propositional questions and wh-questions. In the case of either question, its presupposition must present in the left column before the question is asked.

> (a) The presupposition of a propositional question is of the form "$S_1 \vee S_2 \vee \ldots \vee S_n$", and its possible answers are the S_i's ($i = 1, 2, \ldots, n$).
>
> (b) The presupposition of a wh-question is of the form "$\exists x.S(x)$" and its several answers are of the form "$S(b)$", where b is a name of an object in the domain of M.

The substantial modification is that the presuppositions may be without truth value in some worlds.

(9) If the presupposition of a question is true or without truth value in M, the answer must also be true or without truth value respectively in M.

(10) In the beginning of the process, there are in the left column two sets of formulas, T and RA. Here T consists of the theoretical premises $T_1, T_2, \ldots, T_n$, and the members of RA are of the form "$S_i \vee \sim S_i$". (Note that the latter formulas are only irrefutable, instead of being valid ones.)

The role of formulas of form "$S_i \vee \sim S_i$" as premises in a deductive argument or in an interrogative argument is the same: *to introduce new auxiliary individuals.*

Irrefutable premises open up the possibility for the Inquirer to ask new questions, by serving as presuppositions of yes—no questions which the Inquirer could not otherwise have asked.

As far as the presuppositions of yes—no questions are concerned a VG frame eliminates only the most drastic consequences, those which might force the player into self-contradiction (or lying).

The presupposition of a yes—no question "? A" is "$A \vee \sim A$". The response 'No answer' will mean that the presupposition (and its negation) are without a truth value: $|A \vee \sim A| = 2$, and $|A \& \sim A| = 2$. I think, these considerations are valid for the epistemically explicit formulations of A too.

A sketch of the principles certainly can be no substitute for the definition of the new rules of the game, which have to accept semantic value gaps for the demonstration that the players can find an optimal strategy. This detailed analysis have to be conducted elsewhere, where we show how a VG interrogative model works. Here I mention only that provided the definition of an interrogative model is extended into that of a VG interrogative model, the rules of evaluation can be stated in a 'world-line' terminology; as 'world-lines' are functions pick out individuals from the relevant possible worlds. Let us assume that these functions are partial ones, and apply VG-techniques to them.

From my previous investigations, however, I draw the conclusion that the combination of these two approaches yields a new framework, a new paradigm which can eliminate the unresolved problems of both approaches. I think that a VG interrogative model is feasible to investigate some actual problems of the knowledge representations and so gives help both to cognitive sciences, and to epistemology, and most of all to epistemic logic.

REFERENCES

[1] J. Hintikka, with J. Kulas, *The Game of Language*. Reidel, 1983. — The bibliography of this volume lists all GTS literature published before.

[2] J. Hintikka, The semantics of questions and the questions of semantics. *Acta Philosophica Fennica 28*, No. 4, 1976.

[3] J. Hintikka and E. Saarinen, Information-seeking dialogues. Some of their logical properties. *Studia Logica XXXVIII*, 1979.

[4] J. Hintikka, Knowledge representation and the interrogative model of inquiry. (Unpublished.)

[5] J. Hintikka, Logic of conversation as a logic of dialogue. In: E. Grandy and R. Warner (eds.), *Conditions on Rationality: Values, Intentions and Ends*. Clarendon, 1986.

[6] J. Hintikka, Exploring possible worlds. Forthcoming in the Proceedings of the 1986 Nobel Foundation Symposium on Possible Worlds in Arts and Sciences.

[7] J. Hintikka and M. B. Hintikka, Sherlock Holmes confronts modern logic: Toward a theory of information-seeking through questioning. In: E. Barth and J. L. Martens (eds.), *Argumentation: Approaches to Theory Formation*. Benjamins, Amsterdam, 1982.

[8] J. Hintikka, Game-theoretical semantics as a synthesis of verificationist and truth-conditional meaning theories. In: E. Le Pore (ed.), *Current Trends in Semantics*. Academic Press, London and New York, 1987.

[9] A. Madarász, The roads of formal pragmatics. *Logique et Analyse 29*, 1986.

[10] A. Madarász, A new approach to modal and intensional logic. In: *Intensional Logic, History of Philosophy and Methodology*. Akadémiai K., in print.

[11] I. Ruzsa, *Modal Logic with Descriptions*. M. Nijhoff/Akadémiai Kiadó, 1981.

[12] I. Ruzsa, Intensional logic and semantic value gaps. *Logique et Analyse 29*, 1986.

[13] T. J. Smiley, Sense without denotation. *Analysis 20*, 1960.

DEPARTMENT OF SYMBOLIC LOGIC AND METHODOLOGY OF SCIENCE
L. EÖTVÖS UNIVERSITY, BUDAPEST, HUNGARY

KLAUS WUTTICH

AN EPISTEMIC LOGIC WITHOUT NORMALITY PRESUPPOSITIONS

1. Preliminaries

A justified criticism of many systems of epistemic logic (e.g. of Jaakko Hintikka's pioneering system KB in [3] and of Wolfgang Lenzen's system in [4]) refers to the fact that in their epistemic propositions the speakers are assumed (tacitly) to be logically omniscient or logically infallible persons. They belive and know all true sentences, and they have to know all consequences of their knowledge. It is clear that such demands are not compatible with realistic notions of knowledge and belief. Therefore many ways have been proposed in order to avoid e.g. logical omniscience. Suffice it to mention the papers of V. Rantala and I.A. Gerasimova in [5]. But their solutions lead, among other things, to the elimination of the epistemic subjects in the formulas of epistemic logic.

To many authors it appears to be quite natural that in epistemic logic the speakers are consistently believing subjects, i.e. they accept without any hesitation formulas like

$$\text{(a)} \quad B(a,p) \supset \sim B(a, \sim p).$$

(If person a belives that p then he does not believe that non-p.)

I do not, however, think of such expressions to be laws of epistemic logic. In my opinion they are non-logical demands of normality and we have to distinguish them clearly from logical laws (see [8]).

In this paper I want to present a system of epistemic logic without any normality presuppositions. What we have to assume is (and this is asking for a lot) that the epistemic subjects are physically normal individuals capable of possessing, understanding, believing, accepting, rejecting, knowing etc. propositions. This system belongs to non-modal epistemic logic for its structure shows no analogy to that of alethic modal logic. (About the distinction between modal and non-modal epistemic logic see [11] and [12]).

2. The system SE^b - an epistemic logic for any speakers

2.1 The logical basis of SE^b

The system SE^b is built up by adding the schemes A^b1-A^b7, the definitions D1-D9, and the rule RK (all listed in 2.2 below) to the systems of strong logical entailment (S^s), of degenerate entailment (S^d) and of non-traditional predication theory (S^p) (see [7], [10]). The reason why I use the mentioned systems of logical entailment instead of classical logic is that I want to avoid the well-known paradoxes of material implication in SE^b. Non-traditional predication theory allows for distinction between internal and external negation and gives the logical foundation to the employment of a second negation sign. The inner negation is not an independent logical operator but a part of the operator of denial.

2.2 Additional axiomatic schemes, definitions and rules of SE^b

The system SE^b is built up by the following additions to the logical systems mentioned above. The symbols $\wedge$, $\vee$, $\neg$, $\sim$, $\vdash$ are, respectively, the operators of conjunction, disjunction, internal negation, external negation and the predicate of strong logical entailment. The expression $X \dashv\vdash Y$ will be used as an abbreviation for $X \vdash Y$ and $Y \vdash X$.

Before coming to the proper axiomatic schemes of SE^b we have to give the following additions to the above mentioned logical theories:

Addition to the alphabet:

1. a list of variables for terms, denoting speakers, epistemic subjects;

2. P, U, B, C, Aw-epistemic predicates of possession, understanding, belief, conviction, awareness.

Addition to the definition of a predicate formula:

1. If a is a variable for names of speakers and X is a propositional formula, then P(a,X), U(a,X), B(a,X), C(a,X), $\neg$ P(a,X), $\neg$U(a,X), $\neg$B(a,X) and $\neg$C(a,X) are predicate formulas.

2. If a and a' are variables for names of the same speakers and if Q^1(a,X) is one of the predicate formulas defined in 1, then Q^2(a,Q^1(a',X)) and $\neg Q^2$(a,Q^1(a',X)) are predicate formulas, where Q^2 stands for P, U, B, C or Aw.

The definition of the occurrence of a propositional formula in a propositional formula should take the following form:

A propositional formula occurs in a propositional formula in the following and only in the following cases:

1. X occurs in $\sim X$;

2. X^i (i = 1, 2, ..., n) occurs in $X^1 \wedge X^2 \wedge \ldots \wedge X^n$ and in $X^1 \vee X^2 \vee \ldots \vee X^n$;

3. if X occurs in Y, and Y in Z, then X occurs in Z.

At this point I would like draw the reader's attention to the fact that X occurs in epistemic statements not as the statement X, but just as a part of a term, because epistemic statements have in general the form $Q(a,tX)$, where t is a term-producing operator, which forms the term "the statement X" from the statement X. For the sake of brevity I shall always write $Q(a,X)$ instead of $Q(a,tX)$ in the formulas of SE_b.

Other epistemic predicates and their inner negations will be introduced by the following definitions:

D1. $K(a,X) =_{df} P(a,X) \wedge U(a,X) \wedge C(a,X) \wedge X[a'/a]$

D2. $\neg K(a,X) =_{df} \sim(P(a,X) \wedge U(a,X) \wedge C(a,X)) \wedge X[a'/a]$

D3. $Bm(a,X) =_{df} B(a,X) \wedge \sim K(a,X)$

D4. $\neg Bm(a,X) =_{df} \neg B(a,X) \vee K(a,X)$

D5. $A^i(a,X) =_{df} Aw(a,B(a',X))$

D6. $R^i(a,X) =_{df} Aw(a,\neg B(a',X))$

D7. $I(a,X) =_{df} \sim A^i(a,X) \wedge \sim R^i(a,X)$

D8. $\neg A^i(a,X) =_{df} R^i(a,X) \vee I(a,X)$

D9. $\neg R^i(a,X) =_{df} A^i(a,X) \vee I(a,x)$.

The expressions $K(a,X)$, $Bm(a,X)$, $A^i(a,X)$, $R^i(a,X)$ and $I(a,X)$ will be read as "a knows that X", "a merely believes that X", "a accepts inwardly that X", "a rejects inwardly that X" and "a is indifferent with respect to X".

The symbol $X[a'/a]$ denotes the statement, which will be obtained from the statement X, which does not contain iterations of the form $Q^1(a',Q^2(a',Y))$, where Q^1 and Q^2 are epistemic predicates (not necessarily different), by replacing in X the variable of speakers' names a' wherever it occurs by the variable a. If X contains iterations of the above mentioned form, only the outer occurrence of a' will be replaced by a in the iterated expression. For example, if X is the statement $U(a',p) \wedge P(a',X)$, then $X[a'/a]$ is the statement $U(a,p) \wedge P(a,p)$. But if X has the form $U(a',P(a',p))$ or $B(a',p) \wedge U(a',P(a',X))$, then $X[a'/a]$ is the expression $U(a,P(a',X))$, or $B(a,p) \wedge U(a,P(a',X))$ respectively. In the following axiom schemes and theorem schemes I have employed the abbreviation X' instead of $X[a'/a]$, i.e. I accept the definition:

$$X' =_{df} X[a'/a].$$

Additional axiom schemes of SE^b:

A^b1. $U(a,X) \vdash U(a,Y)$, where Y occurs in X

A^b2. $Q(a,X) \vee \neg Q(a,X) \vdash P(a,X)$, where Q stands for U, B, C, A^i or R^i

A^b3. $\sim Q(a,X) \vdash \sim P(a,X) \vee \neg Q(a,X)$, where Q stands for U, B, C, Bm, A^1 or R^1

A^b4. $C(a,X) \vdash B(a,X)$

A^b5. $\sim X' \vdash \sim K(a,X)$

A^b6. $K(a,i_1 K(b^1,\ldots,i_n K(b^n_n, jK(c,X)))\ldots) \vdash$
$K(a,i_1 K(b^1,\ldots,i_n K(b^n,X))\ldots)$,

where $n \geq 0$ and $i_1,\ldots,i_n, j$ stand for the presence or the absence of the internal negation (in arbitrary combinations)

A^b7. $Aw(a,iQ(a',X)) \Vdash K(a,iQ(a'X))$, where Q stands for an arbitrary epistemic predicate, i stands for the presence or the absence of the internal negation and a' is another (maybe the same) name of the speaker a.

Additional rule of inference:

In logical theories with the predicate of strong logical entailment the rule of inference, which corresponds to the contraposition rule, i.e.

If $X \vdash Y$ is a theorem of the system,
$\sim Y \vdash \sim X$ is a theorem of the system too,

is not valid in general, but only if X and Y involve the same propositional variables and predicate formulas. In order to increase the deductive force of the system SE^b and to avoid the increase of the axiom schemes, I will accept the following rule of inference:

RK. If $X \vdash Y$ is a theorem of SE^b and it is not of the form $p \wedge q \vdash p$, $\sim Y \vdash \sim X$ is a theorem of SE^b too, provided that the same variables (propositional variables and variables for speakers' names) occur in X and in Y. If $X \vdash Y$ is a theorem of SE^b and it is of the form $p \wedge q \vdash p$, then $\sim Y \vdash \sim X$ is a theorem of SE^b only if the same variables and predicate formulas occur in X and in Y.

The conditions of RK are strong enough to exclude the provability of formulas similar to the well-known paradoxes of material implication.

In the axiom schemes A^b2 and A^b3, Q must be replaced always

by the same predicate, i.e. either everywhere by U or everywhere by B or everywhere by C etc. It was necessary to postulate A^b5 because this significant formula is not provable from the remaining postulates, in spite of the "liberale" rule RK. The axiom scheme A^b6 allows to infer from iterated knowledge statements, saying something about the knowledge or ignorance of different speakers, to less iterated or non-iterated knowledge statements.
Axioms of this scheme are for example the formulas

$$K(a,K(b,p)) \vdash K(a,p)$$

$$K(a,\neg K(b,p)) \vdash K(a,p)$$

A^b7 connects the predicate Aw with the other epistemic predicates.

2.3 Some theorem schemes of SE^b

In the following I give some theorem schemes of SE^b. The proof will be mostly omitted. With the help of the following valid rule from definition theory we can get in SE^b theorems as consequences from the definitions D1-D9.

If $X =_{df} Y$, then $X \dashv\vdash Y$ (see [1], S. 54).

In square brackets I will refer to the axiom schemes, theorem schemes or definitions needed for the proof of the corresponding theorem scheme. The symbols S^s, S^d and S^p indicate that valid formulas, definitions and rules of inference of the strong logical entailment, of the degenerated logical entailment and of the non-traditional predication theory are applied in the proof.

T1.	$U(a,X) \vdash P(a,X)$	[A^b2]
T2.	$\neg U(a,X) \vdash P(a,X)$	[A^b2]
T3.	$B(a,X) \vdash P(a,X)$	[A^b2]
T4.	$C(a,X) \vdash P(a,X)$	[A^b2]
T5.	$\neg B(a,X) \vdash P(a,X)$	[A^b2]
T6.	$\neg C(a,X) \vdash P(a,X)$	[A^b2]
T7.	$A^i(a,X) \vdash P(a,X)$	[A^b2]
T8.	$R^i(a,X) \vdash P(a,X)$	[A^b2]
T9.	$\neg A^i(a,X) \vdash P(a,X)$	[A^b2]
T10.	$\neg R^i(a,X) \vdash P(a,X)$	[A^b2]

T11.	$K(a,X) \vdash P(a,X)$	[D1]
T12.	$K(a,X) \vdash U(a,X)$	[D1]
T13.	$K(a,X) \vdash C(a,X)$	[D1]
T14.	$K(a,X) \vdash X'$	[D1]
T15.	$\neg K(a,X) \vdash X'$	[D2]
T16.	$Bm(a,X) \vdash P(a,X)$	[D3, T3, S^s]
T17.	$\neg Bm(a,X) \vdash P(a,X)$	[D4, T5, T11, S^s]
T18.	$K(a,X) \vdash B(a,X)$	[T13, A 4, S^s]
T19.	$\neg K(a,X) \dashv\vdash \sim P(a,X) \vee \sim U(a,X) \vee \sim C(a,X) \vee \sim X'$	[D1, S^s]
T20.	$K(a,X) \vdash \sim K(a, \sim X)$	[T14, A^b5, S^s]
T21.	$K(a,X) \vdash \sim K(b, \sim X)$	[T14, A^b5, S^s]
T22.	$K(a,K(b,X)) \vdash K(a,X)$	[A^b6]
T23.	$K(a, \neg K(b,X)) \vdash K(a,X)$	[A^b6]
T24.	$K(a,K(b,K(c,X))) \vdash K(a,X)$	[A^b6]
T25.	$\vdash \sim K(a, X \wedge \sim X)$	[T14, S^d]
T26.	$\vdash \sim (K(a,X) \wedge K(a, \sim X))$	[T14, S^s, S^d]
T27.	$\vdash \sim (K(a,X) \wedge \neg K(a,X))$	[S^s, S^p, S^d]
T28.	$\vdash \sim (\neg K(a,X) \wedge \neg K(a, \sim X))$	[S^d, D2, S^d, S^s]
T29.	$Aw(a,K(a'X)) \vdash K(a,K(a',X))$	[A^b7]
T30.	$Aw(a, \neg K(a',X)) \vdash K(a, \neg K(a',X))$	[A^b7]
T31.	$Aw(a,B(a',X)) \vdash B(a,X)$	[A^b7, T14, S^s]
T32.	$A^i(a,X) \dashv\vdash Aw(a,B(a',X))$	[D5]
T33.	$A^i(a,X) \vdash K(a,B(a',X))$	[T32, A^b7, S^s]
T34.	$A^i(a,X) \vdash C(a,B(a',X))$	[T33, T13, S^s]
T35.	$A^i(a,X) \vdash B(a,B(a',X))$	[T33, T18, S^s]
T36.	$A^i(a,X) \vdash B(a,X)$	[T33, T14, S^s]

T37. $\sim I(a,X) \dashv\vdash A^i(a,X) \vee R^i(a,X)$ [D7, S^s]

T38. $\sim I(a,X) \vdash P(a,X)$ [T37, T7, T8, S^s]

T39. $Bm(a,X) \dashv\vdash B(a,X) \wedge \sim K(a,X)$ [D3]

T40. $Bm(a,X) \vdash B(a,X)$ [T39, S^s]

T41. $Bm(a,X) \vdash \sim K(a,X)$ [T39, S^s]

T42. $K(a,X) \vdash \sim Bm(a,X)$ [T41, RK]

T43. $K(a,X) \vdash B(a,X) \wedge \sim Bm(a,X)$ [T17, T42, S^s]

T44. $A^i(a,X) \vdash \sim Bm(a,B(a',X))$ [T33, T42, S^s]

If we accept the following definition

D10. $K^w(a,X) =_{df} K(a,X) \vee K(a,\sim X)$,

where $K^w(a,X)$ will be read as "a knows whether X", we can prove that the expression "a knows whether X" is derivable from "a knows that X".

T45. $K(a,X) \vdash K^w(a,X)$

Proof:

1. $\sim K(a,X) \wedge \sim K(a,\sim X) \vdash \sim K(a,X)$ [S^s]
2. $K(a,X) \vdash \sim(\sim K(a,X) \wedge \sim K(a,\sim X))$ [1, RK]
3. $K(a,X) \vdash K(a,X) \vee K(a,\sim X)$ [2, S^s]
4. $K(a,X) \vdash K^w(a,X)$ [3, D10] QED.

Further theorem schemes with K^w:

T46. $K(a,\sim X) \vdash K^w(a,X)$ [S^s, RK, D10]

T47. $K^w(a,X) \vdash U(a,X)$ [T12, A^b1, S^s]

T48. $K^w(a,X) \vdash P(a,X)$ [T47, A^b2, S^s]

The following expression, which is valid in the system of rational belief, knowledge and assumption of P. Weingartner (see [9])

$$\text{(b)} \quad K^w(a,K(b,X)) \vdash K(a,X)$$

is not provable in SE^b. This formula would be provable in SE^b, if in D10 the internal negation of $K(b,X)$ occured, for according to T22 and T23 the expression $K(a,X)$ follows both from $K(a,K(b,X))$ and $K(a,\neg K(b,X))$. Of course, it does not follow

from $K(a, \sim K(b,X))$, for the negation of $K(b,X)$ may be true because of X is false, and if it is so, then a cannot know that X. The acceptance of $K^w(a,K(b,X)) \vdash K(a,X)$ is probably due to the mixing up of the two forms of negation. Another reason for the rejection of this formula is that there are cases, where the premiss is true but the conclusion is false. If e.g. $K(a,X)$ is false because $K(a, \sim X)$ is true, $K^w(a,K(b,X))$ may nevertheless be true.
Also the formula

$$(c) \quad K(a, \sim X) \vdash K^w(a,K(b,X))$$

is not a theorem of SE^b although it appears to be plausible. It would be a theorem in an epistemic logic for logically infallible speakers. Such a system we could get from SE^b by addition of the rule R^u:

R^u. If $X \vdash Y$ is a theorem, $K(a,X) \vdash K(a,Y)$ is a theorem too. (But it would be enough also to supply SE^b by the scheme

$$(u) \quad K(a, \sim X) \vdash K(a, \sim K(b,X)),$$

saying that the speakers of this logic know the axiom scheme A^b5.)
Proof of the formula (c) in the mentioned extension of SE^b:

1. $K(a, \sim X) \vdash K(a, \sim K(b,X))$ [A^b5, R^u or (u)]
2. $K(a, \sim K(b,X)) \vdash K(a, \sim K(b,X)) \vee K(a,K(b,X))$ [S^s, RK]
3. $K(a, \sim X) \vdash K(a,K(b,X)) \vee K(a, \sim K(b,X))$ [1, 2, S^s]
4. $K(a, \sim X) \vdash K^w(a,K(b,X))$ [3, D10] QED.

In SE^b, it follows from $K(a, \sim X)$ only $\sim K(a,K(b,X))$, for the following theorem scheme is valid

T49. $K(a,X) \vdash \sim K(a,K(b, \sim X))$ [T21, A^b5, S^s]

Although it is not allowed to apply RK to T49, the contraposition of T49 is provable:

T50. $K(a,K(b, \sim X)) \vdash \sim K(a,X)$

Proof:

1. $K(a,K(b, \sim X)) \vdash \sim X'$ [A^b6, T14, S^s]
2. $\sim X' \vdash \sim K(a,X)$ [A^b5]
3. $K(a,K(b, \sim X)) \vdash \sim K(a,X)$ [1, 2, S^s QED.

T51. $K(a,K(b,\sim X)) \vdash \sim K(a,K(b,X))$ [A^b6, T49, S^s]

3. Consistency of SE

In order to prove the consistency of SE, before applying the semantic rules of the general theory of logical entailment and of the non-traditional predication theory from [7] we shall apply the following semantic transformations to the proper axiom schemes of SE^b in the given succession:

SR1. If in a formula of the form $\alpha \vdash \beta$ a formula U(a,X) occurs in α, the formula U(a,Y) occurs in β, and if Y occurs in X, then U(a,Y) will be replaced by U(a,X).

SR2. After the transformations of the preceding sort we shall make the following substitutions:

1. a formula of the form P(a,X) will be substituted by X,
2. a formula of the form U(a,X) will be substituted by X,
3. a formula of the form $\neg U(a,X)$ will be substituted by $X \wedge \sim X$,
4. a formula of the form B(a,X) will be substituted by X,
5. a formula of the form $\neg B(a,X)$ will be substituted by $X \wedge \sim X$,
6. a formula of the form C(a,X) will be substituted by X,
7. a formula of the form $\neg C(a,X)$ will be substituted by $X \wedge \sim X$,
8. a formula of the form Aw(a,X) will be substituted by X.

After the rules SR1 and SR2 the semantic rules of general theory of logical entailment ([7], p. 221) and of non--traditional predication theory ([7], p. 248f.) are to be applied.
After the transformations above the axiom schemes A^b1 - A^b5 take the following form:

1. $X \vdash X$

2. $X \vee (X \wedge \sim X) \vdash X$, if $Q = U, B, C, A^{i}, R^{i}$

3. $\sim X \vdash \sim X \vee (X \wedge \sim X)$, if $Q = U, B, C, A^{i}$
$\sim(X \wedge \sim X) \vdash \sim X \vee X$, if $Q = R^{i}$, Bm

4. $X \vdash X$

5. $\sim X \vdash \sim X$.

After the transformations the axiom scheme $A^{b}6$ can take the form of one of the following formulas depending on the succession and combination of the inner negation:

6. $X \vdash X$

7. $X \wedge \sim X \vdash X$

8. $X \wedge \sim X \vdash X \wedge \sim X$.

After the transformations the axiom scheme $A^{b}7$ takes the form of one the following formulas, depending on the occurrence of the inner negation:

9. $X \wedge \sim X \vdash X \wedge \sim X$

10. $X \vdash X$.

According to the semantic rules of general theory of logical entailment formulas 1-10 are tautologies. The transformations of SR1 and SR2 are not relevant for the formulas of general theory of logical entailment, thus the rules of inference of this theory remain valid.

It remains to be shown that after the above transformations the particular cases of the additional axiom of the predication theory with the epistemic predicates U, B, Bm, C, A^{i}, R^{i}, and K become tautologies of the general theory of logical entailment. We must look at the following special cases of the axiomatic scheme A^{p} from S^{p} (let Q be any epistemic predicate):

$A^{p}S$. $\neg Q(a,X) \vdash \sim Q(a,X)$.

If Q stands for the predicates U, B, C, A^{i} or K, we get from $A^{p}S$ after the given transformations

$$X \wedge \sim X \vdash \sim X.$$

If Q stands for the predicates Bm or R^{i}, we get from $A^{p}S$ after the given transformations

$$X \vdash \sim(X \wedge \sim X).$$

Both formulas are tautologies according to the semantical rules of the general theory of logical entailment. From this fact and from the consistency of this theory and of non-traditional predication theory follows the consistency of the system SE^b in the sense of the following metatheorem:

MT1. For the given interpretation all theorems of SE^b are tautologies.

The independence of SE^b can be easily shown, but the proof is laborious and therefore we shall omit it in this paper.

Some remarks on the completeness - or rather - non--completeness of SE^b. A proof of completeness cannot be given without a corresponding semantics for the system. I did not provide any semantics for SE^b because in my opinion the right way is to analyze the epistemic notions and the relations between them before building up the system itself. When this has been done, we can construct the calculus as the syntactical reconstruction of our informal considerations. If we now provide a semantics for the calculus, we can say in the semantics only the same as in our pre-logical, informal considerations-namely which of the syntactically well-formed sentences are valid in this logic and which are not valid.

The construction of a formal semantics moreover contains the danger that the author in order to attain its adequacy with the intuitively valid principles, "dresses up" the pre-logical, informal considerations if they are not in accordance with the possibilities of the semantic apparatus.

Because of these reasons I did not build up a semantics for the system SE^b, I could not define a notion of semantical completeness in this logic. But similarly as W. Stelzner (in [8], pp. 132ff), I could give the following definition:

DC. A system E of epistemic logic is complete according to a system of deduction D, if:

If $X^1, \ldots, X^n \vdash_D Y$, then

$$[B(a,X^1) \wedge \ldots \wedge B(a,X^n) \vdash_E B(a,Y)] \vee$$
$$[C(a,X^1) \wedge \ldots \wedge C(a,X^n) \vdash_E C(a,Y)] \vee$$
$$[K(a,X^1) \wedge \ldots \wedge K(a,X^n) \vdash_E K(a,Y)], \quad \text{where } n \geq 1$$

The validity of the following metatheorem is obvious, because SE^b was built up precisely with the aim to exclude deductive infallibility as it is shown in the completeness relating to a less trivial system of deduction.

MT2. SE^b is not complete according to an arbitrary system of deduction D, which allows other inferences than such of the form $X \vdash X$.

The formulas $B(a,X) \vdash B(a,X)$, $C(a,X) \vdash C(a,X)$ and $K(a,X) \vdash K(a,X)$ are theorem schemes of SE^b. This shows that this system is complete according to a (as one may say) degenerated system of deduction, which allows only inferences of the form $X \vdash X$.

NOTE

1. Such an analysis was done in [11].

REFERENCES

[1] Dölling, J.: 'Definitionen in empirischen Wissenschaften'. In: H. Wessel (ed.), Logik und empirische Wissenschaften. Berlin, 1977. pp. 38-62.

[2] Gerasimova, I.A.: 'Okrestnostnye semantiki dlya epistemicheskikh logik'. In: [5], pp. 207-219.

[3] Hintikka, J.: Knowledge and Belief. Ithaca, N.Y., 1962.

[4] Lenzen, W.: Glauben, W[illegible] und Wahrscheinlichkeit. Systeme der epistemi[illegible] Logik. Wien — New York, 1980.

[5] Modal'nye i intensional'nye logiki i ikh primenenie k problemam metodologii nauki. Moskva, 1984.

[6] Rantala, V.: 'Semantika nevozmozhnykh mirov i logicheskoe vsevedenie'. In: [5], pp. 199-207.

[7] Sinowjew, A. — H. Wessel: Logische Sprachregeln. Eine Einführung in die Logik. Berlin, 1975.

[8] Stelzner, W.: Epistemische Logik. Zur logischen Analyse von Akzeptationsformen. Berlin, 1984.

[9] Weingartner, P.: 'A system of rational belief, knowledge and assumption'. In: Grazer philosophische Studien. 12/13 (1981.) pp. 243-263.

[10] Wessel, H.: Logik. Berlin, 1984.

[11] Wuttich, K.: Modale und nichtmodale epistemische Logik. Berlin, Humboldt-Universität, Sektion Marxistisch-leninistische Philosophie, Bereich Logik, Diss. B. 1986.

[12] Wuttich, K.: Some remarks on the history of epistemic logic. (forthcoming).

HUMBOLDT UNIVERSITY, BERLIN, GDR

ELENA D. SMIRNOVA

AN APPROACH TO NONSTANDARD SEMANTICS AND THE FOUNDATION OF LOGICAL SYSTEMS

1

The present approach to the construction of a nonstandard semantics is closely connected with the analysis and investigation of the nature of logical knowledge. The appearance of logical systems of most diverse kind in modern logic makes the problem of their foundations especially acute. If one treats logical laws and forms as the laws and forms of thinking, considered as the process of mental activity of human beings — as it was done by the representatives of psychologism in logic — then these laws bear an empirical character and change, accordingly, together with the nature of our mind; and what is more important, the problem of foundation of logical systems disappears and the question concerning the ontological presuppositions of logical laws and forms became meaningless. Accordingly, semantic considerations in this casé have nothing to do with logical laws and structures.

From our point of view, considering "empiricism" of logical laws, it is necessary to differentiate strictly the problem of interpretation of logical laws as the laws of some natural processes of mental activity of human beings on the one hand, and the problem of their informativity and their relation to the sphere of objects, on the other hand.

I consider, also, that the appearance of logical systems dealing in their semantics with some characteristics of the human agent such as his knowledge, attitudes, etc. does not mean a return to some form of empiricism and psychologism in logic.

One of the focal ideas of our approach is the assumption that logical systems are controlled by the semantics of the language, and it is the semantic phenomena that makes a stepping stone for profound insights into philosophical issues concerning the nature of logical reasoning.

In contrast to most of generally accepted point of views I don't believe that the semantics of logical systems has to be purely formal. They must be essentially epistemological in their nature. I think that only such kind of semantic theories would provide a suitable basis for formal logical systems. The main problem is to disclose the inner interaction between language semantics, its ontological and epistemic motivations, and logical systems. The crucial question here is what kind of such motivations does really play a role in semantical analysis.

Thus, languages of logical systems together with their semantics fulfil a special function: they play the role of analytical methods in logic and are languages of a special kind — the languages of logical reasoning.

I want to propose a nonstandard approach to the construction of semantics, on the basis of some non-formal considerations. Logical laws and structures on one and the same level. It seems to me rational to distinguish two kinds of such presuppositions. Then, two types, two levels of logical laws appear, correspondingly.

Logic, from our point of view, does not directly depend on the empirical data, on the empirical characteristics of the objects of discourse. It is a theoretical and not an empirical science. But logic depends on the assumed abstractions and idealizations. In other words, it depends on the type of ideal entities with which we deal in semantics. That is why we consider it reasonable to subdivide the laws of logic into two types. The first type depends on the particular ontological assumptions, that is, on assumptions referring to the objects of discourse, on the conditions of their introduction and on accepted idealizations.

The laws of the second kind do not depend on the limitations imposed on the universe of discourse. They depend only on our notions of truth, falsity, logical entailment, and so on.

It is interesting to note that the division of logical laws into laws of logic proper and laws of metalogic was marked in the Russian logic even at the beginning of the 20th century and it is connected with the name of the Russian logician Vasiliev. The laws of logic proper Vasiliev considered to be empirical in the sense that they change with the change of the character of the objects of discourse. In contrast to them, the laws of metalogic connected with the ideas of truth, falsity, judgement and so on, were considered by Vasiliev to be unchangeable, constant.

I assume that even those laws of logic which depend on the concepts of truth, falsity, and entailment are also able to change. Moreover, it is possible to show that both the laws of the first type and the laws of the second type may determine the systems of logical reasoning.

It is well known that the construction of a series of semantics is based on the conception of nonstandard worlds (impossible possible worlds, imaginable worlds, etc.), or, in other terminology, nonstandard state-descriptions (contradictory and incomplete ones). Of course, it is a very interesting and fruitful approach on the basis of which very interesting semantics have been worked out. However, this kind of construction of semantical theory accepts (to a certain extent) some assumptions of ontological character, referring to the sphere of the objects of discourse.

In contrast to this approach I am going to show that adequate semantics can be constructed without using impossible possible worlds and without the concepts of contradictory and incomplete state descriptions. In any case, these concepts are not taken as a background and no assumptions are made with respect to the objects of discourse.

Instead of this, partially defined predicates are accepted. We assume that the predicates of truth and falsity belong to this kind – they can be partially defined. Further, we proceed from the idea of the symmetry of the concepts of truth and falsity (and this is very important). Falsity is considered to be an independent notion and not as absence or negation of truth. Correspondingly, this attitude is applicable in defining the logical connectives.

Let us consider the principles of building language semantics. I shall construct my semantics using the notion of possible worlds. Let W be a non-empty set of possible worlds, φ a function ascribing a pair of sets $\langle H_1, H_2 \rangle$ to propositional variables where $H_1 \leq W$, $H_2 \leq W$.

$\varphi_T(p) = H_1$ is the class of worlds in which p holds (the domain of p).

$\varphi_F(p) = H_2$ is the class of worlds in which p does not hold (the anti-domain of p).

We shall use a propositional language with the logical connectives &, V, ⊃, ~. Let us introduce conditions of ascribing truth values to complex formulas as follows:

$$\varphi_T(A \,\&\, B) = \varphi_T(A) \cap \varphi_T(B) \qquad \varphi_F(A \,\&\, B) = \varphi_F(A) \cup \varphi_F(B)$$
$$\varphi_T(A \vee B) = \varphi_T(A) \cup \varphi_T(B) \qquad \varphi_F(A \vee B) = \varphi_F(A) \cap \varphi_F(B)$$
$$\varphi_T(A \supset B) = \varphi_F(A) \cup \varphi_T(B) \qquad \varphi_F(A \supset B) = \varphi_T(A) \cap \varphi_F(B)$$
$$\varphi_T(\sim A) \;\; = \varphi_F(A) \qquad \varphi_F(\sim A) \;\; = \varphi_T(A)$$

Let us introduce the concepts of truth and falsity in a given world α:

$$\alpha \models^{\varphi} A \iff \alpha \in \varphi_T(A), \qquad \alpha \;{}^{\varphi}\!\!=\!\!\| \; A \iff \alpha \in \varphi_F(A).$$

A formula A is tautological iff $\forall\varphi(\varphi_T(A) = W)$.

A formula A is irrefutable (non-falsifiable) iff $\forall\varphi(\varphi_F(A) = \emptyset)$.

The relation between the classes $\varphi_T(p)$ and $\varphi_F(p)$ may but need not satisfy the following conditions:

(1) $\varphi_T(p) \cap \varphi_F(p) = \emptyset$, (2) $\varphi_T(p) \cup \varphi_F(p) = W$.

Accepting both (1) and (2) we get the standard semantics. Accepting (1) and rejecting (2) is the semantics with truth value gaps; accepting (2) and rejecting (1) is the semantics with glut evaluations (which permits of the overlap of truth and falsity); rejecting both (1) and (2) we get relevant semantics. – If both conditions (1) and (2) are accepted, the class of tautological formulas coincides with the class of irrefutable formulas and is identical to the class of tautologies of classical logic.

If (1) is fulfilled and (2) is not (we shall write shortly (1), $(\overline{2})$), that is, if we have truth value gap semantics, and if the connectives are defined as above, then the class of tautologies is empty, and the class of non-falsifiable formulas coincides with that of classical propositional logic.

If ($\overline{1}$), (2) take place, then the class of tautological formulas coincides with that of classical logic, and the class of non-falsifiable formulas is empty. If both ($\overline{1}$) and ($\overline{2}$)take place, then both the class of tautologies and the class of non-falsifiable formulas are empty.

Thus we have defined the semantics of the language. But we have not yet given the logic proper. Semantics does not determine a logical system as long as the notion of logical entailment (or validity) is not defined.

Usually, logics are determined and given by introducing the class of tautologies. However, as we have seen, classes of tautological formulas in some semantics may be empty. Thus, in the logic of Kleene [4] the class of tautologies is empty and the class of non-falsifiable formulas coincides with that of classical logic. We think that it would be better to consider the concept of logical entailment as being the main concept. It is just this concept and not the concept of tautological formulas that characterizes logical systems.

We shall introduce several entailment relations of different types, proceeding only on the basis of the concepts of the domain $\varphi_T(A)$ and the antidomain $\varphi_F(A)$ of a sentence A.

In our approach, it is possible to introduce not a single, but a whole class of different relations of logical entailment. As a whole, it is possible to introduce sixteen relations of logical entailment. However, due to the interrelations between them, it is sufficient to treat only nine of them. We define these relations of logical entailment in terms of interrelations between domains and antidomains of formulas. We shall denote the complement of the class $\varphi_F(A)$ by $\varphi_{\overline{F}}(A)$.

Now "A entails B" is definable by one of the following nine relations:

R1. $\varphi_T(A) \cap \varphi_{\overline{F}}(A) \leq \varphi_T(B)$

R2. $\varphi_T(A) \cap \varphi_{\overline{F}}(A) \leq \varphi_{\overline{F}}(B)$

R3. $\varphi_T(A) \cap \varphi_{\overline{F}}(A) \leq \varphi_T(B) \cup \varphi_{\overline{F}}(B)$

R4. $\varphi_T(A) \leq \varphi_T(B)$ [a]

R5. $\varphi_T(A) \leq \varphi_{\overline{F}}(B)$ [d]

R6. $\varphi_T(A) \leq \varphi_T(B) \cup \varphi_{\overline{F}}(B)$

R7. $\varphi_{\overline{F}}(A) \leq \varphi_T(B)$ [c]

R8. $\varphi_{\overline{F}}(A) \leq \varphi_{\overline{F}}(B)$ [b]

R9. $\varphi_{\overline{F}}(A) \leq \varphi_T(B) \cup {}_{\overline{F}}(B)$

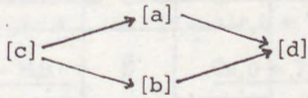
The omitted relations may be reduced to the conjunctions of some of these nine relations.

Different logics are based, according to our approach, on these various relations of logical entailment, combined with the acceptance or non-acceptance of conditions (1) and (2) (that is, combined with conditions imposed on the relations between the domain and the antidomain of a sentence).

If both conditions (1) and (2) are accepted, that is, if $\varphi_T(p) = \varphi_{\overline{F}}(p)$, then all relations of logical entailment coincide and we get only one relation of logical entailment which is formalized by the system C (classical propositional calculus).

If condition (1) is accepted and condition (2) is not (that is, if we deal with semantics with truth value gaps), then the above mentioned nine relations are reduced to the following four ones: R4, R5, R7, R8. Let us mark them correspondingly: [a], [d], [c], [b]. It is easy to see that they are connected in the following way:

(1), $(\overline{2})$

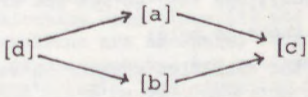


Under these conditions, Modus Ponens holds for entailment of type [a], as $\varphi_T(A) \cap \varphi_T(A \to B) \leq \varphi_T(B)$; but Deduction Theorem does not hold.

For the entailment of type [b], under the same conditions, Modus Ponens does not hold, but Deduction Theorem is valid. Entailment of type [d] does not depend on truth value gaps and behaves as classical if condition (1) holds. The relation of type [c] in semantics with partially defined truth is empty, that is, no pair of formulas belong to this relation.

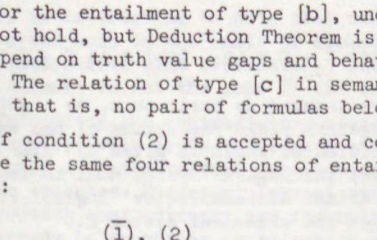
If condition (2) is accepted and condition (1) is not, that is $(\overline{1})$,(2), we have the same four relations of entailment, but they are differently connected:

$(\overline{1})$, (2)



Logics based on semantics with glut evaluations are dual to logics built within the frames of semantics with partially defined predicate of truth.

In semantics admitting both truth value gaps and glut evaluations, we have the following connections among the types of entailment:



3

Now let us consider the problem of the formalization of logics described above. Let us formulate some well known logical systems in sequential form. The following sequent rules are common for these systems:

$$\frac{\Gamma \to \Theta, A \quad \Gamma \to \Theta, B}{\Gamma \to \Theta, A \,\&\, B} \qquad \frac{\Gamma \to \Theta, \sim A, \sim B}{\Gamma \to \Theta, \sim(A \,\&\, B)} \qquad \frac{A, B, \Gamma \to \Theta}{A \,\&\, B, \Gamma \to \Theta} \qquad \frac{\sim A, \Gamma \to \Theta \quad \sim B, \Gamma \to \Theta}{\sim(A \,\&\, B), \Gamma \to \Theta}$$

$$\frac{\Gamma \to \Theta, A, B}{\Gamma \to \Theta, A \vee B} \qquad \frac{\Gamma \to \Theta, \sim A \quad \Gamma \to \Theta, \sim B}{\Gamma \to \Theta, \sim(A \vee B)} \qquad \frac{A, \Gamma \to \Theta \quad B, \Gamma \to \Theta}{A \vee B, \Gamma \to \Theta} \qquad \frac{\sim A, \sim B, \Gamma \to \Theta}{\sim(A \vee B), \Gamma \to \Theta}$$

$$\frac{\Gamma \to \Theta, \sim A, B}{\Gamma \to \Theta, A \supset B} \qquad \frac{\Gamma \to \Theta, A \quad \Gamma \to \Theta, \sim B}{\Gamma \to \Theta, \sim(A \supset B)} \qquad \frac{\sim A, \Gamma \to \Theta \quad B, \Gamma \to \Theta}{A \supset B, \Gamma \to \Theta} \qquad \frac{A, \sim B, \Gamma \to \Theta}{\sim(A \supset B), \Gamma \to \Theta}$$

$$\frac{\Gamma \to \Theta, A}{\Gamma \to \Theta, \sim\sim A} \qquad \frac{A, \Gamma \to \Theta}{\sim\sim A, \Gamma \to \Theta}$$

The structural rules are the usual ones. Cut elimination theorem holds. We get different systems by choosing different basic sequents. Let us consider the following four types:

(S1) $A, \Gamma \to \Theta, A$; (S2) $A, \sim A, \Gamma \to \Theta$; (S3) $\Gamma \to \Theta, B, \sim B$; (S4) $A, \sim A, \Gamma \to \Theta, B, \sim B$.

If only sequents of type (S1) are assumed as basic ones, we get De Morgan's logic (M). (S1) and (S2) together yield the system of Hao Wang (WH) in sequential form (axiomatic construction of which is given by Alan Rose in [3]). Assuming (S1) and (S3), we get the logical system dual to Hao Wang's logic (DWH). (S1) and (S4) yield a variant of Łukasiewicz' logic. Finally, by assuming (S1), (S2), and (S3) we get the classical system C.

The following theorems hold:

THEOREM 1. In the semantics with truth value gaps, the relation R7 is empty; the relation R4 is formalized by WH; the relation R8 by DWH; and the relation R5 by C.

THEOREM 2. In the semantics with glut evaluations, the relation R5 is empty; the relation R4 is formalized by DWH; the relation R8 by WH; and the relation R7 by C.

THEOREM 3. In the semantics free from the conditions (1) and (2), relations R7 and R5 are empty; relations R4 and R8 are formalized by M; relations R1 and R2 by WH; relations R6 and R9 by DWH; and the relation R3 by C.

The proofs of these theorems can be find in [6].

It is possible to treat the problem of adequate formalization of more complex relations of logical entailment, combining some of the the above mentioned nine relations of entailment. For instance, let us introduce the entailment [e] as a conjunction of the relations of entailments [a] and [b].

We admit different relations among sentences, depending on their domain and antidomain, and we proceed the formalization of these relations by formal systems of different kinds, taking into account the relations between the concepts of truth and falsity (between the sets $\varphi_T(A)$ and $\varphi_F(A)$). The result of the formalization of the analysed relations is summarized in the following chart:

conditions \ entailment	[a]	[b]	[d]	[c]
(1), (2)	C	C	C	C
(1), $(\bar{2})$	WH	DWH	C	Ø
$(\bar{1})$, (2)	DWH	WH	Ø	C
$(\bar{1})$, $(\bar{2})$	M	M	Ø	Ø

It is interesting to note that one and the same formal system may be constructed both on the basis of semantics with value gaps and on the basis of semantics with glut evaluations, but in this case the relation of logical entailment is not the same in the two systems.

Incomplete and contradictory state descriptions usually correspond to the semantics with truth value gaps and glut evaluations, respectively.[1] As a rule, only a single standard relation of logical entailment, marked by us as [a], is accepted. Different logical systems appear as a result of accepting or rejecting contradictory and incomplete state descriptions — and this corresponds to accepting or rejecting conditions (1) and (2). If contradictory and incomplete state descriptions are accepted, we get De Morgan's logic. If incomplete state descriptions are accepted but contradictory state descriptions are not, we get Hao Wang's logic. And accepting contradictory but rejecting incomplete state descriptions we get the logic dual to Hao Wang's.[2] If neither contradictory nor incomplete state descriptions are accepted, we get classical logic. (Cf. [2].)

In contrast to the above mentioned attitude, our approach allows us, as we have seen, to receive Hao Wang's logic, its dual, and the classical logic either by accepting truth value gaps or by accepting glut evaluations.

In our analysis, instead of introducing additional truth values or contradictory and incomplete possible worlds, it is sufficient to assume that the predicates of truth and falsity are partially defined and that these notions have been introduced independently of one another.

4

Let us sum up: The approach suggested by me is based on some principles:

1. The notion of impossible possible worlds and its analogues are not admitted in the semantics as judged to be less clear.

2. Instead of this, the idea of partially defined predicates and functions is employed (in particular, predicates of truth and falsity are assumed to be partially defined).

3. The ascribing of the domain and antidomain to a proposition is done independently of one another. This, in fact, means the introduction of the notions of truth and falsity independently of one another.

4. Dealing with such objects as the classes $\varphi_T(A)$ and $\varphi_F(A)$ it is possible to establish different relations between them, to accept or not to accept conditions (1) and (2). It is possible to accept one of them and reject the other, for they are independent of one another.

5. The function of ascribing values to propositional variables is given in a generalized form: not the truth values in a given world, that is, not the objects t and f, are ascribed to propositional variables, but special "intensional objects" — classes of worlds $\varphi_T(p)$ and $\varphi_F(p)$. It is this that gives the intensional character to the propositional connectives, cf. [1].

6. Moreower, when defining logical connectives, no limitations are imposed on the relations between the classes $\varphi_T(A)$ and $\varphi_F(A)$. The independence in ascribing domains and antidomains to propositions allows us to treat the operation of negation in a generalized way. As a result of the above mentioned principles we get semantics with truth value gaps and with glut evaluations.

7. The key concept of logic is the concept of logical entailment. On the basis of concepts of domain and antidomain of a proposition different relations of entailment may be introduced, independently of the conditions (1) and (2). (Correspondingly, independently of state descriptions and limitations imposed on them.)

8. It is these notions of entailment, combined with conditions (1) and (2), that determine different logical systems.

9. One and the same formal system can be based on different semantics and on different assumptions concerning the relations between the classes $\varphi_T(A)$ and $\varphi_F(A)$. However, the relation of entailment which is formalised in this case, may change.

Coming back to the questions which we put at the beginning of this paper we must note that no ontological assumptions concerning the universe of discourse have been taken into consideration. Only the notions of truth and falsity and their correlates were changing. Correspondingly, the notion of logical entailment was changing as well.

It is assumptions of metalogical character which determine, according to our approach, laws and rules of logical systems which have been considered here.

NOTES

[1] In particular, in our country, semantics for relevant logic have been succesfully worked out on the basis of generalized state descriptions. Cf. [2].

[2] Moreover, E. K. Voyshvillo formulates such limitations for state descriptions as a result of which we get a fragment of Łukasiewicz' logic. In this case he accepts in semantics the condition that every contradictory state description is to be complete. Cf. [2].

REFERENCES

[1] Belnap, N. D., Jr., A useful four-valued logic. In; Epstein, G. and J.M. Dunn (eds.), *Modern Uses of Multiple-Valued Logic*. Proceedings of the 1975 International Symposium of Multiple-Valued Logic. Reidel, Dordrecht, 1976.

[2] Voyshvillo, E. K., Semantics of relevant logic and the question about the nature of logical laws. In: *Mind and culture*. Moscow University, 1983, in Russian.

[3] Ermolaeva, N. M., Muchnik, A. A., Modal extensions of Hao Wang's type logical calculi. In: *Investigations in formalized languages and non-classical logics*. Moscow, 1974, in Russian.

[4] Kleene, S. C., *Introduction to Metamathematics*. New York, 1952.

[5] Smirnova, E. D., Semantics with truth value gaps, glut evaluations, and the concept of logical entailment. In: *Intensional logics and the logical structure of theories*. Tbilisi, 1985, in Russian.

[6] Smirnova, E. D., *Logical semantics and philosophical foundations of logic*. Moscow University, 1986, in Russian.

MOSCOW STATE UNIVERSITY
FACULTY OF PHILOSOPHY
DEPARTMENT OF LOGIC

HORST WESSEL

NON-TRADITIONAL PREDICATION THEORY AND ITS APPLICATIONS

1. THE INTUITIVE BASIS OF NON-TRADITIONAL PREDICATION THEORY

We presuppose the classical sentential logic with the classical or external negation ~. The external negation concerns the whole sentence. In natural languages, we do not have a single word for external negation. If we want to express the external negation in natural languages, we have to use expressions like "It is not the case that ...".

In our construction of the non-traditional predication theory, we assume the pre-logical capability of man to distinguish between subject terms and predicate terms. We represent elementary sentences by schemes of the form

$$s \leftarrow P,\ (s_1,\ldots,s_n) \leftarrow P,\quad s \not\leftarrow P, \text{ and } (s_1,\ldots,s_n) \not\leftarrow P$$

where $n \geq 2$. In these schemes, s, s_1, ..., s_n are subject terms (names of objects which are spoken about in the sentence), P is a (one-place, n-place) predicate, the operators $\leftarrow$ and $\not\leftarrow$ are used as the English "is" ("has") and "is not" ("has not"). In classical and also in intuitionistic logic often $s \not\leftarrow P$ is identified with $\sim(s \leftarrow P)$. I think that is a logical mistake. At least in the following cases the two sentences $s \leftarrow P$ and $s \not\leftarrow P$ do not exhaust all possibilities:

1) If the meaning of the predicate P is undefined for subjects of the type s (e.g. "The moon is honest"), then

(*) $$\sim(s \leftarrow P) \wedge \sim(s \not\leftarrow P)$$

will be true.

2) If the meaningful use of $s \leftarrow P$ or $s \not\leftarrow P$ presupposes the truth of another sentence A (e.g., "N. has stopped beating his wife" or "N. was excluded from the party"), then the case of (*) will be possible, if $\sim A$ applies.

3) If it is impossible to state whether $s \leftarrow P$ or $s \not\leftarrow P$ holds then it will be also true the sentence (*) (e.g.,"In the decimal development of π, zero occurs 10^{10} times in succession.").

4) If there does not exist an object named s, (*) will be true too (e.g., "Round squares are round").

5) If P is a vague predicate like "bald" or "red", then, in addition to $s \leftarrow P$ and $s \not\leftarrow P$, (*) will be true for some instances.

In all these cases, in addition to $s \leftarrow P$ and $s \not\leftarrow P$, the case of uncertainty $\sim(s \leftarrow P) \wedge \sim(s \not\leftarrow P)$ is possible. Even if the reasons for the occurrence of this uncertainty are different, we are always concerned with the same situation regarding the logical respect: apart from affirming or denying a predicate to a subject, it is still necessary to consider the case that the predicate involved is neither affirmed nor denied to the subject. Such cases of uncertainty can be found in the practice of language and therefore they must be taken into account in a logical predication theory.

2. SEMANTICAL RULES AND AXIOMATIC CONSTRUCTION OF THE NON-TRADITIONAL PREDICATION THEORY

Instead of $s \leftarrow P$ we write in the following $P(s)$ (or only p), instead of $s \not\leftarrow P$ we write $\neg P(s)$ (or only $\neg p$), and instead of $\sim(s \leftarrow P) \wedge \sim(s \not\leftarrow P)$ we write $?P(s)$ (or $?p$). The symbol $\neg$ is called the *internal negation*. We say that the formulas A and $\neg A$ are *contrary to each other*. In accordance with our intuitive considerations about elementary sentences we accept the following semantic rules in addition to the semantic rules of classical sentential logic.

R1. Predicate formulas of form $^{n}(s)$ may have the values "true" (t) and "false" (f).

R2. If $P(s)$ has the value t, $\neg P(s)$ must have the value f.

R3. If $P(s)$ has the value f, the value of $\neg P(s)$ does not depend on the value of $P(s)$, that is, $\neg P(s)$ can have the value t or f.

R4. The value of $?P(s)$ is the same as that of $\sim P(s) \wedge \sim\neg P(s)$.

The non-traditional theory of predication is decidable.

For the axiomatic construction of the non-traditional theory of predication we are adding the following axiom scheme to a consistent and complete axiomatic construction of calssical sentential logic:

A1. $\sim(f(a) \wedge \neg f(a))$

where a is a subject variable or a sequence of subject variables and f is accordingly a one-place or an n-place predicate variable.

This axiom system of non-traditional predication theory is semantically consistent and complete. (See [10].) Let us now consider some applications of the non-traditional theory of predication.

3. NON-TRADITIONAL PREDICATION THEORY (PT) AND INTUITIONISTIC SENTENTIAL LOGIC (ISC)

A logical analysis of the so-called intuitionistic counterexamples against some laws of classical logic leads to the conclusion that the intuitionists wrongly identify the external and the internal negation, i.e., they do not distinguish between the two different forms of negation. Their counterexamples only show that the formulas $p \vee \neg p$ and $\sim\neg p \supset p$ are not valid, but do not concern the classical logical laws $p \vee \sim p$ and $\sim\sim p \supset p$. Because the formula $\sim p \equiv \sim\sim\sim p$ is also valid in the intuitionistic logic, the differences between the classical and intuitionistic negation are connected with the negations occurring immediately before a sentential variable. This leads to the conjecture that the following statement holds:

If A is a classically provable formula of sentential logic, and A' is obtained from A via substituting all external negations $\sim$ occurring immediately before a variable by internal negations $\neg$, then A' is provable in the non-traditional theory of predication if and only if the formula A is provable in the ISC.

But this conjecture is wrong.

The following comparison of PT with ISC concerning negations shows that intuitionistic negation is a confusion of external and internal negation. We use the following abbreviations:

CSC – classical sentential calculus,
ISC – intuitionistic sentential calculus,
PT – non-traditional theory of predication.

Let us adopt the following definitions:

D1. The formula obtained from a formula A of ISC via substituting all intuitionistic operators by the corresponding classical ones is called *the C-representative (C-R) of the formula A in CSC*.

D2. The formulas obtained from a formula A of ISC in the following way are called *P-representatives (PR) of A in PT:*

1) if A contains no negation immediately before a variable, the P-representative of A will be its C-representative;
2) if A contains negations immediately before a variable, the P-representatives of A in PT will be the formulas obtained from A via substituting at least one negation which occurs immediately before a variable by the internal negation $\neg$ and by substituting all other intuitionistic operators by the corresponding classical ones.

With the help of the proposed terminology we get the following classification of all formulas of ISC:

THE FORMULAS OF ISC

provable in ISC		
All PR's are provable	Some PR's are provable, some are not	No PR's are provable
$\sim p \supset (p \supset q)$	$p \supset \sim p \supset \sim p$	$(p \equiv q) \supset (\sim p \equiv \sim q)$
$p \supset \sim\sim p$	$\sim\sim\sim p \supset \sim p$	$\sim\sim(p \vee \sim p)$
$\sim(p \wedge \sim p)$		

not provable in ISC			
CR is provable in CSC			CR is not provable in CSC
No PR's are provable	Some PR's are provable, some are not	All PR's are provable	
$\sim p \vee p$	$\sim p \supset \sim q \supset (q \supset p)$	$p \supset \sim q \supset p \supset p$	$p \wedge \sim p$
$\sim\sim p \supset p$	$\sim p \supset \sim q \supset (\sim q \supset p)$	$p \vee (p \supset \sim q)$	

This classification shows that the intuitionistic negation is a confusion of external and internal negation. In the framework of the non-traditional predication theory, a more detailed analysis of the intuitionistic problems is possible without any restriction of classical logic. (See [9].)

4. DIALETHEISM AND NON-TRADITIONAL PREDICATION THEORY

Dialetheism is the mystical belief that there are true logical contradictions. I will show that dialetheism is founded on logical mistakes. Dialetheism emerged in the framework of the so called paraconsistent systems of logic.

The dialetheists state that the logical law of the excluded contradiction is not valid. There are different formulations of the law of the excluded contradiction, for example the following ones:

$$\sim(A \wedge \sim A)$$

$$\sim((s \leftarrow P) \wedge (s \nleftarrow P))$$

$$\sim((s \leftarrow P) \wedge (s \leftarrow \overline{P}))$$

$$\sim(E(\downarrow(s \leftarrow P)) \wedge E(\downarrow(s \nleftarrow P)))$$

[where E is the predicate of existence, $\downarrow$ is the operator "the fact that ...", and $\overline{P}$ represents the complementer of P]

$$\sim(E(\downarrow(s \leftarrow P)) \wedge E(\downarrow(s \leftarrow \overline{P})))$$

All these formulations are true by logical and not by empirical reasons. Therefore, it makes no sense to search for empirical facts by which the principle of contradiction would be rejected. The validity of the principle of contradiction is a necessary presupposition for any reasonable human communication. Nevertheless some dialetheists, for instance G. Priest, believe that they have found true logical contradictions by the descriptions of empirical changes.

Priest [5] distinguishes the following three types of changes: Let us suppose that before a time t_0 a system S is in a state S_0, and after t_0, S is in a state S_1. Accordingly, at t_0 it changed from S_0 to S_1. What state was it in at t_0? There are three possibilities:

α) S is in exactly one of S_0, S_1.

β) S is neither in S_0 nor in S_1.

γ) S is both in S_0 and in S_1.

Priest calls these three kinds of changes a type-α, a type-β, and a type-γ change, respectively.

Pries starts from the right presupposition that not all changes are type-α changes. Under this presupposition, he wants to show that there are at least some type-γ changes. Suppose, a change from p to $\sim p$ is not a type-α change. Now Priest wants to show: If this change is a type-β change, it will be a type-γ change too. He argues in the following way: If a change from p to $\sim p$ is a type-β change, then at the instant of change the statements p and $\sim p$ are both false. Since p is false, $\sim p$ is true, and since $\sim p$ is false, $\sim\sim p$ is true. Hence, both $\sim p$ and $\sim\sim p$ (and therefore p) are true and we have a type-γ change. And Priest believes that he has found a true logical contradiction.

If we use the non-traditional predication thory for describing the changes, we can accept the existence of type-β changes without accepting the existence of contradictory type-γ changes.

If we use p for the sentence "S is in the state S_0" and $\neg p$ for the sentence "S is in the state S_1", we can describe type-β changes by $\sim p \wedge \sim\neg p$. Because in non-traditional predication theory p does not follow from $\sim\neg p$, we do not get the contradictory type-γ change $\sim p \wedge p$. Priest has got the contradiction only because he used the non-valid rule $\sim\neg p \supset p$.

Priest gives also another argumentation for his statement: If there are type-β changes, then there will be type-γ changes. He uses the following valid formula:

$$\sim(p \vee \sim p) \equiv (\sim p \wedge \sim\sim p).$$

The left-hand side appears correctly describing the situation at the instant of a type-β change from p to $\sim p$, and the right-hand side describes the situation at the instant of a type-γ change from $\sim p$ to $\sim\sim p$. It seems that type-γ changes and type-β changes hold or fail together.

Using the non-traditional predication theory, we describe type-β changes by $\sim p \wedge \sim\neg p$ (or $\sim(p \vee \neg p)$) and type-γ changes by $p \wedge \sim p$ or by $p \wedge \neg p$ or by $\neg p \wedge \sim\neg p$. But in non-traditional predication theory the following formulas are not valid:

$$\sim(p \vee \neg p) \equiv (p \wedge \sim p)$$

$$\sim(p \vee \neg p) \equiv (p \wedge \neg p)$$

$$\sim(p \vee \neg p) \equiv (\neg p \wedge \sim\neg p)$$

We have seen that dialetheists did not found true logical contradictions, but made logical mistakes. The statement of dialetheism is simply logically false. (See [11].)

5. EXISTENCE AND PREDICATION

If I say to my son: "The peonies in my cabinet have a wonderful red colour", and he answers me: "No, this is not the case", then he is right and makes a true statement. The reasons for rejecting my phrase may be different. It may be the case that there are some peonies in my cabinet, but they are white, or it may be the case that there are some red peonies in my cabinet, but their red colour is not wonderful. But I have in mind another case. My son is a reasonable young man, but he does not know at all the truth-gap-theory of modern logic, therefore he rejects my statement because there are no peonies in my cabinet. Nevertheless, his rejecting statement "It is not the case that the peonies in your cabinet have a wonderful red colour" is true. If I would say under the same circumstances to my son "The peonies in my cabinet do not have a wonderful red colour", he also could reject "It is not the case that..." and his assertion would be true.

This example shows: there is a difference between external and internal negation in natural languages. An elementary statement of the form $s \leftarrow P$ or $s \not\leftarrow P$ can only be true if the subject term s denotes an object. If we have a statement of the form $s \leftarrow P$ or $s \not\leftarrow P$ and its subject term s is empty, this elementary statement will not be true.

In the truth-gap-theory of modern logic, which stems from Frege, such elementary sentences are not ascribed a truth value at all. – Frege wrote:

> "Wenn man etwas behauptet, so ist immer die Voraussetzung selbstverständlich, dass die gebrauchten einfachen oder zusammengesetzten Eigennamen eine Bedeutung haben. Wenn man also behauptet, 'Kepler starb im Elend', so ist dabei vorausgesetzt, dass der Name 'Kepler' etwas bezeichne; aber darum ist doch im Sinne des Satzes 'Kepler starb im Elend' der Gedanke, dass der Name 'Kepler' etwas bezeichne, nicht enthalten. Wenn das der Fall wäre, müsste die Verneinung nicht lauten
>
> 'Kepler starb nicht im Elend',
>
> sondern
>
> 'Kepler starb nicht im Elend, oder der Name 'Kepler' ist bedeutungslos'.

Dass der Name 'Kepler' etwas bezeichne, ist vielmehr Voraussetzung ebenso für die Behauptung

'Kepler starb im Elend'

wie für die entgegengeseztte." ([3], p. 54-55.)

On Frege's quotation I would like to make the following remarks:

1) If the subject term of an elementary sentence is empty, in Frege's opinion the sentence will be referenceless (bedeutungslos), i.e., it has neither the value t nor the value f. In my opinion, such a sentence with an empty subject term cannot be true, but it must not be without a truth value.

2) Frege does not distinguish between the external and the internal negations, i.e., he identifies this two different kinds of negation. In our predication theory with the two kinds of negation we can ascribe the value f to elementary sentences with empty subject terms without contradictions.

3) Frege excludes elementary sentences with empty subject terms from his logic at all, i.e., he does not have the possibility in his logic to handle sentences with empty subject terms. In the Fregean logic, we can neither assert nor reject such sentences with empty subject terms.

Inspired by Frege, A. N. Prior constructed a modal sentential logic, called system Q, which permits truth value gaps. (See [6].)

In the book by Imre Ruzsa we find a further development of this idea in all logical details. In the semantics of this modal logic we have three truth values 0, 1, and 2, where 0 and 1 stand for the usual truth values 'false' and 'true' whereas 2 stands for 'unstatable', i.e., 2 represents the truth-value-gap. [8]

The semantical rules of the system are selected in such a way that a compound formula will have the value "unstatable" if a part of this formula has the value "unstatable". As a consequence of this convention, in the object language of such a logic we cannot reject an unstatable sentence, because the negation of an unstatable sentence is unstatable too.

In the case of sentences with empty subject terms we can ascribe the value f as well to $s \leftarrow P$ as to $s \not\leftarrow P$ accordingly to our predication theory. Besides that we can reject in our object language such sentences with the help of external negation.

6. VAGUE PREDICATES AND NON-TRADITIONAL PREDICATION THEORY

There are different approaches to the problem of vague predicates. I refer to the papers of M. Dummett [1], K. Fine [2], L. Zadeh [12], R. Parikh [4], and H. Putnam [7]. It is not possible to analyse here all these approaches in detail.

M. Dummett has stated that vague predicates like "small", "red", "heap", "possible in practice" are not predicates in the sense of Frege.

R. Parikh makes the proposal to reconstruct a language with vague predicates as a locally consistent but globally inconsistent language. In such a language proofs which are short enough never lead to a contradiction, but long proofs can not be trusted.

Kit Fine has made the proposal that statements about the predicate "red" be taken true iff they hold for all possible ways of making "red" precise, i.e., for all predicates which classify as red those colours that we all agree red, and which classify as non-red the colours that we all agree are not red. The colours about which we are uncertain are also classified as red or not-red, but differently in different precisification.

H. Putnam has proposed to use the intuitionistic logic in argumentations with vague predicates.

I do agree with all the above mentioned authors in the point that vague predicates are not accidental, but that they necessarily occur in every observational language.

I think that the proposed solutions of these authors are not satisfying. In my opinion, the non-traditional predication theory can help to solve some problems connected with vague predicates. Take this for an example.

Imagine, there is a long line of men which are ordered in such a way that at the first place is a man with 0 hairs, at the second place is a man with 1 hair, at the third place is a man with 2 hairs, ..., at the n-th place is man with n-1 hairs and at the end of the line is a man with plenty hairs. In such a case, the following inference seems to be correct:

1) A man with 0 hairs is bald.

2) $(\forall n)$(If a man with n hairs is bald, then a man with n+1 hairs is bald).

3) $(\forall n)$(A man with n hairs is bald).

The sentence 1) is true, the sentence 3) is false, also the sentence 2) must be false, i.e., the sentences

$\sim(\forall n)$(If a man with n hairs is bald, then a man with n+1 hairs is bald),

and

$(\exists n)$(A man with n hairs is bald, and a man with n+1 hairs is not bald).

But this sentence contradicts our intuitions with vague predicates.

If we use the non-traditional predication theory, the paradox does not emerge, because we have the possibility to use besides of $s \leftarrow P$ and $s \nleftarrow P$ in the case of uncertainty the formula $\sim(s \leftarrow P) \wedge \sim(s \nleftarrow P)$.

REFERENCES

[1] M. Dummett, Wang's paradox. In: *Truth and other enigmas*. Cambridge 1978.

[2] K. Fine, Vagueness, Truth, and Logic. *Synthese 30*, 1975.

[3] G. Frege, *Funktion, Begriff, Bedeutung*. Göttingen 1975.

[4] R. Parikh, The problem of vague predicates. In: R.S. Cohen, M.W. Wartofsky, eds., *Language, Logic, and Method*. Dordrecht, Boston, London, 1983.

[5] G. Priest, To be and not to be:Dialectical tense logic. *Studia Logica* 1982, 2-3.

[6] A.N. Prior, *Time and Modality*. Oxford, 1957.

[7] H. Putnam, Vagueness and alternative logic. In: *Realism and Reason, Philosophical Papers*, Vol. 3. Cambridge 1983, pp. 271-286.

[8] I. Ruzsa, *Modal Logic with Descriptions*. Budapest, 1981.

[9] H. Wessel, Kritische Bemerkungen zur intuitionistischen Logikkonzeption. In: *Frege Conference 1984*, ed. G. Wechsung. Berlin, 1984.

[10] H. Wessel, *Logik*. Berlin, 1986.

[11] H. Wessel, Dialetheismus: Mystik in logischen Gewande. In: *VIII. Internationaler Kongress für Logik, Methodologie und Philosophie der Wissenschaften 17-22. 8. 1987, Moskau, DDR-Beiträge*. Berlin, 1987.

[12] L. Zadeh, Fuzzy logic and approximate reasoning. *Synthese 30*, 1975.

HUMBOLDT UNIVERSITY

BERLIN, GDR.

IMRE RUZSA

ON INTERNAL AND EXTERNAL NEGATION

In the literature of logic, one occasionally meets the view that there are several sorts of negation. In particular, some authors claim that we have to distinguish internal and external negation. Such a view is represented in the papers of Professor Horst Wessel and Dr. Klaus Wuttich (in this volume). In these essays, the mentioned point of view is embedded into the framework of the so-called Non-traditional Predication Theory (NPT, for short).[1]

The essence of NPT consists in distinguishing two sorts of negation. Classical (or external) negation '$\sim$' applies to any sentence whereas internal (inner) negation, denoted by '$\neg$', (a symbol borrowed from intuitionistic logic) only applies to atomic sentences of form "$P(s)$" where P is an n-place predicate and s is an n-tuple of "subject terms" ($n \geq 1$).[2] Now, $P(s)$ and $\neg P(s)$ are assumed to be contraries: not both could be true, but both could be false.

My first critical remarks concerning this sort of inner negation are of rather formal nature. As '$\neg$' only applies to atomic sentences, the uniformity of the grammatical rules of logic perishes. (The application of other logical symbols is regulated only by syntactic type criteria.) Further, the notion of atomic sentence is vague in natural languages, and hence, the applicability of the internal negation depends on the depth of the logical analysis of a sentence. Moreover, the truth value of "$\neg P(s)$" is not uniquely determined by that of "$P(s)$" (nor by the meaning of "$P(s)$") – a deviation from classical logic.

Of course, one may counter these objections by saying that the gains of using NPT amply compensate a loss in formal uniformity. If there are serious problems in logic only solvable by the introduction of this sort of internal negation, we must not hesitate to include it into the logical theory.

Let me note, by the way, that, in a sense, classical negation is a sort of internal negation. For, in first-order logic, a negation sign standing in front of a compound formula always can be put farther in (assuming we use the full list of logical symbols: $\forall$, $\exists$, $\&$, $\vee$, $\supset$, $\equiv$, $\sim$) until finally '$\sim$' will occur only in front of atomic formulas. (This holds even in first-order modal logic.) And, if p is atomic, then "$\sim p$" is expressible in natural language by putting a negative particle inside p – producing a paradigmatic form of the internal negation of NPT.

SEMANTIC VALUE GAPS

Of course, the starting point of NPT is that sentences of forms

"*s* is not *P*" and "it is not the case that *s* is *P*"

are not synonymous. According to NPT, the case that neither "*s* is *P*" nor "*s* is not *P*" holds is possible. Prof. Wessel mentions two convincing cases of this possibility (his other cases will be treated later on):

(i) The subject term *s* is empty, i,e., without denotation.

(ii) The meaning of the predicate *P* is undefined for the subject *s* (Wessel's example: 'The moon is honest').

Both cases belong, obviously, under the phenomenon *semantic value gap*. In (i), the term *s* is without denotation; in (ii), the semantic value of the predicate *P* is a partial function undefined for some objects (belonging to the universe of discourse). As quoted by Prof. Wessel, Frege stated that a semantic value gap is hereditary from a part of a sentence to the whole. Then, the mentioned sources of semantic value gaps lead to truth value gaps of sentences.

Prof. Wessel does not accept Frege's view on this point. According to him, in the case of (i) or (ii), both "*s* is *P*" and "*s* is not *P*" are false, and their external negations are true. And, he argues, this is a better solution than accepting truth value gaps, for in the latter case we are unable to refute the mentioned sentences in our object language. He is right *provided our object language is a poor one.*

In this debate, I vote for Frege. (But not because of his prestige.) My starting points are the following two assumptions:

(1) The main sources of semantic value gaps are:

(a) individual names without actual denotations,

(b) functors (of any extensional type) whose extensions are partially defined functions.

(2) Semantic value gap is hereditary via *extensional* contexts. (Of course, this needs further precisification with respect to quantification. But this is unimportant in the present discussion.)

Frege stated that in a pure logical language there is no room for semantic value gaps. At this point, I deviate from Frege. For twenty years, I have worked on constructing modal and intensional systems permitting semantic value gaps.

Noting that the sentence functors of classical logic are extensional ones, one can use them in a system with semantic value gaps without much ado. Their classical semantic rules remain unchanged – one has only to take into account principle (2) which, in this case, says that a truth value gap is hereditary via the functors ~, ⊃, &, etc.

The introduction of semantic value gaps into first-order logic seems to me of small importance. The very usefulness of semantic value gaps comes to light at the presence of intensional functors. Hence, it was quite obvious to begin the work in modal logic – as initiated by A. N. Prior.

Applying value gaps semantics – VGS for short – to languages with intensional functors it becomes necessary to introduce another negation which does not transmit truth value gap. Thus, I agree with Prof. Wessel in that we need two sorts of negation. If you will, they could be called *internal* and *external* negation, respectively. However, my internal negation is the classical one (denoted by '~') which transmits the truth value gap and which might be called 'internal' in a sense, as argued earlier. Using 'NON' for the 'external' negation, its semantic rule is as follows:

If p is true, NON(p) is false. In other cases (i.e., if p is false or without a truth value), NON(p) is true.

Thus, NON is the symbolic counterpart of 'it is not the case that' or 'it is not true that'. By its use, one can refute unstatable sentences as well. Hence, the charge that in VGS, one is unable to refute sentences without a truth value has lost its power.[4]

Up to this point, NPT and VGS seem to be of equal force, and it seems to be a matter of taste to choose between them (if one disregards the formal inelegances of NPT mentioned at the beginning of this paper). Now, my vote for VGS is seriously motivated by some claims of linguists who are not fully satisfied with the handling of some problems in the framework of logic – including Montague's intensional logic. Many linguists state that semantic value gap is a real phenomenon in the use of language – including truth value gaps of sentences as well.[5] Furthermore, VGS is able to give a fine distinction of meanings standing very near both to the intuitive and to the linguistic use of meaning difference. Consider, e.g., the pair of formulas:

(3) $F(a,b) \vee {\sim}F(a,b)$ (4) $F(b,a) \vee {\sim}F(b,a)$

According to classical logic, both formulas are tautologies, true in all "possible worlds", and hence, their meanings are not distinguishable. But what about the particular cases (5) and (6) of these formulas?

(5) John visits or does not visit the House of Parliament.

(6) The House of Parliament visits or does not visits John.

If you deny the synonymity of (5) and (6) then you rightly claim that logic must not judge (3) and (4) as synonymous for pure logical reasons. Now, VGS fulfils this demand: it is possible that in some worlds the two-place predicate F is defined for the pair $\langle a,b \rangle$ but is undefined for the pair $\langle b,a \rangle$ and in such a case, (3) is true but (4) is without a truth value. Hence, the meanings of (3) and (4) need not coincide on purely logical grounds. (Two extensional expressions are per definitionem logically synonymous iff their 'extensions' coincide in all possible worlds of every interpretation.)

Further examples: If 'raven' and 'black' are partially defined predicates, then the following sentences are not synonymous:

Ravens are black (things).

Not-black things are not-ravens.

Under analogous assumptions, the following sentences (proved to be tautologies in classical logic) might be false and are not synonymous:

Every boy is either a student or not a student.

Every student is either a boy or not a boy.

THE INTUITIONISTS' SIN

According to Prof. Wessel, the third case in which neither "s is P" nor "s is not P" is true occurs when "it is impossible to state, whether $s \leftarrow P$ or $s \nleftarrow P$ applies". Illustrations are Brouwerian sentences about the decimal development of the real number π. Nobody knows, e.g., whether the decimal development of π involves a succession of zeros of 10^{10} length. Probably, but not surely, we never will know it. However, it is clear that this sort of example confuses a statement p with "it is (currently) unknown whether p".

Here we are touching on intuitionistic (or constructivistic) mathematics. The intuitionistic calculus looks like a formal system of logic. However, according to its original interpretation (related to intuitionistic mathematics), it is really a formal theory of universally solvable (mathematical) problem patterns. As to its correct interpretation, the following example might be instructive. As one may guess, it is not the case that if p is an arbitrary (mathematical) problem, then either it is solvable or else its solvability can be reduced *ad absurdum*; hence, "$p \vee \neg p$" must not belong to the list of universally solvable problem patterns (in misleading formulation: it is not a theorem of "intuitionistic logic").[6]

In the light of these observations, Wessel's statement that "intuitionistic negation is a confusion of external and internal negation" does not seem to be well-founded. If p stands for a problem, the intuitionistic "$\neg p$" stands for the problem to show the unsolvability of p. (Or, if p means the provability of a statement, then "$\neg p$" means the refutability of the same statement - another interpretation of the intuitionistic calculus.) Intuitionistic "negation" is not a sentence negation at all (neither external nor internal) in accordance with the fact that intuitionistic "logic" is not a logic at all (in the proper sense of the word). (However, I do not want to state that the embedded use of '¬' in intuitionistic calculus is consistent.)

THE MAN WITH SO MANY HAIRS

As the fourth case of applying NPT, Prof. Wessel mentions vague predicates. He treats a case of the paradox of the heap re-worded here as follows:

(P1) A man with 0 hairs is bald.

(P2) For all n: If a man with n hairs is bald, then a man with n+1 hairs is bald.

Consequently:

(P3) For all n, a man with n hairs is bald.

Now, he says, (P1) is true, (P3) is false, hence, (P2) must be false, i.e.,

(P4) For some n, a man with n hairs is bald, but a man with n+1 hairs is not bald

must be true. But, he argues, this contradicts our intuition on the vague predicate 'bald'.

My first remark is that the inference from (P1) and (P2) to (P3) is justified by no system of logic. The correct inference needs a third premise: the principle of mathematical induction which may be worded as follows:

(P0) If the number 0 has a property P, and if this property P is hereditary from each natural number to its successor, then every natural number has the property P.

But (P0) is an accepted principle, is it not? Yes, it is accepted (perhaps as an axiom) in the theory of natural numbers with the obvious restriction that the property P must be expressible in the language of arithmetic. Clearly, the property spoken of in the inference above [λn(a man with n hairs is bald)] is hardly expressible in the language of arithmetic.[7]

I think the moral of this "paradox" is, Do not apply mathematical induction to vague predicates. I would like to increase the severity of this principle as follows: Do not apply mathematical induction to empirical predicates.

If the principle of mathematical induction is assumed to belong to mathematics but not to the realm of pure general logic, fuzzy predicates are harmless with respect to the relation of logical consequence. No remedy is needed.

The fifth case of useful applications of NPT is, according to Prof. Wessel, a sentence involving the presupposition of another sentence. Wessel's example: "N. has stopped beating his wife." Its presupposition is "In the past, N. (always/sometimes /often) had beaten his wife". If this presupposition is false, then both the former sentence and its internal negation are not-true, that is, in Wessel's view, both are false.

Let us note that in this case, the presupposition can be included into the logical reconstruction of the sentence. Let p stand for "N. beates his wife", and let q stand for Wessel's sentence. Then:

$$q = P^{+}(p) \& \sim p$$

where $P^{+}(p)$ is true at a moment t iff for some t', p is true in the interval $[t',t)$ (closed to left and open to right). And the "inner negation" of q (i.e., "N. has not stopped beating his wife") will be

$$P^{+}(p) \& p.$$

(Of course, another reading of q is possible as well. But the formulation just given is sufficient to show that q is not an atomic sentence.)

Now, if the presupposition "$P^{+}(p)$" is false, then it may occur that both q and its "inner negation" are false. But, as we have seen, all this is expressible in the frames of a suitable tense logic without any contradiction.

However, there are cases in which the presupposition of a sentence cannot be formulated in the above way. Examples are the epistemic sentences of form

(7) a knows that p.

In most epistemic theories, it is assumed that the presupposition of (7) is the sentence p. This can be expressed by the meaning postulates

(8) (a knows that p) $\supset p$,

(9) $\sim$(a knows that p) $\supset p$.

Here the antecedent of (9) is assumed to be the inner negation of the antecedent of (8), i. e., it is to be spelled as "a does not know that p".

Now, in classical logic, the following inference pattern is sound:

(10) $\{A \supset p, \ \sim A \supset p\} \vdash p.$

Then, accepting all meaning postulates of forms (8) and (9) compels us to accept all sentences whatsoever.

To avoid this paradox, K. Wuttich abandons classical logic as the logical frame of his epistemic theory (see his paper in this volume). Instead, he chooses a somewhat complicated "axiomatic" version of NPT. Let us see how this problem is solvable in VGS in a natural way.

Accepting (8) and (9), both sentences (11) and (12) imply the truth of p:

(11) a knows that p

(12) a does not know that p.

Permitting semantic value gaps, the paradox consequences of our meaning postulates (8) and (9) are avoidable by the following third meaning postulate:

(13) If p is *not true* (i.e., if p is false or without a truth value) then "a knows that p" is without a truth value.

Then, by the hereditariness of semantic value gaps, we have:

(14) If p is *not true* then "a does not know that p" is without a truth value.

Furthermore, if p is not true, then both (8) and (9) are without a truth value. What does it mean, then, in VGS to accept them as meaning postulates? It means to accept them – not as *always true* ones, but only – as *never false* ones (leaving open the possibility that some of their instances may be without a truth value). Introducing the definition

$$\mathrm{VER}(p) =_{\mathrm{df}} \sim\mathrm{NON}(p),$$

we can somewhat strenghten them:

(8′) $(a \text{ knows that } p) \supset \mathrm{VER}(p)$

(9′) $\sim(a \text{ knows that } p) \supset \mathrm{VER}(p)$.

What about the classical inference pattern (10)? It remains sound in VGS in a weakened sense: if both premises are *true*, the conclusion *cannot be false* (although it can be without a truth value). Thus, we have shown that the general acceptance of the non-falsity of (8) and (9) [or even of (8′) and (9′)] in VGS does not lead to a paradoxical consequence. Furthermore, we can express the external negation of (11) and (12) by our operator 'NON'. Of course, neither "NON(11)" nor "NON(12)" implies p.

SUMMARY

I can summarize my impressions about NPT as follows. Non-traditional predication theory offers an *ad hoc* solution of some logical problems unsolvable in classical first-order logic. But its remedy seems to treat only the symptoms. I think the right treatment of such puzzles must begin with a deeper analysis of the problem (both at syntactic and at semantic level), and the solution is to be searched in a suitable enlargement of classical logic (tensed intensional logic, value gap semantics etc.) rather than in a contrived system deviating from classical logic. (This also applies to the problem of change formulated very roughly by the dialetheists and also solved very roughly by Wessel.) At least, I think, this is the way we should proceed.

NOTES

[1] NPT originates from A. A. Zinov'ev, see, e.g., [9] and [10]. A more exact version of the theory is due to H. Wessel; cf. [5] and [7].

[2] Strictly speaking, this is not quite so. Zinov'ev combines '$\neg$' with quantifiers as well, differentiating, e.g., "$(\neg \forall x)A$" and "$\sim(\forall x)A$"; and Wessel follows him, although in [7] he only mentions this possibility. Furthermore, K. Wuttich applies '$\neg$' to sentences of form "a knows that p", in contrast to the fact that the predicate 'knows' is not a primitive one in his system (but introduced by definition).

[3] Some natural language expressions may yield new sources of semantic value gaps. E. g., we can assume that a sentence of form "a knows that p" is without a truth value whenever the sentence p is *not true* (i.e., false, or without a truth value). Cf. the section 'PRESUPPOSITIONS' of this paper.

[4] However, Prof. Wessel rightly states that in my modal systems presented in [4], unstatable sentences cannot be refuted. In this book, I followed literally Prior's intention according to which even modal operators transmit truth value gap. Let me note that although this book appeared in 1981 its manuscript was finished already in 1977. Not later than 1979 I faced the fact that semantic value gaps are not hereditary via intensional contexts. Then I have constructed a new modal semantics with value gaps in which modal operators do not transmit truth value gaps. If "$\Diamond p$" stands for "it is possible that p", then Prior's original possibility is expressible by

$$(p \supset p) \ \& \ \Diamond p$$

which shows that Prior's possibility is somewhat stronger than (and is expressible by) the one not transmitting truth value gap. Due (partly) to difficulties in publishing, most of my results remained *intra muros* in our Department. (Some Hungarian publications now are in print.)

Concerning 'NON', I mentioned it in some Hungarian publications 1985. Of course, its introduction needs no logical talent: whenever you meet the problem, the solution is obvious. What was a problem for me was whether 'NON' is definable by means of the primitives of (my system of) intensional logic. The answer turned out to be NO.

[5] I am in the comfortable position of being able to refer to the first pages of the nice paper of Sebastian Löbner in this volume. Concerning semantic value gaps, his standpoint is nearly the same as that of mine.

[6] Of course, this interpretation is not my invention. See, e.g., Kolmogorov [2]. Kurt Gödel construes the intuitionistic calculus as a (fragment of a) system of provability (Beweisbarkeit), cf. [1].

[7] I must add: or in the language of any branch of mathematics. For as it is known, the principle of mathematical induction is sometimes used, e.g., in geometry. For example, one finds in textbooks the proof "by ind-

uction on n" of the theorem that the sum of the inner angles of a poligon with n edges $(n > 2)$ is $(n-2) \cdot \pi$. Of course, this proof belongs to a combined system of geometry and arithmetic.

REFERENCES

[1] K. Gödel, Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematischen Kolloquiums 4* (1933), 39-40.

[2] A. Kolmogoroff, Zur Deutung der intuitionistischen Logik. *Mathematische Zeitschrift 35* (1932), 58-65.

[3] S. Löbner, The conceptual nature of natural language quantification. *This volume,* 81- 94.

[4] I. Ruzsa, *Modal Logic with Descriptions.* Akadémiai Kiadó and M. Nijhoff, 1981.

[5] A. Sinowjew/H. Wessel, *Logische Sprachregeln. Eine Einführung in die Logik.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1975.

[6] H. Wessel, Non-traditional predication theory and its applications. *This volume,* 231-239.

[7] H. Wessel, *Logik.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1984.

[8] K. Wuttich, An epistemic logic without normality presuppositions. *This volume,* 209-220.

[9] A. A. Zinov'ev, *Kompleksna'a logika.* Nauka, Moscow, 1970.

[10] A. A. Zinov'ev, Kompleksna'a logika. In: *Issledovani'e logicheskih sistem,* Nauka, Moscow, 1970, 166-203.

DEPARTMENT OF SYMBOLIC LOGIC AND METHODOLOGY OF SCIENCE

L. EÖTVÖS UNIVERSITY, BUDAPEST, HUNGARY

[illegible]

REFERENCES

[1] K. Gödel, Eine Interpretation des [illegible] Aussagenkalküls, [illegible]

[2] [illegible]

[illegible]

[6] [illegible] non-traditional predication theory and its [illegible]

[illegible]

[10] [illegible] Nauka, Moscow, 1970, [illegible]

DEPARTMENT OF [illegible]

THE AUTHORS

AUWERA, JOHAN VAN DER * University of Antwerp (UIA), Germaanse Filologie, Universiteitsplein 1, B - 2610 Wilrijk, Belgium

CASADIO, CLAUDIA * Institute of Philosophy, University of Chieti, (private address) Via Villa Glori n. 1, 48100 Ravenna, Italy

HUNYADI, LÁSZLÓ * Department of Slavic Philology, L. Kossuth University, H - 4010 Debrecen, Pf. 10, Hungary

KÁLMÁN, LÁSZLÓ * Institute of Linguistics (of the Hungarian Academy of Sciences), H - 1250 Budapest, Pf. 19, Hungary

KAŃSKI, ZBIGNIEW * Institute of English and General Linguistics, University of Silesia, Bando 10, 41-205 Sosnowiec, Poland

LÖBNER, SEBASTIAN * Universität Düsseldorf, Seminar für Allgemeine Sprachwissenschaft, Universitätstrasse 1, 4000 Düsseldorf 1, Bundesrepublik Deutschland (Federal Republic Germany)

MADARÁSZ, ANNA * L. Eötvös University, Department of Symbolic Logic, H - 1446 Budapest, Pf. 394, Hungary

MARÁCZ, LÁSZLÓ * University of Groningen, Institute for General Linguistics, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands

MEY, SJAAK DE * University of Groningen, Institute for General Linguistics, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands

PEREGRIN, JAROSLAV * ORGATEX, M. Gorkého 1023, 500 02 Hradec Králové, CzechoSlovakia (ČSSR)

PÓLOS, LÁSZLÓ * L. Eötvös University, Department of Symbolic Logic, H - 1446 Budapest, Pf. 394, Hungary

RANTA, AARNE E.J. * University of Stockholm, Department of Philosophy, Universitetsvägen 10 D, S - 106 91 Stockholm, Sweden

RANTALA, VEIKKO * University of Helsinki, Department of Philosophy, Unioninkatu 40 B, SF - 00170 Helsinki, Finland (Suomi)

RUZSA, IMRE * L. Eötvös University, Department of Symbolic Logic, H - 1446 Budapest, Pf. 394, Hungary

SMIRNOVA, ELENA D. * Moscow State University, Department of Logic, 117234 Moscow, USSR

STRIGIN, ANATOLI * (Akademie der Wissenschaften der DDR), Zentralinstitut für Sprachwissenschaft, Prenzlauer Promenade 149-152, DDR - 1100 Berlin.

SZABOLCSI, ANNA * Institute of Linguistics (of the Hungarian Academy of Sciences), H - 1250 Budapest, Pf. 19, Hungary

WESSEL, HORST * Humboldt Universität zu Berlin, Sektion Marxistisch-leninistische Philosophie, Bereich Logik, Unter den Linden 6, DDR - 1080 Berlin

WUTTICH, KLAUS * Humboldt Universität zu Berlin, Sektion Marx.-len. Philosophie, Bereich Logik, Unter den Linden 6, DDR -1080 Berlin

684

Felelős kiadó: Dr. Pölöskei Ferenc, az ELTE BTK dékánja

Készült a CHEMOLIMPEX Házinyomdájában, 480+20 példányban

Kiadói Főigazgatóság e.., sz.: 59942

ISBN 963 462 238 0

PROCEEDINGS OF THE '87 DEBRECEN SYMPOSIUM ON LOGIC AND LANGUAGE