PROPOSITIONAL ATTITUDES IN MODEL-THEORETIC SEMANTICS*

Roger VERGAUWEN

1. *Introduction: Linguistic and Logical Semantics*

Richard Montague's model-theoretic approach to the semantics of languages differs considerably from Chomsky's views on language-theory. According to Chomsky a generative grammar must express the speaker-hearer's knowledge of the language, i.e. his linguistic competence. The brand of semantics accompanying this theory is usually called 'linguistic' and it is aimed at the description of meaning by giving semantic representations of words, word-groups, and sentences. According to J.J. Katz (Katz '72) these semantic representations are built up componentially from basic 'semantic markers'. The markers are to be seen as the representants of concepts with a certain psychological content and it is supposed that with a limited set of them it is possible to give an analytical definition of words. The use of markers in the semantic metalanguage is, however, very problematic. It is far from clear how the elements of this 'markerese', as it was at one time called by D. Lewis, could characterize 'meaning' in any good sense of the word. It is said that a string of semantic markers is not a string of English words, but a collocation of representatives of concepts. As such it is only more unintelligible, unaccompanied as it is by any rules of interpretation (Vermazen '67, 355-356). Explaining semantic properties of natural languages in terms of such a semantic metalanguage is in fact something like an 'ignotum per ignotius'. Opposed to this, there is another tradition in semantics, called 'logical'. This semantics is part of semiotics and metalogic and studies systems of semantic rules. These rules interpret logical calculi by

(*) I want to thank A. BURMS, W. de PATER, F. DROSTE, O. LEYS and H. ROELANTS, who read the manuscript, for their interesting comments and helpful discussion.

assigning denotations to the primitive terms and truth-conditions to the sentences in such a way that it is immediately clear how the meaning of a compound depends upon the meanings of its parts. For many years the logical and linguistic paradigm in semantics have existed side by side without much interference. It was supposed that natural languages were too 'vague' and 'ambiguous' to be studied by the methods used in logical theory. Only the last few years, under the influence of the work of the late Richard Montague, several linguists and logicians became interested in applying his thories to the analysis of natural languages. Montague is not convinced of a fundamental dichotomy between logical and natural languages: "There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians. Indeed, I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory" (Montague '70a, 222). Moreover, the aim of the semantic part of the theory when applied to natural languages, is the same as that commonly found in the study of formal logical language: "The basic aim of semantics is to characterize the notion of a *true sentence* (under an interpretation) and of *entailment*" (Montague '70a, note 2). The goal of this paper, then, is twofold. First, I want to show that there are serious problems in applying such a theory to natural languages. More specifically I will be concerned with fixing the semantic object of 'propositional-attitude verbs' in such a theory. It is well known that with such verbs as 'believe', 'think'... the sentences that serve as their complements cannot be substituted 'salva veritate' even when these sentences are 'logically' equivalent. Moreover, it seems that in the contexts of propositional attitudes 'rigid designators' seem to lose their 'rigidity'. As will be shown, on the one hand Montague's semantics seems to make the wrong predictions in certain contexts (entailment in contexts of propositional attitude), and on the other hand it doesn't match our intuitions regarding the use of certain terms in language (rigidity).

Second, and most important, I would want to say that, unlike some authors for whom the existence of these problems in Montague-semantics is a reason either to dismiss the theory or turn it into something more 'psychological', it could be possible to solve them within the model-theoretic framework itself. I will propose a possible

solution at the end of the paper. Having summarized the problems, let us now turn to a discussion of what exactly they are.

2. *Possible-Worlds and Logical Equivalence.*

In Montague's ontology, an *interpretation* (for intensional logic) can be understood as a quintuple (1).

(1) $< A, I, J, \leq, F >$ (Montague '73, 258)

A is a set of (possible) individuals, I a set of possible worlds, J a set of moments of time, and $\leq$ is a linear ordering relation over J.F is a function having as its domain the set of all constants ('words') and assigning them intensions (functions from possible worlds and times to extensions of a certain type). The interpretation of sentences is carried out 'bottom-to-top,. First the lowest constituents (words) are interpreted and then larger units are built up and interpreted compositionally by means of semantic rules that work along the lines of a function/argument relation. According to B. Partee, Montague's logico-mathematical view of semantics contains certain idealizations which are incompatible with some psychological claims about language and the brain. These idealizations are the following: (Partee '79,3)

(i) The objects of Propositional attitudes are propositions.
(ii) The intensions of sentences are propositions.
(iii) The intensions of sentences are compositionally determined, i.e. recursively built up from the intensions of their parts.
(iv) Intensions are functions from possible worlds to extensions.
(v) Words have intensions.

The psychological or, for that matter 'linguistic' assumptions which are incompatible with those are

(vi) people know their language.
(vii) The brain is finite. (Partee '79,3).

This incompatibility is particularly clear in contexts with verbs of propositional attitude, and, if it is true, it could be very devastating for a possible worlds semantics like Montague's: it follows from Monta-

gue's theory that two expressions which have the same extension on all points of reference (possible world/times) do in fact have the same intension and are '*logically equivalent*' (or: 'synonymous' in a certain sense). Now, if two expressions have the same intension, it should be possible in a compound expression to substitute the one for the other without affecting the truth value of the whole (substitution 'salva veritate'). Nevertheless, such a substitutability of logical equivalents is not always possible, as can be seen readily from the following examples. Let us assume that *P* and *Q* are two sentences whose propositions (cfr. ii) are logically equivalent. So, normally it could be expected that they mutually entail one another. This does not come out true because, even when P and Q are logically equivalent, we cannot validity make an inference from (2) to (3) in the context of the verb of propositional attitude 'believe'.

(2) x believes that P.
(3) x believes that Q.

The counterintuitivity of the inference becomes clear if *P* and *Q* are replaced by two sentences.

(4) Thales believed that *two plus two equals four.*
(5) Thales believed that *the square root of two is irrational.*

The sentences that in (4) and (5) function as complements of the verb 'believe' have in all possible worlds the same truth-value (viz. 'true') and thus express the same proposition but (4) is certainly not entailed by (5) or vice versa. This problem is called '*The logical-equivalence problem*', and it is mainly due to the idea that it is propositions that are the objects of propositional attitudes, and that sameness in the intension implies sameness in meaning. It is not so clear what Montague himself thinks about this. On the one hand he seems to be convinced that we should tolerate inferences like (4)-(5) because they are inevitable:'' I should point out, however, that if φ and ψ are logically equivalent sentences (with respect to given analyses f and g), then 'John believes thet φ' and 'John believes that ψ' will turn out to be logically equivalent (with respect to analyses that contain f and g as parts). This may at first appear strange, but it is a conclusion that I believe we should accept'' (Montague '70b, 218). In 'Pragmatics and intensional logic', however, he seems to have some doubts about this

unrestricted interchange of logical equivalents:'' A second objection might concern the fact that if φ and ψ are any logically equivalent sentences, then the sentence $\beta[J, \hat{}\varphi] \rightarrow \beta[J, \hat{}\psi]$ is logically true, through it might under certain circumstances appear unreasonable... its counterintuitive character can perhaps be traced to the existence of another notion of belief, of which the objects are sentences or, in some cases, complexes consisting in part of open formulas'' (Montague '70c, 139). Nevertheless, he remains convinced of the unrestricted substitutability approach because it is the most elegant one:'' But even to those who, like myself, believe that the best and most elegant approach is to permit unrestricted interchange on the basis of logical equivalence, it may be of some interest to learn that this approach has genuine alternative'' (Montague '70a, 231).

Allowing unrestricted substitution of logical equivalents clashes with the assumptions (vi) and (vii). If it is in fact true that people know their language and if they have certain intuitions about what can be entailed by what and about the notion of a valid inference, it is unclear why the inference (4)-(5) is generally deemed to be impossible. Moreover, if the brain is finite it is virtually impossible to acquire the knowledge that is necessary to deduce (5) from (4). In this case it should some how be possible for someone to know all tautologies (whose extension is 'true' in all possible worlds): Montague's system could, then, only be a description of the semantic competence of a speaker who is able to know all the logical consequences and entailments of everything he believes, which is generally impossible for mortal beings: ''The impossibility of logical omniscience for finite beings thus seems to be the reason both for the limits on our knowledge of the semantics of our language and for the failure of substitutivity of logically equivalent sentences in propositional attitude sentences. Another way of putting the point is to view Montague's semantics as a super-competence model, e.g. a semantics for English as spoken by God'' (Partee '82, 90). Sentences like (4) and (5) are a serious problem for Montague's theory. As it stands it looks as if is too 'powerful', it 'predicts too much'. Furthermore, it follows from the discussion that the objects of propositional attitude sentences are not propositions (functions from possible worlds to truth values). Inferences like (4)-(5) cannot be tolerated in our semantics, even if unrestricted substitutability is an 'elegant approach', because they go

counter to our intuitions. But if it is not propositions we are looking for, what then are the semantic objects of propositional attitudes? Most probably, the problems are due to our considering the proposition as an unstructured entity without internal cohesion. Of the idealizations listed in the beginning of this part, there may at least be one which is helpful in finding an adequate solution. Idealization (iii) says that propositions are structured entities built up compositionally from the intensions of their parts. With this in mind, let us now look at a possible solution for our problem in which the internal structure of the proposition is taken into account.

3. *David Lewis and why 'meaning' is not enough*

In his 'General Semantics' (Lewis '76) David Lewis defines 'intention' essentially as we have done it here (as a function from possible worlds/times to extensions) but he proposes to use a second technical notion which he calls '*Meanings*'. Meanings are introduced to be able to ditinguish between sentences that express the same proposition. Roughly, a meaning can be identified with the intension of the whole sentence and the intensions of each of its subconstituents arranged in correspondence to the syntactic structure of the sentence (Dowty' 81, 173). Lewis hopes to avoid problems of logical equivalence by introducing 'meanings': "we have already observed that intensions for sentences cannot be identified with meanings since differences in meaning – for instance between tautologies may not carry with them difference in intension... The same goes for other categories, basic or derived. Differences in intension, we might say, give us coarse differences in meaning. For fine differences in meaning we must look to the analysis of a compound into constituents and the intensions of the several constituents" (Lewis '76, 14). All this can be illustrated by the following sentences.

(6) Missiles are dangerous or they aren't.
(7) Snow is white or it isn't.

Both sentences express the same proposition. They are tautologically true. Though the intensions of (6) and (7) are the same, this is, according to Lewis, not the case for their 'meaning'. There are serious

differences between the intensions of the subconstituents (in this case, the different words) of (6) and (7). It is precisely this difference in meaning that would block substitution salva veritate in (8) and (9).

(8) Ronald believes that missiles are dangerous or they aren't.

(9) Ronald believes that snow is white or it isn't.

There is sameness of intension with the complement sentences of 'believe that' but as there is difference in intension between the lexical constituents there is difference in 'meaning' and therefore substitution is impossible: "only when we come to non-compound lexical constituents can we take sameness of intension as a sufficient condition of synonymy" (Lewis '76, 14). Let us note, by the way, that Lewisian meaning is also an excellent tool for distinguishing between (4) and (5). The complement sentences in (4) and (5) differ considerably in meaning. Not only is there a difference in the intensions of the lexical constituents but also both sentences have different syntactic analyses. So, on the face of it, problems of logical equivalence seem to disappear if we are willing to introduce 'meanings' in our model-theoretic semantics. But it is the strength of Lewis' theory which is at the same time its weakness. Lewis takes into account the structure of the proposition and seems to be clearly on the right track, and at the same time – as appears from the quotation just given – comes up with a notion of lexical synonymy as a decisive condition on what in Montague's terminology could easily be translated as 'logical equivalence'. But Lewis' criterion for synonymy as identity of meaning and thus of logical equivalence is still not rigorous enough. Unfortunately it is possible to construct sentences in which identity of meaning up to the lexical constituent level does nevertheless not allow for substitutivity salva veritate in propositional attitude contexts. The existence of such sentences, to which I will turn presently, is intertwined with idealization (v) from part 2 and gives rise to problems of rigid designation and their consequences for a model-theoretic semantics.

4. *The rigid Designator problem*

In a number of papers Kripke and Putnam (Kripke '80, Putnam '75) have developed a theory of how reference for proper names and

natural kind terms comes about. Their theory, also called the 'causal theory of reference' is set up to show that the relation of 'reference' and meaning (in the general sense of the word) is not to be identified with the existence of a certain 'psychological state' in the mind of the language-users. By means of proper names we can refer to persons we ourselves have never known and who have died long ago (e.g. 'Moses', 'Frege'..). Nevertheless it is possible to refer to these persons because, as the causal theory says, there is a kind of '*causal chain*' that links the names and their referents: ''a rough statement of a theory might be the following: An initial 'baptism' takes place. Here the object may be named by ostension, or the reference of the name may be fixed by a description. When the name is 'passed from link to link', the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it ''(Kripke' 80, 96). Something analogous happens in the reference to natural kinds (terms designating naturally occurring stuffs like 'water', 'gold'...). In this case the 'original baptism' is effected by ostension or by giving a description, e.g. 'Gold is the substance instantiated by the items over there, or at any rate by almost all of them' (Kripke '80, 135). According to Kripke and Putnam proper names and natural kind term are typically what they call '*rigid designators*': rigid designators are terms that designate the same object in every possible world, whenever that object exists in that world. In the terminology of possible worlds semantics we could say that these terms do have intensions but that their intensions are *constant functions*. In the case of proper names this would mean that a name as 'John' in all possible worlds designates the same individual as is designated by that name in this world. With all this in mind, let us examine carefully the following sentences.

(10a) Eye-doctors are eye-doctors.
(10b) Eye-doctors are ocultists.
(11a) John believes that eye-doctors are eye-doctors.
(11b) John believes that eye-doctors are ocultists.
(12a) Hesperus is Hesperus.
12b) Hesperus is Phosphorus.
(13a) The ancient Greeks believed that Hesperus was Hesperus.

(13b) The ancient Greeks believed that Hesperus was Phosphorus.

It is fairly plausible to assume that in (10) the words 'eye-doctor' and 'oculist' are synonymous (Cresswell '82, 69, Chomsky '55, 39). On the face of it, it would be difficult to find a possible world in which the set of individuals denoted by 'eye-doctor' is completely disjoint from the set of individuals denoted by 'oculist' (but, cfr. infra). With respect to (12a-b) some things should be kept in mind. Suppose that in ancient Greece a certain star was seen in the evening and it was called 'Hesperus'. Furthermore, another star was seen in the morning and it was called 'Phosphorus'. Then, in fact, we find that it's not a star, but it is the planet Venus and that Hesperus and Phosphorus are in fact the same. According to Kripke, both proper names are rigid designators, which in all possible worlds refer to the same planet (Venus). The identity expressed in (12b) is a *necessary a-posteriori truth* (Kripke '80, 100n), and therefore true in all possible worlds. But, though all this be true, intersubstitutivity in contaxts of propositional attitudes does not seem to be possible. In (11), it does not follow from the fact that John believes (10a), and notwithstanding the identity (10b), he does therefore believe (10b). Substitutability of the equivalents (10a-b) is simply impossible in (11a-b) without changing the truth values. The same goes for (12-13). The identity between Hesperus and Phosphorus expressed in (12b) and its equivalence with (12a) due to the rigid designation of the proper names, does therefore not justify substitutability salva veritate of (12a) by (12b) in the belief-context (13a-b). Moreover, it is very well possible that rigid designation can be ascribed to such common nouns as 'eye-doctor' or 'oculist', which are in a certain sense 'non-natural kind terms' (this is a still ongoing discussion, cfr. Grayling '82, 185). If this were so, it would be all the more strange that substitution salva veritate is blocked in certain contexts also for these terms. From these examples one could conclude that apparently there are cases where rigid designators lose their rigidity or, conversely, that a very important principle of the semantics has to be given up: it is the principle that states that the intension of a term completely determines its extension (Putnam '75, 219). This would clearly be an undesirable result for a model-theoretic semantics and I don't think we should accept it, but I will come back

to that in the final part of this paper. Now, again with respect to sentences (10-13), it is perhaps not too difficult to point out informally why substitution is blocked there. In fact, it all comes down to the inability of the natural-language speaker-hearer to know everything. In Kripkean terminology we could account for non-substitutability of equivalents by saying that instead of having one unique 'causal chain' leading to the same referent in the 'Hesperus-Phosphorus' case, perhaps with certain people either there are two causal leading chains to different referents or there is no causal chain at all that would for them link certain terms with their referents. The same goes for 'eye-doctor – oculist': to put it even more clearly, it is simply possible that people don't know certain words of their language. If one has never heard or read the word 'oculist' one can hardly know that it is the same as 'eyedoctor'. It is also possible that one 'knows' the *words* but is convinced that 'oculists' have nothing to do with medicine but are members of a bizarre religious sect worshipping the god 'ocul'. What is also more or less useful is Putnam's hypothesis of the universality of the division of linguistic labor (Putnam '75, 228). Applied to our problem, this thesis would say that in every linguistic community there are a subset of the speakers who know certain words with their associated criteria and who are in fact able to detect synonymy between them. All this is nevertheless very unsatisfactory when it comes to applying it to a model-theoretic semantics, even apart from the difficulties that are involved in Kripke's and Putnam's theories. Barbara Partee considers the importance of psychological factors in the discussion. With respect to (12-13) she writes: " 'Hesperus is Phosphorus' is necessarily true but not known apriori. The ancients may have used 'Hesperus' and 'Phosphorus' as rigid designators for the same planet, but they didn't know it. The two names were not psychologically equivalent for them (nor are they for us). Sentence (13b) can have a different truth value from sentence (13a) *because* it involves a psychological modality, applied to a sentence which is (psychologically)informative although metaphysically necessary" (Partee '79, 7). It is not so clear what is meant here by 'psychological modalities'. It can certainly not mean that 'meaning' is somehow 'in the head', because the problems with rigid designation seem to suggest the opposite as Partee would readily agree to (Partee '79, 7). On the one hand we cannot accept the conclusion that people

do *not* know their language, because it would be counterintuitive (cfr. assumption vi in part 2), on the other hand, however, it is all the more clear that if one is to allow unrestricted substitution in propositional attitude contexts Montague's model-theoretic semantics is not able to fulfil its own promises, viz. the characterization of *entailment* in natural language (cfr. part 1). So, there are both good epistemological and logical reasons to find alternatives. The sentences (10-13) show at least one very important thing, if at all we are to look for a serious description of their semantics. They show that Lewis' notion of 'meaning' is not enough in the case of propositional attitude sentences: In (10-13) we have, in fact, counterexamples to his proposals: 10a-b, 12a-b are sentences expressing the same proposition (there is sameness in intension between 10a-b and 12a-b). But there is more. On the face of it, there is also no difference in the internal structure of the proposition. Up to the lexical constituents the propositions seem equivalent (that somehow they are not, can only be detected in contexts like (11a-b) and (13a-b), but to this I will come in the next part), and therefore they also don't differ in 'meaning'. Of course, at this point it would be possible to change tactics and to reconstruct the semantics of propositional attitudes as part of a certain psychological theory, which is one of Partee's proposals: "I agree with Brian Loar that semantics is part of propositional attitude psychology, and stands or falls with it "(Partee '82, 96). This might, then, also mean giving up the framework of a model-theory and repalcing it by some other 'psychological' paradigm. I am not convinced that we should do that, without denying, however, that 'psychological factors' should be taken into account at some point in the theory. Rather, I think, we should examine whether it is not possible to construct a kind of *logical counterpart* within a model-theoretic semantics which would account for just these psychological characteristics, following the lead of Church: "Let us take it as our purpose to provide an abstract theory of the actual use of language for human communication.... we must demand of such a theory that it have a place for all observably informative kinds of communication, including such notoriously troublesome cases as belief statements... or at least that it provide a theoretically workable substitute for them" (Church '64, 438). The treatment for propositional-attitude sentence that I want to propose is an attempt to modify Lewis' notion of 'meaning' in such a way that it

can accomodate sentences like (10-13). It is to be based on the notion of 'supervaluation' developed by van Fraassen (van Fraassen, '66,', '71) and is consistent with a model-theoretic semantics.

5. *Supervaluations and Propositional Attitudes.*

Gochet (in Gochet '80, 182) mentions a proposal by Partee to represent statements of propositional attitude in a certain psycho-logical way that fits nicely with her interest in psychological factors. For a sentence like (14a) she proposes a representation by the following formula (14b).

(14a) Charles believes that Hesperus is Phosphorus.

(14b) ($\exists$i) ($\exists$m) believes (Charles, m (i(Hesperus) = i(Phosphorus)))

"where *m* designates Charles' own psychological performance factors and *i* his idiosyncratic interpretation of lexical items" (Gochet '80, 182-183). It is clear, then, that as a consequence of the relativization of the interpretation of (14a) to the parameters *m* and *i*, the interpretation of the whole may be affected due to considerable variations in *m* and *i*. Such a treatment makes it more or less understandable why equivalents may not be substituted salva veritate in certain contexts. Applied to e.g. sentence 13 this would mean that for the ancient Greeks the belief in the identity expressed in (13a)is not the same as that expressed in (13b) because the Greeks 'didn't know' that Hesperus was Phosphorus, i.e. their psychological performance factors and their idiosyncratic interpretations of the lexical items were different in (13b) from (13a). Nevertheless, such a formalisation remains unclear (what exactly is the *content* of such notions as 'psychological performance factors' and 'idiosyncratic interpretation'?) and is not very useful in a modeltheoretic framework. Let me, therefore, try to formulate an alternative which is, I hope, more useful in these circumstances. In an article in 1966 and in some later publications the American logician van Fraaassen tried to explore the consequences for logic and formal semantics of the view that in some circumstances syntactically well-formed sentences may be neither true nor false. Notoriously, such circumstances are those in which some singular

term occurring in the statement does not have a referent, e.g. in 'The king of France is wise' or 'Pegasus has a white hind leg'. In his discussion he introduces the idea of a *supervaluation*. Let us consider a firstorder language with predicates, names, variables, connectives and quantifiers. Basically, constructing a specific interpretation for this language (call it 'L') means specifying at least two things. First of all the delimitation of a certain *domain of discourse D*. D is also called 'model structure' (Thomason '74, 30) and contains the 'ontological furniture' for the theory, i.e. the individuals one wishes to consider. Moreover, an interpretation for L also requires a function f, called the 'model assignment' (Thomason '74, 30). It assigns values from the domain D to the expressions of L. Suppose, now, that F is a predicate from L. If F is a predicate of L and f(F) is the extension of F in D, then the couple (f, D) is called a *model for the language L*. Truth in such a model is defined in the usual way, e.g. the formula $\forall(x)Fx$ is *true* in the model if all elements from the extension of F are included in the domain. The function of models as just described is essentially the same as the model described in (1).

There is, however, a difference. It is interesting to note that van Fraassen considers 'models' as "the reconstructed counterparts of Leibniz' possible worlds" (van Fraassen '66, 485). This means that for van Fraassen models have the same function as the set I of possible world from (1). This need not bother us too much, the more as it gives us an idea about a further important notion, viz. the idea of a *valuation*. In van Fraassen's terminology, models can be identified with (possible) states of affairs. Such a view leads naturally to a definition of 'valuation': "we can thus identify a state of affairs with a function that associates with each proposition a truth value. The term *valuation* is commonly used for such a function" (Mc Cawley '81, 157). A valuation is, then, an assignment of truth values to all sentences of a language. Valuations are called *classical* if the following conditions are fulfilled:

(15) "a classical valuation over a model is a function *v* that assigns T or F to each statement, subject to:
- if A is an atomic statement containing no nonreferring names, then v(A) is determined by the model in the indicated manner, and,

– if A is a complex statement, then v(A) is determined by what v assigns to the simpler statements, in the usual manner" (van Fraassen '66, 486).

Suppose now that our language L contains exactly one one-place predicate F, and two names *a* and *b*. Suppose the function f *(model-assignment)* is defined for *a*, but not for *b*. If we would consider b to be a non-referring term, this would mean that F(b) would not always have the same value in different valuations (according to whether f(b) belongs to the domain D or not). For L there are exactly two (classical) valuations v_1 and v_2 and a supervaluation s. All this can be illustrated by the diagram (16). Let us assume that $f(a) \in f(F)$ and that all of D is included in f(F).

(16) valuations and supervaluation for L.

	v_1	v_2	s
Fa	T	T	T
$\sim$ Fa	F	F	F
Fb	T	F	–
$\sim$ Fb	F	T	–
Fb $\vee \sim$ Fb	T	T	T
$(\forall x)$ Fx	T	T	T
$(\forall x)$ Fx $\supset$ Fb	T	F	–

(Van Fraassen '66, 487).

It can easily be deduced from (16) that in the supervaluation s the values T(F) are assigned to exactly those sentences to which T(F) was assigned in the valuations $v_1$and v_2. If in the distinct classical valuations distinct values are assigned, then in the supervaluation there appears a truth value gap. In (16) truth value gaps are indicated by means of a dash ('–'). It is to be noted that these gaps are not the representations of what would be a third value in the logic, but merely they indicate the *absence of* truth value for the formula (cfr. Haack '74, 58). The exact definition of a supervaluation follows quite naturally.

(17) "A valuation is a supervaluation for a language L iff there is a nonempty set K of admissible valuations for

> L such that for all sentences A of L;
> s(A) = T iff v(A) = T for all v ∈ K.
> s(A) = F iff v(A) = F for all v ∈ K.
> s(A) is not defined ('–') otherwise" (van Fraassen '71, 95).

I think the use of supervaluations so defined may be of extreme interest in the solution of logical equivalence and rigid designation problems in propositional attitude context. In Montague's theory common nouns like 'eye-doctor' and 'oculist' are semantically defined as *sets*. These sets consist of 'individual concepts' (intensions of individuals) and may be represented in the language of semantic type-theory by ((s, e)t); t is the type for 'truth value, e for 'individual' and s shows that it is about the intensions of individuals e (so, it is not to be confused with the s of 'supervaluation' in (17)). Of course, in natural-language sentences common nouns frequently get combined with quantifiers (determiners) and when this happens their type converts into a set of *properties* of individual concepts. What is important, however, is that sets (e.g. sets of individual concepts) can be represented by their *characteristic function*. Instead of enumerating all members of the set, the set itself is identified with a function which, for every element (individual concept), determines unambiguously whether the element is part of the set or whether it is not. Every individual in the domain is assigned the value True (1) or False (0) according as it belongs to the set or not. In Montague-semantics common nouns typically get the same type as one-place predicates, such as e.g. 'walk' or 'run'. Of course, common nouns are assigned values within the valuations over the models of a natural language. From a purely logical pount of view it is perfectly well possible that several valuations which are mutually different might come into consideration. It is possible that in distinct valuations different objects (or, for that matter, individuals or individual concepts) or sets of objects are assigned via characteristic functions to the same word. Variation of domains in different models may have a serious influence as to what gets attached to what. Montague himself has a passage in which he compares models with dictionaries, thus allowing for variation in meaning: "It should, perhaps, be pointed out that throughout our development models play the role of possible dictiona-

ries, with the one oddity that alternative definitions are never allowed within a single dictionary, but instead involve the consideration of several dictionaries'' (Montague '70b, 210). To put it still otherwise: in different interpretations (or 'intensional models') for the intensional logic it may happen that e.g. the values assigned to the common noun *'oculist'* are different from those assigned to *'eye-doctor'* in such a way that the sets of individuals assigned to the extension of the one are disjoint from the sets assigned to the extension of the other. Let us call the valuations in which this happens *idiosyncratic valuations*. They are to be considered as a logical analogon for the parameter *i* from (14b) ('idiosyncratic interpretations'). Of course in some idiosyncratic valuations it may be that no objects at all are assigned to the nouns just mentioned (someone who doesn't even know the words certainly has no interpretation for them). Rigid designation is another important point. In Montague's 'The Proper Treatment of Quantification in ordinary English' (Montague '73) rigid designation for proper names is assured by the formula (18):

(18) $\exists u \, \Box[u = \alpha]$ where α is j.,m . . . (Montague '73, 263)

In (18), *u* is a variable over individuals, $\Box$ is the necessity operator and j,m are abbreviations for proer names such as 'John', 'Mary'. . . It says that there is an individual that in every possible world is the same as the individual denoted by that name in this world. A formula like (18) is also called a *meaning-postulate*. The aim of meaning-postulates is mainly to delimit the class of possible interpretations for a language to the class of 'intended' or 'logically possible' interpretations: ''Not all interpretations of intensional logic, however, would be reasonable candidates for interpretations of English. In particular, it would be reasonable in this context to restrict attention to those interpretations of intensional logic in which the following formulas are true (with respect to all, or equivalently some, worlds and moments of time)'' (Montague '73, 263).

Most probably, Montague wanted to safeguard rigid designation because otherwise he would almost certainly have had to accept *'counterparts'* in his semantics, which is an additional problem. I will not dwell on the counterpart problem-more about it is to be found in Loux '79 – but the fact that for Montague rigid designation is guaranteed by a meaning postulate is, I think, significant. In his

semantics meaning postulates are that part of the theory that 'could have been otherwise' but is accepted because it is 'reasonable'. Dropping rigid designation in some sense would not be too drastic a modification of the theory. Anyhow, let us now come back to the relation between the several valuations already mentioned and the idea of a supervaluation to see how it can be of any help in solving our problem. It does not seem unreasonable to suppose that language-users have amongst them a certain common interpretation-basis. For people to communicate effectively (and that's what happens most of the time) it is necessary that there is a large part of common knowledge, including knowledge about the meanings of most words. It is highly probable that within a linguistic community the 'knowledge systems' and even the 'belief-systems' of the language-users run parallel, at least in a large part. Supposing this is not so leads in the end to a futile solipsism in which nothing at all can be said about any communicative process whatsoever.

The effectiveness of any communication seems to depend heavily on an already present common basis of knowledge and belief: "if someone's whole belief structure, including beliefs about what many words mean, differs radically from mine, I cannot hope to describe his or her belief in my language. A compositional semantics may be impossible without the assumption of an homogeneous interpretation (both the model theory and the interpretation into it) "(Partee '79, 8). Of course, there are also even large differences in belief and knowledge, which gets the more clear in precisely contexts of propositional attitude. Now, what could all this mean for sentences like (10-13). It has already been said that common nouns like 'eye-doctor' and 'oculist' are semantically to be considered as *predicates* and as such also identifiable with characteristic functions of a certain kind. In montague's theory proper names like 'Hesperus' and 'Phosphorus' are interpreted as property-sets of individual concepts, or characteristic functions thereof. Rigid designation, which makes these proper names the characteristic function for the property-set of the individual they designate in our world, is assured by a meaning-postulate. Let us call the valuations (or 'models') in which these terms are assigned denotations *'as it should be'* (i.e. valuations in which the sets of individuals designated by the common nouns 'eye-doctor' and 'oculist' are not disjoint or in which 'Hesperus' and 'Phosphorus' are

rigid designators for the same planet) *general valuations* (v_g). These general valuations do not necessarily differ completely from the other type of valuation about which we have already been speaking, viz. the set of *idiosyncratic valuations* (v_i). Idiosyncratic valuations may assign different, partly different, or identical valuations as the general ones (v_g). The role and function of idiosyncratic valuations comes to the fore especially in contexts of propositional attitudes. In these contexts they are to be taken into account for the construction of a *supervaluation* as defined in (17). By the end of part 4 of this paper it had become clear that Lewis' notion of 'meaning' was insufficient for logical equivalence and rigid designation, because equivalent propositions with the same syntactic structure could not be handled. Now, I think it is possible to modify Lewis' idea of a 'meaning' and to replace it by something else that can be used in explaining phenomena of non-substitutability of equivalents, viz. *'supervaluated meanings'*. Supervaluated meanings in this sense are something like Lewis' 'meanings' except that it is not the intension of the components of the proposition that is important but the semantical values assigned to them by means of a supervaluation. The following tentative rule for propositional attitude contexts could, then, be formulated.

(19) *rule of supervaluated meanings*

Substitutability salva veritate in propositional attitude contexts is only possible if the elements so substituted have the same supervaluated meaning, or, conversely; If substitutability of equivalents in propositional attitude contexts is impossible salva veritate this is due to a difference in the supervaluated meanings of the elements so substituted.

(19) signifies that the semantic complements of propositional attitudes are not always propositions or sentence-intensions, neither 'meanings' but another type of objects, viz. *'supervaluated meanings'*. Let me illustrate this with the help of sentences (10-11). Suppose that (20) and (21) give a hypothetical representation of the semantic interpretations of the predicates *'eye-doctor (x)'* and *'oculist (x)'* respectively. Let 'x' be a variable over individuals (and not individual concepts, as Montague would have it. This is, though, only an unimportant change in the terminology). $O_1, O_2 \ldots O_n$ are the respective sets of (possible)

individuals (cfr. (1)) in the interpretation(s). If a set of individuals satisfies the predicate it is given the value 1 (or: 'True'), if it does not, it is given the value O (or: 'False'). Once again, v_g represents the 'general valuations'(denotations are assigned 'as they should be') and v_i the 'idiosyncratic valuations' (where different assignments might turn up). In the same way as illustrated in (16) the supervaluation is given by *s*.

(20) *eye-doctor(x)*

	v_{g_1}	v_{g_2} ..... v_{g_n}	v_i	v_s
O_1	1	1	1	1
O_2	1	1	1	1
O_3	0	0	0	0
⋮				
O_n	1	1	1	1

It is clear that, because the value-assignments in v_g and v_i correspond, the same values turn up in the supervaluation v_s. Suppose, now, that for 'oculist(x)' in the general interpretations the same values are assigned to sets of individuals as for 'eye-doctor(x)', but that the idiosyncratic interpretation is very different. Then, of course, this will have its effects in the supervaluation v_s as is illustrated in (21).

(21) *oculist(x)*

	v_{g_1}	v_{g_2}	v_{g_n}	v_i	v_s
O_1	1	1	1	0	–
O_2	1	1	1	0	–
O_3	0	0	0	1	–
⋮					
O_n	1	1	1	0	–

(20) and (21) are, of course only illustrative in character, but it is easily seen that in (21) the supervaluation v_s contains truth-value gaps, illustrated by the dashes '–'.

Comparing (20) to (21), it turns out that in (20) and (21) the sets v_i in (20) are disjoint from those designated by v_i in (21). Nevertheless, in v_g there is no difference in assignment for (20) and (21). That in (21) v_i and v_g differ may be due to the fact that v_i of (21) is the logical counterpart of someone's psychological inability to perceive the identity between (20) and (21) either because he doesn't know the word 'oculist' or thinks that the word means something else (this is, by the way, only a rough sketch of what may psychologically be going or; other elements may also be very important). Sentence (10a) can be formalised by means of the implication (22).

(22) $\forall x$ (eye-doctor(x) $\rightarrow$ eye-doctor(x))

General valuations as well as idiosyncratic valuation (v_i) assign the predicate 'eye-doctor(x)' the same values, as can be seen in (20). Consequently, in the supervaluation, the same value is assigned as in v_g and v_i. There will be no truth-value gap and the implication (22) is assigned the value 1 (true). In the propositional-attitude context (11a) rule (19) is applied, but the value of the implication remains unchanged. It is a tautology and is assigned the value 'True'. As for sentence (10b), it can be formalised by means of the implication (23).

(23) $\forall x$ (eye-doctor(x) $\rightarrow$ oculist(x)).

In the valuations v_g there is no problem for this implication. Due to the coextensiveness of 'eye-doctor' and 'oculist' in v_g of (20) and (21), the implication between the antecedent from (20) and the consequent from (21) results in the value 'True'. The big difference is in the valuations (v_i) from (20) and (21). For (21), the differences between v_g and v_i result in the truth-value gaps in the supervaluation v_s. As rule (19) puts it, in the belief – context (11b) it is the supervaluated meanings of the antecedent and the consequent that are to taken into account. Moreover, supervaluations also induce altered truth-tables for the material implication. This truth-table is given by (24).

Now, if we take as the antecedent the v_s-value of 'eye-doctor(x) in (20) and as a consequent the v_s-value of 'oculist(x)' in (21), theses values are '1' and '–' respectively. From (24) it follows that the supervaluated implication between 'true' and 'truth-value gap'results in a truth-value gap. The value of (23) is thus a truth-value gap '–'.

(24) supervaluated truth-table for $(p \rightarrow q)$.

p \ q	T	F	–
T	T	F	–
F	T	T	T
–	T	–	T/–

(McCawley '81, 245).

If we would do the same for (22) the supervaluated meaning and its implication in (22) would result in 'true', even taking into account (24). The conclusion of all this must inevitably be that there is a difference in the supervaluated meanings of (10a) and (10b). That is of course the reason why substitution salva veritate of the one by the other is impossible in contexts of propositional attitude, though they are v_g-equivalent (i.e., in the general valuations these expressions are genuine equivalents). For the cases with rigid designators (12-13) one could construct an analogous argument.

If the supervaluated meaning for (12b) differs from the supervaluated meaning for (12a) because e.g. the one results in a truth-value gap while the other is true (in v_i), then, by rule (19), substitution in propositional attitude contexts is impossible. I think that this treatment may also well incorporate the idea that for Kripke (12b) expresses a necessary but a *posteriori truth*. It is a truth that is known a priori. The same holds for the two terms which in (10b) designate the same two 'non-natural' kinds' (as has been pointed at in 4, the discussion about the rigid designation of these terms is still going on, so we should consider this with a certain reserve). The interpretation of (12b) or (10b) results in a truth-value gap in the supervaluated meaning.

This does, however, not mean that there would be no truth-value whatsoever: "The principle upon which van Fraassen's semantics rests, so far from being specially appropriate to the accomodation of items which lack truth-value altogether, is plausible precisely on the assumption that the items in question have truth-values, though perhaps unknown truth-values" (Haack '74, 58). It is indeed so that the truth of (12b) does not depend on what goes on in the head of the speaker, but it was discovered at some time (though previously it was

unknown) and appears then necessary albeit a posteriori. Truth-value gaps in supervaluations indicate well this idea. A truth-value gap is not a third value, but may be seen as a postponement of truth value assignment. This postponement is, of course, in the case of (12b) only visible in the supervaluated meaning in propositional attitude contexts (this is also the case for (10b) because in v_g there certainly is a truth-value but that is the value 'as it should be', without taking v_i into account). Supervaluations, as used here, are also good logical counterparts for the parameters *i* and *m* from (14b). Not only can *i* be matched with v_i, but it does not seem too unreasonable to assume that *m* (the 'psychological performance factors') and the supervaluation are on a part. It is in v_s that the knowledge accessible to *an individual* in the interpretation is (hypothetically) fixed.

In this paper I have tried to show that the problems for a model-theoretic semantics are manifold. Especially in contexts of propositionel attitudes, this becomes all the more clear. Logical equivalence and rigid designation are cases in point. Solving them may mean the introduction of a new notion of *supervaluated meaning*. It should be noted that treating the semantic complements of propositional attitudes as supervaluated meanings not only can be brought in accordance with a model-theoretic semantics like Montague's, but allows us to predict the conditions under which substitution salva veritate of equivalents is (not) possible. If this is true it is perhaps not necessary to make an appeal to a vaguely defined propositional attitude psychology.

Research Assistant at the National Fund for scientific Research (Belgium)
Parkstraat 106, bus 11
3000 Leuven

Roger VERGAUWEN

REFERENCES

Bäuerle R., Egli U., Von Stechow A. (eds) 1979 *Semantics from different points of view* (Berlin/Heidelberg, N.Y.).

Chomsky N. 1955 *"Logical syntax and semantics. Their Linguistic relevance"*. Language 31, 36-45.

Chomsky N. 1975 *Aspects of the theory of syntax* (Cambridge-Massachussets).

Church A. 1964 *"The need for abstracts entities in semantic analysis"*. In: Fodor and Katz (eds) 437-445.

Cresswell M. J. 1982 *"The autonomy of semantics"*. In: Peters S. and Saarinem E. (eds), 69-86.

Dowty D., Wall R., Peters S. 1981 *Introduction to Montague Semantics* (Dordrecht).

Fodor J. and Katz J. (eds) 1964 *The structure of language: readings in the philosophy of language* (New Jersey).

Gochet P. 1980 *Outline of a nominalist theory of propositions: an essay in the theory of meaning and in the philosophy of language* (Dordrecht).

Grayling A. C. 1982 *An introduction to philosophical logic* (Sussex/New Jersey).

Haack S. 1974 *Deviant Logic* (London-New York).

Kanger S., and Ohman S. (eds) 1981 *Philosophy and Grammar* (Dordrecht).

Katz J. J. 1972 *Semantic Theory* (New York).

Kripke S. 1980 *Naming and Necessity* (London).

Lewis D. 1976 *"General semantics"*. In: Partee '76(ed), 1-50.

Loux M. (ed.) 1979 *The possible and the Actual: readings in the Metaphysics of Modality* (Cornell).

McCawley J. D. 1981 *Everything that linguists have always wanted to know about logic, but were ashamed to ask* (Chicago-Oxford).

Montague R. 1970a *"Universal Grammar"*. In: Thomason (ed), 222-246.

Montague R. 1970b *"English as a formal language"*. In Thomason (ed), 188-221.

Montague R. 1970c *"Pragmatics and intensional Logic"*. In: Thomason (ed), 119-147.

Montague R. 1973 *"The proper treatment of quantification in ordinary English"*. In: Thomason (ed) 247-270.

Partee B. (ed) 1976 *Montague-Grammar* (London-New York).

Partee B. 1979 *"Semantics-Mathematics or Psychology"*. In: Bäuerle R., Egli U., van Stechow A. (eds), 1-14.

Partee B. 1981 *"Montague Grammar, mental representations and reality"*. In: Kanger S. and Ohman S. (eds), 87-106.

Partee B. 1982 *"Belief-sentences and the limits of semantics"*. In: Peters S. and Saarinen E. (eds), 87-106.

Peters S. and Saarinen E. 1982 *Processes, Beliefs, and Questions* (Dordrecht).

Putnam H. 1975 *"The meaning of 'meaning' "*. In: Putnam '75, 215-271.

Putnam H. 1975 *Mind, language and reality. Philosophical Papers*, volume 2 (London-New-York).

Thomason R. H. (ed) 1974 *Formal Philosophy: selected papers of Richard Montague*. Edited and with an introduction by Richmond H. Thomason (New Haven-London).

Van Fraassen B. 1966 "Singular terms, truth-value gaps and free logic". *Journal of Philosophy 63*, 481-495.

Van Fraassen B. 1971 *Formal semantics and logic* (London).

Vermazen B. 1967 Review of Katz and Postal '64': an integrated theory of linguistic descriptions' and Katz 1966 'the philosphy of language'. *Synthese 17*, 350-365.