FORMAL SEMANTICS OF METAPHORICAL DISCOURSE

TEUN A. VAN DJUK

1. INTRODUCTION

1.1 The aim of this paper is a discussion of some of the problems in the formal, i.e., logical, semantics of metaphorical languages. A METAPHORICAL LANGUAGE is a language with metaphorical sentences or discourses like natural language. A formal semantics specifies the conditions under which such metaphorical sentences may be said to have a truth value, viz. it defines the notion of an interpretation for a language including metaphorical sentences. The idea is not, as usual, to assign the value 'false' or a neutral value e.g., 'non-sense', to all metaphorical sentences, but to reconstruct formally the idea that metaphorical sentences may be true in a given context.

1.2 It goes without saying that only a fragment of a serious theory of metaphor can be covered by the formal semantics approach. I therefore must presuppose well-known the current linguistic and psychological work on the structure and the functions of metaphorical discourse and will focus my attention upon the properly logical and philosophical problems involved. Within these domains, again, I shall neglect both syntactic and pragmatic aspects of the problem. The idea of a semantics for metaphorical sentences will not be based on an independent syntax, so that the interpretations will be quasi-formal. As far as the pragmatic aspects are concerned: although it will briefly be argued that metaphorical sentences can be interpreted only in context-determinate models, I shall have to leave out a detailed discussion of the pragmatic status of metaphorical aspects of conversation.

1.3 The semantics which will be introduced here must have a rather

* I am indebted to Dorothea Franck and Uwe Mönnich for very useful preliminary discussions about this paper. Suggestions by Hans Mooij and Francis Edelina could, regrettably, not be integrated before the paper went to press.
unusual feature in the theory of logic, viz. it must be sortal. That is, it only interprets those sentences which are sortally correct. In linguistic terms: it is a semantics accounting for selection restrictions. The problems involved here, both linguistic and logical, are considerable, and only part of them can be treated here: without many positive results, I am afraid.

The framework for the sortal semantics is basically the one constructed by Van Fraassen and elaborated by Thomason. The novelty of our approach is the assumption that under specific conditions sentences which, ‘superficially’, are sortally incorrect may be assigned an interpretation. This idea brings us close to recent work in modal (‘possible world’) semantics for conditionals and counterfactuals, e.g., by Thomason, Lewis, Åqvist and others, in which the concept of ‘similarity’, e.g., between worlds, plays such a central role. Although research on these and related topics is merely in its first stage — where, indeed, ‘theory’ and ‘metaphor’ are still close —, it may be hoped that such a formal, model-theoretic account provides a better basis for a theory of metaphor in natural language than the current linguistic attempts.

2. A REMARK ON THE PRAGMATICS OF METAPHOR

2.1 A sentence can be assigned a metaphoric interpretation only when a set of specific contextual conditions have been satisfied, especially with respect to the intentions of the speaker. As has been demonstrated by Dorothy Mack (Mack, 1975 [this issue]) the specific use of metaphors in conversation requires an examination of the speech act status of metaphorical utterances. Since the formal semantics will make abstraction from all relevant pragmatic features, making metaphorical utterances more or less ‘appropriate’, I make a remark on those pragmatic aspects before attacking the proper semantic problems (although it is difficult to fully separate semantics and pragmatics).

Within the general framework of the basic principles of rational conversational interaction as presented by Grice (1967), the utterance of a metaphorical sentence would, when taken ‘literally’, violate the basic cooperative principles of truth, relevance and manner. That is, a non-metaphorical interpretation would make the sentence false or nonsensical, irrelevant with respect to situation and discourse and — even when interpreted metaphorically — often a circumstantial ‘way’ of saying things or, more often, especially in literature, an obscure way of saying things.
However, as Grice has demonstrated, very often deviations from basic pragmatic principles have a reason, i.e., they remain within the scope of the Cooperation Principle for rational interaction. Metaphor is an example in case. The deviation, thus, is merely apparent, and the hearer knows that the speaker intends (to mean) something different than the 'literal' meaning of the sentence or at least implies additional meaning besides the literal meaning. The precise reconstruction of 'literal meaning' must await the following sections, and will here be taken in its usual sense, viz. as the 'normal', conventional meaning of an expression.

2.2 The appropriate use of metaphorical utterances in interaction is thus based on a general principle, e.g., something like:

(1) By uttering a sentence $S$ in a context $c$, such that for all possible literal interpretations of $S$ the utterance of $S$ violates the principles of Truth and Relevance, the speaker implicates (that the hearer knows that the speaker intends) a meaning of a sentence $S^*$, satisfying the mentioned principles and being semantically related with the meaning of $S$.

The pragmatic principle of metaphorical utterances might be formulated more precisely, but it will do for the moment, the more so while we cannot here go into the details of a pragmatic theory. Part of the principle, however, is semantic, and it must be made clear below under what conditions the literal meaning cannot 'receive' an interpretation. The purely pragmatic aspects involved are the following, among others:

(i) The speaker knows that the hearer knows that $S$ is false or nonsensical in $c$;
(ii) The speaker assumes that the hearer knows the semantic relation (operation) between $S$ and $S^*$ by virtue of his knowledge of the language and of a system of perception related to this language;
(iii) The speaker wants that the hearer interprets the uttered sentence $S$ as if it were the sentence $S^*$;
(iv) The speaker, by uttering $S$ and intending the meaning of $S^*$, wants the hearer to know that this indirect speech act has a specific reason (which may be non-conscious, like the wish to self-assessment by 'showing' linguistic virtuosity) — which in aesthetic contexts may be accomplishment of the specific speech act itself.
3. SENTENCE TYPES

3.1 Taking both formal and natural languages as (infinite) sets of sentences, it is theoretically useful to distinguish between different types of sentences. Thus, in classical formal languages a clear distinction is made between those syntactically well-formed formulas which can be assigned a truth value, and those wff's which, for some reason or other, are not assigned a truth value. In natural language sentencehood is determined by the grammar of that language, determining the possible phono-morphological, syntactic and semantic structures of each sentence belonging to the language. Sentencehood in natural language, however, is a rather fuzzy concept, and the grammar is merely an idealizing approximation to the notion 'sentence of a given language'. These methodological problems of linguistics will for the moment be put aside in favour of a more restricted analysis of some sentence types in formal languages.

3.2 The type of sentence which can be assigned a truth value is traditionally considered to be the declarative sentence in the present tense. Moreover, such classical valuations are bivalent, i.e., a sentence is either true or false, but not both.

It is well-known, however, that the set of meaningful or interpretable sentences of a language is much larger. First of all, tensed sentences are naturally true or false. Secondly, non-declaratives such as uttered in questions, commands, advices, requests, etc. may also receive some sort of interpretation: they may be said to be appropriate or non-appropriate in a given context, or — sticking to truth values — it may be argued that such sentences are made true or false in a given context. Finally, it has been attempted to introduce other, non-bivalent, truth values, e.g. 'indeterminate', 'zero', etc. for those cases where a (meaningful) sentence is neither true nor false.

One of the standard examples in the latter case is the non-satisfaction of the presuppositions of a sentence. Thus 'The a is F' is neither true nor false if a does not exist. Such and similar solutions are given in the valuation of some logical constants, especially material implication and negation.

From this brief discussion it appears that it is sensible to distinguish at least between the 'meaningfulness' of sentences on the one hand, and their 'valuation/interpretation' on the other hand. Probably the first property is a necessary condition for a sentence to have the second
property, i.e., before we may know whether a sentence is (made) true or false we must know what it means. In other words, the extension of a sentence depends on its intension, viz. on the proposition it expresses.

3.3 Without pressing too much the extremely controversial distinctions mentioned above, we may distinguish a set of sentences which are meaningless for some reason. Meaninglessness may be determined at several levels in natural language: it may result from inadequate performance of phonemes and morphemes, the syntactic structure of the sentence may be ill-formed and finally, at the level of semantic representations, predicates may be used which do not ‘apply’ to the referring phrases (or at the object level: properties are assigned to individuals which these individuals cannot possibly have). In this case meaninglessness results from the sortal incorrectness of the sentence, i.e., a predicate is applied to an object of the wrong type, category or sort. Since the terms ‘type’ and ‘category’ have numerous other meanings in logic and grammar the terms ‘sort’ and ‘sortally (in-) correct’ will be used here. Examples of sortally incorrect sentences are:

(2) The typewriter has a headache.
(3) The square root of Susy is happiness.
(4) May I please flow under your door?

These sentences are sortally incorrect because in most normal contexts typewriters are not the ‘sort’ or ‘kind’ of thing which may have a headache, Susy (when understood to refer to a girl) cannot have a square root, nor can square roots be identical with, or have as a value, something of the sort of happiness. Finally, I (i.e., the speaker, hence a human) cannot have the property ‘liquid’, and hence I am not able to flow. In usual linguistic terms, we say in these cases that selection restrictions are violated.

A characteristic property of sortally incorrect sentences is that their logical negations are also sortally incorrect. This fact prevents us from simply treating sortally incorrect sentences as false, although it might be argued that (external) logical negation should simply be read as ‘It is not the case that ...’, which makes some of the sentences meaningful, but only at the meta-level where the meanings of expressions can be discussed.

As we shall see below, however, meaningfulness is a relative and gradual concept. Sometimes, a sortally incorrect sentence expresses a proposition denoting an ‘impossible fact’, i.e., a state of affairs which
cannot even be conceived mentally. This is certainly not the case in sentence (4), because I may well imagine a world where speakers-humans are non-compact (e.g. in Asimov’s *The Gods Themselves*).

3.4 Finally, there is a type of sentence which is either plainly false or zero (e.g., for lack of presupposition satisfaction) or which is sortally incorrect, but which may nevertheless be assigned a normal (classical or non-classical) truth value, viz. truth in a given context. Those sentences are traditionally called metaphorical. Some examples are:

(5) *The car protested against such driving.*
(6) *We have been beaten up by the pigs.*
(7) *The sun laughed high in the sky.*
(8) *This mind made war* [cummings].

Since these sentences are, when taken literally, sortally incorrect, their truth value is at least ‘indeterminate’. However, we have the linguistic ability to interpret such sentences in a way such that they can be true or false. In the following sections it will be attempted to spell out the conditions under which sortally incorrect sentences may nevertheless be assigned a normal valuation. Notice, that a sentence need not be sortally incorrect in order to receive a metaphorical interpretation. Take:

(9) *The king is back in office.*

where *king* may be metaphorical for an authoritative boss. But at the same time a literal interpretation of the sentence might, in a specific context (where kings normally work in their office), yield normal truth values.

### 4. CONDITIONS ON A THEORY OF METAPHOR

4.1 A serious theory of metaphor must satisfy a number of methodological and empirical requirements. More specifically, a formal semantics for metaphorical sentences, part of such a more general theory, must account for a certain number of semantic facts. Some of these facts are the following:

(i) some sortally incorrect sentences can, given the appropriate context, be interpreted metaphorically:
(ii) not all sortally incorrect sentences can be interpreted in the sense of (i);
(iii) for each metaphorically interpreted sentence there is at least one reading, the so-called ‘literal’ reading, such that the sentence is sortally incorrect under this reading or false/indeterminate in each possible world compatible with the context of utterance;
(iv) there are sentences which in a context $c_1$ can be interpreted metaphorically, but not in a different context $c_2$;
(v) a sortally correct sentence which is false or zero may, in an appropriate context, be interpreted metaphorically.
(vi) some metaphorical sentences have non-metaphorical counterparts having the ‘same meaning’, i.e., which are semantically equivalent [the translation principle for a class of metaphors];
(vii) the metaphorical character of a sentence does not merely depend on its meaning or intension, but also on the extensions of its terms, i.e., it is determined by the (properties of the) individual referred to [the referentiality principle].

Other facts and principles the theory should describe and explain will be mentioned below.

4.2 At the same time the theory should evaluate the traditional treatments of metaphorical sentences. Some major ideas from this tradition, which will not further be analysed here, are the following:

(i) the substitution hypothesis: [the sun] laughs $\rightarrow$ [the sun] shines brightly;
(ii) the deletion hypothesis: the sun (shines so brightly as if it) is laughing;
(iii) the extension/reduction hypothesis: the predicate laughing does not only apply to humans but also to bright objects or to objects which make humans laugh.

Other versions and other proposals can be found in the literature. Our problem is: can such hypotheses be made explicit? In particular, under what conditions may predicates be substituted, deleted, extended, or reduced? Below, it will become clear that the mentioned hypotheses all focus on a specific aspect of the interpretation of metaphorical sentences.
5. SORTAL SEMANTICS

5.1 There are different ways to build up a semantics of natural language which is 'sortally sensitive'. I shall begin with a brief summary of the attempt by Thomason, and shall subsequently discuss some of the particular and general problems involved in this approach or in any attempt to provide a sortal semantics.

5.2 Thomason (1972), in his article "A Semantic Theory of Sortal Incorrectness", adopts several features from the semantic work of Van Fraassen (1967, 1969, 1971). Valuations are given with respect to a sortal specification. Only those sentences satisfying the sortal specification can receive a bivalent interpretation. A sortal specification, then, is a function which assigns to each predicate of the language a region of logical space. A logical space is conceived as a conceptual network by Thomason, and Van Fraassen speaks of a set of 'points', 'possible individuals', or 'possible objects'. Such a possible object, which in Thomason's terms may be viewed as an abstract 'role' in semantic space, is constructed out of abstract properties, such that predicates of the language can either apply or not apply to these possible objects. Thus, the application of the predicate 'large' requires the presence of a feature 'having dimensions', say. Individuals from a domain $D$ are said to be able to 'occupy' a point in logical space, i.e., they may become an instance of a possible object. The function determining this relation is the location function: \( \text{loc} \), whereas a reference function: \( \text{ref} \), selects the intended individual from the domain for a given term of a sentence. Predicates are assigned a subset \( I(P) \) of the appropriate sort \( E(P) \). A sentence is true if \( \text{ref}(a) \), say \( d_i \in D \), is located at a point \( \text{loc}(d_i) \) in the subregion \( I(P) \).

Let me give a concrete example to illustrate these ideas. Take the simple sentence

\[ (10) \text{The boy is ill.} \]

The valuation of this sentence has basically two phases. It is first determined whether it is sortally correct and then it is determined what (further) conditions make the sentence true or false. The sentence is sortally correct because the possible object \([\text{boy}]\) lies within the region \( E(\text{ill}) \), i.e., it is a member of the set of possible objects of which it may be true or false to be ill. In other words: the predicate \( \text{ill} \) applies to individuals to which also the predicate \( \text{boy} \) applies, probably because
of the fact that the possible object $\text{boy}$ has the property $\langle$organic$\rangle$ or $\langle$animate$\rangle$ required for an object to be sane or ill. The rest of the valuation of (10) is normal. That is, the predicate $\text{ill}$ or being $\text{ill}$ is assigned part of the region $E(i\text{ll})$. This part is the interpretation $I(\text{ill})$ of the predicate. Next, the definite referring phrase $\text{the boy}$ is assigned a particular individual, say $d_i$ — e.g. Peter —, element of the domain of discourse. This individual must be the only ‘instance’ of the possible object $\text{boy}$ i.e. $\text{loc}(d_i)$ must be an element of the $\text{boy}$-region of logical space. The sentence is true iff $\text{loc}(d_i) \in I(\text{ill})$.

5.3 Such a sortal semantics requires a brief discussion about its philosophical foundations. It is clear from the literature on ‘sorts’ or ‘types’ that our understanding of the relations between (in-) significance in language and the structure of logical space is still very fragmentary (see e.g., Sommers, 1963; Drange, 1966). Thomason’s notion of sort is close to the ideas put forward by Russell on the subject, where two objects are of the same type when the same (monadic) predicate is significantly (i.e., truly or falsely) predicable of both. Sommers himself proposes a more sophisticated approach combining ideas from Russell, Ryle, and Black. That is, he distinguishes between linguistically determined types and ontological types, in the following way:

- **A-type**: set of predicates of which every member spans every member of an $\alpha$-type (and no member a non-$\alpha$-type member);
- **$\alpha$-type**: set of objects spanned by a monadic predicate;
- **B-type**: set of predicates significantly predicable of a thing;
- **$\beta$-type**: set of things to which every member of a B-type applies.

In this terminology an A-type seems to correspond with what has been termed the ‘content’ of a term, or, in C. I. Lewis’ terms the ‘intension’ of a predicate, also known as its ‘connotation’. Hence it refers to those predicates which are included in a given predicate determining an $\alpha$-type. Thus, the predicates speak and think belong to the A-type corresponding with the $\alpha$-type spanned by the predicate human (see Van Fraassen, 1967).

The $\alpha$-type would in that case correspond with the ‘comprehension’ (Lewis) of a predicate (called ‘intension’ by Carnap), where the set of objects is conceived of as a set of possible objects, in the sense of Van Fraassen and Thomason. In case the set of objects is meant of which the predicate is actually true — at a given moment — we would have to speak of the ‘denotation’ or ‘extension’ of a term or predicate.
Notice that the $x$-type is not very homogeneous. According to this
definition all things which are 'thinkable' form a type, the things which
are 'yellow' (or rather: which may be yellow), etc. Hence an $x$-type is a
rather arbitrary finite subset of the power set of all possible objects.

A B-type is what we could also term a 'description' or 'characterization',
or perhaps rather a 'possible description' of an object. However, a
description is rather an (ordered) consistent subset of a B-type. At this
point a distinction between generic and non-generic predicates should
be made, and we enter the debate about essentialism: which properties
are essential and which are non-essential of a given thing? The B-type
could be split up accordingly. Thus, the predicate heavy or heavy should
somehow be distinguished from falling or being painted yellow.

The $\beta$-type, finally, is perhaps closest to our intuitive notion of a sort,
type or kind. In case we are able to make a serious distinction between
'necessary' (analytic) and 'possible' (synthetic, contingent) predicates,
the $\beta$-type would be specified in the same way, where the set of things
to which a set of analytic predicates applies would be a 'natural kind'
(for detailed discussion, see Kripke, 1972).

There are a large number of questions arising already from this
brief discussion on the different possibilities to introduce sorts in the
ontology for the semantics we need. What, for example, is the status of
higher predicates? What constraints are needed to allow a sentence like
The boy runs quickly and to mark off, somehow, a sentence like The boy
runs deeply? Taking adverbs as higher predicates would suggest that we
predicate something of a predicate, which would not very well fit into
any of the kinds discussed above. Hence the need to allow processes and
events to be serious individual objects, an assumption which would at
the same time account for the admissibility of a quick run (or the run
was quick) and the inadmissibility — in most contexts — of a deep run
(or the run was deep).

Another difficulty is the status of predicates of higher degrees. In
most of the discussion on sorts monadic predicates are used. Formally,
the kind for a binary predicate like love would be a set of ordered pairs,
viz. those of which the first member is human and the second any type
of object. Nevertheless, in practice we would rather use the complex
(abstracted) predicate ‘being able to love something/somebody’ as
being applicable to humans, or in general to those objects ‘having emo-
tions’. Similarly for converses like ‘visible’, for those objects on which
‘x is able to see [it]’ is applicable.
5.4 What, then, are the consequences for a formal semantics taking sorts into account? Some predicates, like happy and grunt have precisely natural kinds as their range, viz. humans and pigs, say. Other predicates, like big and yellow do not seem to lie within 'natural' kind boundaries. The applicability of predicates, as Thomason suggests, is thus determined by the conceptual structure of the possible objects. It may be the case that some form of B-type, in Sommers' sense, determines these possible objects, viz. as a set of those possible individuals in logical space satisfying an n-tuple of predicates. Thus, the applicability of big would require a conceptual structure with a feature 'having dimensions', and 'yellow' would require 'coloured' or 'visible' as features. These 'features' themselves are identified by predicates, which in turn require a conceptual structure of certain possible objects in order to be applicable, say 'concrete' for both 'visible' and 'having dimensions'. Apparently, the set of conceptual features defining a possible object is (hierarchically) ordered.

The applicability of, say, happy on a possible object [boy] seems thus being determined by the existence of a possible object [happy boy]. Such an object is possible if there is a possible world in which there is an individual instantiating the possible object or individual concept. Similarly, happy is not applicable to [table], because [happy table] is not a possible object, since a possible object with a feature [happy] must also have the higher ordered feature [having emotions], requiring in turn [animate], whereas [table] has at least a feature [non-animate]. The non-applicability of predicates, leading to the construction of impossible objects, is determined by a contradiction between two predicates for the basic conceptual features of an object.

Note that it is not easy to establish a clear distinction between linguistic (in-)significance and ontological possibility and absurdity (see Routley and Routley, 1969: 215ff.). The second case pertains to objects (things, including events, etc.) which are possible or impossible, i.e., which may have an instantiation in some or in no possible world. The first case can be decided only on the basis of the 'meaning' or 'content' of the predicates used to refer to an object or to a property or relation. Thus the sentence The yellow thing is square is linguistically significant, but in case the definite description the yellow thing refers to a circle, the sentence denotes an impossible fact. In other terms, a fact-concept like a yellow thing is square which is 'intensionally' possible, may have fact-individuals as values which are impossible, because the 'yellow'-concept may pick out things with properties which are inconsistent with those required for 'square' being applicable. In this example, thus, the
sentence is sortally incorrect under a given interpretation, which is determined by the context, viz. by the referents intended and their further properties. A sentence like *This circle is square*, however, is absurd under all interpretations, independently of context. Curiously, the last example, although expressing a plain contradiction, does not seem to be sortally incorrect, because the predicate *square* requires of a possible object to have (two) dimensions, which as such is not inconsistent with the predicate *circle*. Moreover, the sentence *A (the) circle is not square* is true, and would be indeterminate in case the non-negated version would be sortally incorrect.

5.5 It is clear that we cannot even attempt to solve here the numerous philosophical and logical problems involved. From the discussion above it seems that sortally incorrect sentences express a proposition which, under a given interpretation, 'denotes' a fact that is impossible in all physically possible worlds (situations, states of logical space) compatible with our actual world, i.e., those worlds having the same set of basic postulates (see Goble, 1973).

A serious insight into sortally incorrect sentences at the linguistic level and absurd or impossible facts at the ontological level requires previous understanding of the mechanisms of 'predication' and their corresponding operations in logical space. Should we maintain a fundamental distinction between 'substantive' concepts and 'attributive' concepts, the latter somehow presupposing the former? And does such a distinction correspond to a fundamental distinction between 'objects' and 'properties'? The objects of our world, values of (individual) concepts (see Scott, 1970), are topologically based selections of 'properties' of logical space. They are constant functions over changing time-space situations, continuous and structurally stable (for detail see Thom, 1972). The (non-analytic) assignment of properties, an operation expressed by predication on the linguistic level, produces non-stable objects like events/facts/states (Thom's catastrophes), e.g., from *boy* a *happy boy*, where the 'boy'-properties remain constant but the set of properties 'applied' under the property 'happy' may change in different possible situations. Thus, the impossible concept *happy table* does not have a value in any physically possible world, compatible with our own, because the discriminating or constituting feature 'inanimate' of the table-concept conflicts with the 'animate' feature of the happy-concept. Taking 'happy' as a function, we would say that it is only defined for a domain of higher animates or humans, whereas it cannot assign an
emotional state to tables since [table] itself has only values for a domain of inanmites.

5.6 Another feature of logical space to which more attention should be paid in order to be able to work out a complete theory of sortal semantics is the possibility of similarity and difference between its possible objects. Such an account is necessary for the description of degrees of sortal (in-)correctness.

Consider the following sentences:

(11) *The theory of relativity grunts.*
(12) *The horse grunts.*

Intuitively, we somehow find (11) more 'strange' than (12). That is, we have less difficulty imagining horses that grunt than theories that grunt. A horse could have all its characteristic properties, and therefore being still (recognized as) a horse, but one, perhaps rather marginal biological feature, viz. the production of a different sound, could be different. In that case practically nothing of the 'horse'-properties would need to be different in order for a horse to be able to grunt. In other terms: a logical space where [grunting horse] is a possible object is less remote than a logical space where [flying horse] would be a possible object, because such an object would probably require such a horse to have wings. *A fortiori* so, in sentence (11), for grunting theories. A theory would be required to be a concrete object, perhaps even be animate, and at least have a sound-producing mechanism. In fact, we would hardly call such an object a theory, because it gives a fundamentally different 'coupe' from logical space.

A provisional statement about degrees of sortal incorrectness would perhaps run as follows:

(13) A predicate $f$, applied to a term $a$ in a sentence $S_f$, is LESS APPLICABLE than a predicate $g$, applied to a the same term $a$ in a sentence $S_g$, iff the set of predicates entailed by $f$ which are also inapplicable to $a$ is larger than the set of inapplicable predicates entailed by $g$.

A sentence is sortally less correct is its predicates are less applicable.

Thus, in *The horse grunts* the predicate grunt does not entail a predicate which would be also inapplicable to horse, whereas in *The horse is arguing* the predicate argue would entail at least the predicates think, and speak, which are also inapplicable to horse. A semantic theory of similarity
relations in logical space would have to work out the topology of possible objects. Intuitively, it is obvious that \(|\text{horse}|\) has a smaller \text{distance} to the possible object \(|\text{cow}|\) than to \(|\text{theory}|\). \(|\text{Horse}|\) is connected with \(|\text{cow}|\) through all those properties designated by the predicates entailed by \text{mammal}, and with \(|\text{theory}|\) perhaps only through the most general feature \(|\text{object}|\). In other terms: the intersection of the conceptual structure of \(|\text{horse}|\) with that of \(|\text{cow}|\) is greater than that with \(|\text{theory}|\).

Similarly, we would have to account for the hierarchical structure of the possible objects in logical space. The possible object \(|\text{tree}|\) somehow ‘dominates’ the possible object \(|\text{oak}|\); the set of oaks is included in the set of trees, which seems to mean that \(|\text{tree}|\) is one of the features of \(|\text{oak}|\).

We will, however, leave these general discussions about the foundations of sortal semantics and now consider the specific problems in the interpretation of metaphorical sentences.

6. PROBLEMS IN THE SEMANTICS OF METAPHORICAL SENTENCES

6.1 The semantics of metaphorical sentences has a number of problems well-known in the semantics of modal operators and of intensional contexts in general: identity, substitution, opacity, etc.

First of all, Thomason’s \textsc{referentiality principle} in the valuation of sortally sensitive sentences is problematic for metaphorical sentences. That is, a sortally (in-)correct sentence must be so under equivalent substitution. Yet, sentences which may be perfectly correct under one description of an individual may be sortally incorrect under another description, as we saw in some examples above, so that sortal (in-)correctness must be determined at the object level in the interpretation. In metaphor, similarly, we start from a textually or contextually determined discourse referent, remaining identical under any description, including its metaphorical description. The familiar scheme of such sentences is then *\(Fa\), where \(a\) is a constant for the intended referent and *\(F\) is a \textsc{metaphorical predicate}, as in \textit{He is a bear} or \textit{She is an encyclopedia}, significantly assertable of humans in certain contexts.

However, there are also cases in which the metaphorical predicate appears in the descriptive referring phrase as in sentence (9) and in sentences like

(14) \textit{Peter preferred to pick one of the local flowers.}
where *flower* is intended to refer to a girl. Only under this referential condition sentence (14) is metaphorical. Under a literal interpretation it is not even sortally incorrect.

Such 'metaphorical referring phrases', however, must be treated as definite descriptions in general. As a general rule, definite noun phrases like *the Fx* may be derived only when the discourse referent has been introduced as the value of a term in an indefinite description in a previous sentence or in a pragmatically known proposition, including meaning postulates. Thus, the definite noun phrase *the local flowers* is 'grammatical' only when the interpretation specifies that the individuals referred to (viz. the local girls) have been mentioned before or are known to be intended by other contextual information. We will therefore assume that metaphorical referring phrases are derived from underlying structures in which the metaphorical term is introduced predicatively, e.g. as follows: 'the local girls are (like) flowers'. The same is true for such apparent indefinites like *Peter picked a local flower*, where the object noun phrase is equivalent with 'one of the local flowers'. Without previous identifying sentences a sentence like *Peter picked some flowers in the park* would, in a non-specific context, receive a normal, non-metaphorical interpretation.

Now, take the converse case, e.g. in the following sentence:

(15) The flowers in the park smiled at him.

Such a sentence may be metaphorically ambiguous in the following way: either *flowers* is non-metaphorically referring to real flowers and *smile* is a metaphorical predicate for e.g. 'blossom' (in a certain way), or *flowers* is a metaphorical referring phrase, denoting girls, and *smile* is a normal predicate for girls. From this example it is clear again that a sentence is metaphorical only with respect to the model structures in which the previous sentences or pragmatic, contextual knowledge is interpreted.

The same holds for sentences which are not metaphorically ambiguous but which, in isolation, would not have been interpreted metaphorically:

(16) The lion roared.

In this sentence *the lion* may be a metaphorical referring phrase to a strong man, say, where the predicate *roared* would, in order to 'stay in the image', be the adequate predicate for his cry. Such a sentence would be ambiguous in a narrative about a hunting party, but plainly metaphorical in a narrative about a battle in Flanders, say.
6.2 Another familiar problem arising in the interpretation of metaphor is whether identity is preserved under predicate substitution. In case of the local girls: as soon as it is implicitly or explicitly asserted that they 'are' flowers, we may freely use the referring phrase flowers. But the contextual identity girls = flowers will not hold in general, and the following sentences would be strange under a metaphorical interpretation:

(17) The government decided to give higher income to young flowers.
(18) The flowers are requested to dress after the boys.

The semantics of metaphor, thus, must guarantee that any form of identity or equivalence must be restricted to contexts with a specific structure.

In intensional contexts, e.g., in the scope of such verbs as know, believe, dream, pretend, etc., metaphors may be opaque:

(19) Peter thought (dreamt) that the flowers in the park smiled at him.

Several possibilities are open here: (i) presupposed is that girls = flowers, such that Peter's dream objects are girls, (ii) Peter dreams/thinks that girls = flowers, (iii) Peter dreams about real flowers, and it is presupposed that smile = blossom, (iv) Peter dreams/thinks that smile = blossom, (v) Peter dreams about real flowers which are 'really' smiling, which would entail that they somehow have a face and mouth.

The opacity of some texts can of course be resolved in cases where the metaphor is not likely to be under the scope of the intensional verb:

(20) Ford pretended that he had been bugged by those dirty FBI pigs.
(21) The young lady believes that her kid has been saved by the pigs.

The general rule holding for such cases is that the speaker knows, believes, or assumes that the metaphorical identity or equivalence holds. That is, metaphorical equivalence statements, normally implicit in the discourse, are under the scope of pragmatic operators (speaker believes ...). In other contexts it is obvious that the metaphor is under the scope of the intensional verb:

(22) The mayor, opening the new public park this morning, reminded the youth that trees and flowers are our friends.

6.3 There is another case in which the (co)text influences the possibility of metaphorical interpretation. Thus, sentence (15) may very well receive a literal interpretation in a fable or fairy tale. Here animals or
non-animate objects may systematically be assigned properties reserved to humans in our actual worlds. Hence, whether a sentence is metaphorical often depends on the sortal incorrectness of the sentence, but sortal incorrectness, as we saw, is determined with respect to interpretation in model structures with a set of possible worlds compatible with our own physical worlds. Thus, in her reply to Alice (Carroll, *Through the Looking Glass*, Chap. II), Tiger-lily may non-metaphorically utter:

"... it's enough to make one wither to hear the way they go on."

7. METAPHORICAL INTERPRETATION

7.1 With the problems mentioned above in mind, we now must specify the general conditions under which metaphorical sentences may receive an interpretation.

Recall that many metaphorical sentences are sortally incorrect whereas other metaphorical sentences would be contextually false under a literal interpretation. A sentence is sortally incorrect if the discourse referent occupying a given point (the possible object) in logical space is not a member of the interpretation domain of the sortal specification function for a given predicate. In other words: the individual would not satisfy the predicate under any interpretation. Example: *The theory grants.*

The traditional idea associated with metaphorical sentences is that instead of the sortally correct sentence *Ga*, denoting a given fact, a sentence *Fa* is used, such that under a normal interpretation *Fa* is false or sortally incorrect. Since it may be maintained that *Fa* expresses a different proposition than *Fa*, because *Fa* at least entails *Ga* and *Fa* does not, we shall distinguish syntactically between a given sentence *Fa* (receiving a literal interpretation) and its metaphorical counterpart *Fa*. Whether the '*' symbol has a specific operator value, will be discussed below.

Furthermore, it is traditionally supposed that a sentence *Fa* may be used in stead of *Ga* iff the predicates *F* and *G* have some sort of meaning relation.

Following a brief suggestion made by Thomason, this would mean that the conventionally determined logical region of the predicate *F* somehow changes in an expression like *Fa*. This change, in most cases would be an extension in case *F* does not apply to a region of which *a* is a member. In our examples given earlier, this would mean that the predicate flower is extended such that its region not only comprises flowers but also girls (in a given context).
7.2 How can these ideas be made more explicit in the framework of a sortal semantics?

First of all, it must be established that at least two principles must be respected in any account:

(i) the structure of logical space (for a given language) does not change by the use of metaphorical sentences (as it is the case in the fairy-tale context)

(ii) the conventional meaning structures of the predicates of the language do not change by the metaphorical interpretation of a predicate.

The conclusion is that the 'metaphorical predicate' is (ad hoc) used in a specific way, which indeed makes metaphorical interpretations pragmatically.

Nevertheless, although a given predicate may ad hoc be used in a specific way, there are general semantic principles determining under what conditions such ad hoc changes in the interpretation of predicates are possible.

Formally, the account would have roughly the following structure. Call the specific metaphorical meaning of a predicate $F$: $F'$. The interpretation $V$ under a sortal specification $E$ of the metaphorical predicate has a range $VE(F')$ large enough to have the referent $a$ localized in it.

Such a formal account, however, would be simplistic and explain nothing, if other features are not built in. Since the range of $F'$ is larger than that of $F$, the latter predicate must be less general. Hence some of the predicates (features) included in $F$ must somehow be dropped in metaphorical contexts. In other words: $F'$ must be constructed out of $F$ by a selection procedure. The intuitive idea of such a selection procedure is that those predicates entailed by $F$ which are applicable only to a more limited region of logical space are dropped. On the other hand the remaining content of the resulting predicate must be specific enough to convey non-trivial information.

The selection procedure may make use of a specific selection function $s$ with the set $P$ of predicates of a language $L$ as domain and co-domain. That is, $s$ operates an ordering in $P$ according to a similarity principle. In this way, each $f \in P$ is assigned a set $M$ of predicates 'environeing' $f$ such that each element of $M$ is more general than $f$. Hence it follows for any $g \in M$, and for any individual $u$ that $f(u) \vdash g(u)$.

The similarity function, however, must have a relative character. A predicate $g$ may be more general than a predicate $f$ with respect to a
certain 'tag'. Thus, with respect to the tag [sex] the predicate female is more general than girl, whereas youngster is more general than girl with respect to the tag [age]. The function will therefore be indexed with the feature(s) with respect to which it operates on the predicates: \( s_x(f) = g \). In generalizing we keep part of the included predicates constant, e.g. \( \langle \text{young} \rangle \) when going from girl to youngster, whereas others are dropped, e.g. \( \langle \text{young} \rangle \) when going from girl to female.

Since it is possible to drop several 'included predicates' we may introduce a linear set of subsets of \( M \), viz. \( M_1, M_2, \ldots, M_n \). In our example we may thus go from girl, via female to human, dropping \( \langle \text{young} \rangle \) and \( \langle \text{female} \rangle \) respectively.

Each predicate is thus surrounded by sets of concentric circles, depending on the feature 'on which' we generalize. This topological structure of our concept language requires further analysis which cannot be given here.

7.3 We have already noticed that the extension of predicates by dropping part of their content is not arbitrary. The intuitive criterion is that the features which must remain are somehow typical. It is not typical of a pig that it has four legs, nor of a girl that she has hair. Hence the predicates pig and girl may not be used in an extended meaning in which the features \( \langle \text{having four legs} \rangle \) and \( \langle \text{having hair} \rangle \), respectively, are constant.

The consequence of this criterion is that the property underlying the extension may not be too general. We do not use a metaphorical predicate table to identify or qualify a flower because both are concrete objects. A typical property of a possible object (or more in general, at the linguistic level, a typical predicate/feature included in a predicate) need not be ontologically so. The choice of typical criteria for the similarity function is pragmatically determined on the basis of cultural knowledge and beliefs. Thus, a pig need not be fat, physiologically speaking, but is typically fat with respect to other (semi-) domestic higher animals in the eyes of the members of a given culture. Thus we would need at least the following pragmatic postulate:

\[
(23) \ (\text{For most } x, \text{ and most } y) \ [(\text{member of cultural group}) (x) \ & \ (\text{pig}) (y) \ & \ B_x (\text{fat})(y)]
\]

where \( B_x \) is an indexed belief-operator.

So, the first condition is that it is generally believed that a given possible object has a property (or has this property in a remarkable
degree) which other objects belonging to the same, more general kind, do not have (or do not have in this remarkable degree/way).

The second condition is that the usual ‘name’ for the possible object or attribute in question may be used for an individual of which it is predicated that it has this property.

Assuming that the predicate flower has the features ⟨beautiful, animate, fresh ...⟩ typically, we may use this predicate in a context where this abstract feature structure is satisfied, e.g., where the possible object [girl] has an instance. Note that these specific features need not be defining the original concept, but may be contingent, contextually determined, features: the concept [girl] does not have [beautiful], say, as a ‘necessary’ property. In fact, this can be explained by the function of metaphors: we want to pick out and assign a specific contingent property of an individual, not the property the individual has anyway.

The fact that flowers grow is therefore not a sufficient condition to use the predicate flower predicatively for humans, which also grow. In that case only the manner, e.g., the speed, of the growing may be selected for metaphorical predication if this is specifically different from the average type of growing.

7.4 The selection function does not necessarily have predicates as values which have a lexical expression in the language. These values may consist of an ordered n-tuple of features. Again this is precisely often one of the reasons for the use of metaphors. Since we have no predicate ‘taking together’ (conceiving) ⟨fat⟩, ⟨dirty⟩, ⟨animate⟩, ⟨stupid⟩ exclusively, we use a predicate, e.g., pig, signifying a possible object which has these properties inclusively. In case only one predicate has to be asserted this substitution is of course easier: Peter is strong = Peter is a bear.

7.5 The general truth condition for metaphorical sentences like *Fa would thus become roughly the following:

\[(24) V(*Fa) = 1 \text{ iff} \]
\[
\begin{align*}
&\text{(i) } V(Fa) \neq 0 \\
&\text{(ii) there is an } F' \text{ such that } F' \in s_{a}(F) \\
&\text{(iii) } V(F'a) = 1 \\
&\quad \text{(iff: } a \in IE(F'))
\end{align*}
\]

In other terms the sentence is true just if the discourse referent, localized ‘at’ some possible object in logical space, is element of the region of the extended predicate \(F'\), viz. \(E(F')\), including the intended interpretation
of the predicate, viz. \(IE(F')\). Hence the sentence is true iff \(lo(a) \in M1E(F)\), where \(M1\) is the metaphorical interpretation of the predicate \(F\), which is identical with the 'normal' interpretation of the transformed predicate \(F'\), and where \(E\) is a sortal specification. The sentence \(*Fa\) is false iff condition (iii) does not hold, i.e., if \(a\) does not have the property \(F'\). In other cases, e.g., if the similarity function has no values, the sentence is metaphorically speaking neither true nor false but 'indeterminate'. Keeping the first two conditions constant we guarantee that our metaphorical interpretations are bivalent. These truth conditions should be embedded in a more general set of pragmatic appropriateness conditions. That is, although the conditions may hold, and thus a given predicate is metaphorically true of a given individual, it may be contextually inappropriate to use that metaphorical predicate, as in sentence (17) for example.

7.6 From this still very imperfect formal account it emerges that in asserting \(*Fa\) I basically assert \(F'a\), where \(F'a\) is true iff \(a\) has the properties selected from \(F\) by the similarity function. This interpretation is context sensitive. Since a policeman will not in general have the properties (fat, dirty/mean, stupid), \(F'a\) will not always be true and hence \(*Fa\) may be false. The equivalence \(policeman \rightarrow pig\) is therefore not analytic and substitution therefore not possible in all contexts. Nor may we use the predicate \(policeman\) to identify or characterize a pig, because a policeman does not have typical features which are contingently predicable of pigs. An unqualified substitution hypothesis is therefore inadequate.

8. METAPHORS, COMPARISONS, AND COUNTERFACTUALS

8.1 Selection functions have recently been used also in other formal semantic research. Especially in the semantics of certain connectives of the implicational type (material implication, entailment, counterfactual conditionals, causals, etc.) it has been fruitful to introduce selection functions operating on possible worlds. This work has been done by such scholars as Thomason, Stalnaker (see Stalnaker and Thomason, 1970), Åqvist (1973) and Lewis (1973).

In the interpretation of e.g. \(if \ldots then \) — statements, we might formulate the truth condition such that truth is obtained if the consequent is true in all the possible worlds 'selected' by a function based on the truth of
the antecedent. A proposition, which may be false in the actual world (of the speech context) may select those worlds in which it is true, all other things being equal, for which it is further asserted that the consequent holds. Instead of selection functions other formal devices may be used, e.g., indexed alternativity relations between possible worlds \((w_0, Rp_w1)\) (see Lewis, 1973).

8.2 There are a certain number of formal properties of metaphors which recall these aspects of the interpretation of conditionals of different strength. As a first try we might construe the metaphorical sentence as an implicit conditional of which the antecedent must be true and select the possible situations in which the consequent, viz. the apparent metaphorical part may be true. When we take again the trivial example Peter picked a local flower, we would then rephrase it something like:

(25) If a girl would be a flower, Peter picked a flower.

This is a bit a strange indicative conditional, which is true in case Peter picked a flower is true in a possible situation which would be compatible with the actual (speech) situation, with the only difference that girls have 'flower'-properties in that situation. If the antecedent is false, the metaphor is neither true nor false, it would be metaphorically indeterminate, indeed, no metaphor at all. Weakened versions of (25) would be:

(26) If a girl could be called a flower ...
(27) If a girl \{would be\} \{is\} like a flower ...

etc.

In all cases both antecedent and consequent are true/false in the actual world, which makes \(R\) at least reflexive. This is not strange, since the similarity relation is also reflexive: probably no world is as similar as \(w_1\) as \(w_1\) itself. Another reading of (25) would be:

(28) Provided girls are like flowers: Peter picked a flower.

One of the advantages of such an approach would be the possibility to make explicit the METAPHORICAL POSTULATE required to interpret Peter picked a local flower or The local flowers smiled at him, where the definites had been explained earlier by the presence of a previous (textual or contextual, i.e., pragmatical) proposition. We could build out this hypothesis a step further by assuming that the implicit antecedent in fact should be a COMPARATIVE SENTENCE, e.g. like in:

(29) If girls are as beautiful (as ..., as ....) as flowers, ...
where the \textit{tertium comparationis}, which corresponds to the constant feature determining the selection function, is made explicit. In that case the antecedent would select exactly those situations in which girls are like flowers but relative to their property of beauty, say.

This treatment further would specify the exact relation between a \textit{comparison} and a metaphor. A metaphor is not a shorter version of the comparison from which the 'like' and the \textit{tertium comparationis} are deleted (if this would yield a grammatical sentence at all), but a metaphor \textit{presupposes} a comparison. In that case, the comparison should be true for the metaphor to have a truth value, whereas the metaphor would be indeterminate if the comparison is false.

There are of course situations in which presuppositions of a given sentence are not satisfied by the context. The presupposed proposition in such cases may be said to be \textit{obliquely or indirectly asserted} together with the rest of the sentence. Similarly, in metaphorical sentences, the discourse referent may be introduced metaphorically such that it is characterized assertively by the specific predicates carrying the metaphorical transformation. The condition is, of course, that the discourse referent can be properly identified as the value of an individual concept by the following (con-)text.

8.3 In the account of the previous section the \textit{counterfactual} character of metaphors did not yet come fully into picture, and therefore it is not fully satisfactory.

Let us take another example to illustrate this point. The metaphorical sentence \textit{He is a bear} (as in the other examples it does not matter that the metaphor is rather trivial and even conventionalized) may be expanded as follows: \textit{He is as strong as a bear}. The latter sentence is close in meaning to a sentence like:

\begin{equation}
\text{(30)} \quad \text{He is as strong as if he were a bear.}
\end{equation}

Here the counterfactual is apparent in \textit{as if} and in \textit{were}. Now compare (30) with a normal counterfactual conditional like

\begin{equation}
\text{(31)} \quad \text{If he were a bear, he would be strong.}
\end{equation}

In (31) the proposition \textit{He is a bear} is false in the actual world, and no information is implied whether \textit{He is strong} is true in \(w_0\), although conversationally we normally infer that \textit{He is strong} is (also) false in \(w_0\). Sentence (31), when uttered, asserts however that \textit{He is strong} is true in all those worlds where \textit{He is a bear} is true. In fact his bear-hood entails
his strength in those worlds. The alternative worlds in which *He is strong* is true are at least as similar as our own world with respect to the validity of the sentence *A bear is strong* in both.

In (30) the situation is reversed, so to speak. It is asserted (to be true) that he is strong and implied that he is not a bear. We therefore let *He is a bear* again be true in an alternative world: in all those worlds accessible from that alternative world, with the aid of *He is a bear*, it is true that *He is strong*. Now, the problem is that *He is a bear* is not true in *w₀* so that *w₀* does not seem accessible from *w₁*. One solution for this dilemma is to assume that, seen from the point of view of *w₁*, *He is a bear* is true in *w₀*, e.g., because *bear* is a predicate with a larger region in *w₁*. This, however, would require the construction, not only of alternative worlds, but also of ALTERNATIVE LANGUAGES, a task which might theoretically be interesting but of which the difficulties cannot be estimated at the moment. Intuitively, of course, the metaphorical use of predicates seems indeed an ‘alternative’ use of predicates, a use ‘accessible’ through the selection function, with the condition that part of the predicate-structure is kept constant. In order for the (partial) similarity to remain recognizable, we would need a SIMILARITY THRESHOLD, for which cognitive psychological research may provide empirical data.

Still, the explanation given is either too speculative at the moment or simply incorrect. Perhaps we should modify the account of the semantics of *as if* as follows: *p as if q is true* (in *w₀*) iff (i) *p* is true (in *w₀*) (ii) *q* is false in *w₀* (iii) *q* is true in most worlds in which *p* is true (in most worlds: *(p & ~q)*). In fact, this is the semantic structure of an HYPOTHETICAL EXPLANATION. Take for example the following sentence:

(32) *The streets are* (so) *wet, as if it has rained.*

The *as if* clause provides a possible explanation for the antecedent, because in most possible (physically) worlds it is the case that the streets are wet when it is raining. The converse of course does not hold: there may be other causes for the street to be wet, but the fact that in most situations the wetness of streets is caused by rain, is a sufficient reason to use an *as if* clause. Sentence (32), indeed, could even be used in cases where, e.g., the firemen were the cause of the wet streets.

The very fact that other causes may be mentioned as explanation would be a satisfactory account of the possibility to give SEVERAL EQUIVALENT METAPHORS in order to assert the same fact. Indeed, in other cultures we might have *He is as strong as a horse/bison/gorilla* ... . If this approach is correct, a metaphorical sentence would be a shortened
as if clause, of which the entailed consequent is deleted if it is a typical consequence. We will not here explore the further axiomatic properties of the as if connective/operator, although at first glance it would yield a satisfactory explanation of the modal character of metaphorical sentences like *Fa.

University of Amsterdam, winter 1974

REFERENCES

Åqvist, Lennart

Drange, Theodore

Gobic, L. F.

Grice, H. P.
1967 Logic and Conversation”, unpublished lectures (Berkeley, mimeo).

Kripke, Saul A.

Lewis, David

Mack, Dorothy

Routley, R., and V. Routley

Scott, Dana

Sommers, Fred

Stalnaker, Robert C., and Richmond H. Thomason

Thom, René
1972 Stabilité structurelle et morphogénèse (Reading, Mass.: Benjamin).

Thomason, Richmond H.

Van Fraassen, Bas C.

Although much other work, especially linguistic and literary, has not been mentioned in this paper because it has been mentioned by other authors in this issue or because it is well-known to most readers of this journal, I would like to mention another omission: in the discussion on the conceptual structure of logical space I did not refer to the rich recent research in cognitive psychology and artificial intelligence (computer semantics), e.g. by Quillian, Schank, Anderson & Bower, Kintsch, and others, where for example notions like ‘semantic distance’ have been made explicit (see Abraham, in this issue).