



University of
Massachusetts
Amherst

Montague Grammar and Issues of Psychological Reality

Item Type	article;article
Authors	Partee, Barbara H.
Download date	2024-06-30 09:17:50
Link to Item	https://hdl.handle.net/20.500.14394/50167

MONTAGUE GRAMMAR AND ISSUES OF PSYCHOLOGICAL REALITY

Barbara H. Partee

Prepared for the Conference on Language and Psychotherapy
of the Institute for Philosophy of Science, Psychotherapy,
and Ethnics, April 1977

1. Introduction

Philosophers, linguists, and psychologists all share an interest in the semantics of ordinary language, but the same facts which account for the relevance of natural language semantics to such a wide range of fields also lead to frequent disagreement as to what the central questions of semantics are and as to the criteria by which semantic descriptions and semantic theories should be judged. Without trying to be prescriptive, I will begin by indicating three kinds of central problems which I believe are potentially of common concern in all of these approaches to semantics; then I will suggest some ways in which reasonable differences as to goals and criteria of adequacy may be reflected in different idealizations that are invoked by investigators in different fields.

The first problem can be put as follows: each natural language has infinitely many sentences, and native speakers can produce and understand indefinitely many sentences they have never heard or uttered before. The semantic side of this ability is the ability to associate meanings, whatever meanings may be, with each of the infinite set of sentences of the language. The task for a semantic theory then is to specify how this can be done in a finite way, presumably via a system involving a direct specification of the meanings of some finite set of primitive elements and a finite set of rules capable of associating meanings with all of the remaining, non-primitive expressions. The second and third questions relate to narrowing down somewhat the vague concept of meaning. The second is how to appropriately capture entailment relations between sentences within a semantic description;

this assumes that part of what is to be understood by the meaning of a sentence is that by virtue of which certain sentences can be said to follow from other sentences. The psychologist may be interested primarily in the inferences people actually draw, and the philosopher primarily in characterizing inferences which are logically valid, but I believe such differences can be expressed mainly as differences in criteria for the correctness of competing answers to the same question, namely how to characterize entailment relations between sentences. The third question, which is the least universally accepted of the three, is how to characterize the truth-conditions for sentences of a natural language. The importance of this question rests on the idea that a fundamental part of knowing the meaning of a sentence is knowing something like under what conditions it would be true, coupled with the observation that even the most complete specification of entailment relations among sentences will not tie the meanings of any of the sentences down to anything outside of language but only to other sentences of the same language. I will return below to some of the difficulties involved in taking the search for truth-conditions as one of the central problems of semantics, but at the outset I will simply accept it, since it has been taken as fundamental in the kind of semantic theory I want to discuss.

Now let us elaborate briefly on the suggestion above that the different interests in such questions from the perspectives of different fields may lead to different criteria of adequacy and in turn to different idealizations. Richard Montague, who should be regarded as relatively extreme in this respect even for a philosopher, was interested in a very general conception of possible languages, one at least broad enough to cover both natural languages and the artificial languages constructed by logicians; he had no interest in delimiting the class of possible human languages, which is the central concern of linguists. Hence where the linguist idealizes to an ideally competent speaker-hearer in a homogeneous speech community, and asks about the internalized systems of rules and representations that such a language user must have "in his head", so to speak, Montague idealized still farther, and did not talk of speaker-hearers at all. Much as a mathematician can study alternative systems of geometry without being concerned either about which geometry fits physical space or about how people form conceptions or intuitions about geometrical objects, Montague studied languages as formal objects, consisting of pairings between forms and meanings, where the meaning of a sentence is taken to be a structural specification of the conditions under which it would be true in any possible state of affairs, ~~or possible world, to use the favored expression.~~ A psychologist interested in general human cognition and language might accept the linguist's idealizations as a starting point, but a psychologist interested in individual differences among people might not want to accept the assumption of the homogeneous speech community, particularly if he wanted to explore the extent to which individual differences in knowledge, beliefs, and attitudes may be reflected in individual differences in language use within what may otherwise be reasonably regarded as a homogeneous speech community.

If we try to impose the talk of speaker-hearers on Montague, we have to assume a speaker who knows his language perfectly, is ideally rational, and is furthermore capable of conceptualizing a complete description of every possible world, e.g., God; the linguist is inclined to assume something like complete knowledge of the language and idealized if not

perfect rationality but not complete knowledge of what all the possible worlds are like (but rather some internalized model corresponding to conceptual possibility, or what Chomsky¹ has called 'common-sense understanding'); and the psychologist interested in individual differences probably wants to weaken the idealizations still further in various directions. The question of whether an approach to semantics developed with one set of idealizations can be relevant to an inquiry with a different set is a serious one, and of course one which cannot be answered a priori in a general way. My goal in this paper is to suggest that Montague's framework can be of relevance to linguistic and psychological concerns as well as philosophical ones, and that one can profitably attempt to develop linguistic and psychological theories of semantics by a combination of extensions and restrictions on various components of a Montague-like general theory.

In section 2 below I will describe Montague's framework briefly, and sketch some of the kinds of restrictions that a linguist might incorporate to try to characterize a theory of possible human languages as a subtheory of Montague's general theory. In section 3 I will suggest that the underlying models of Montague's model-theoretic semantics might be interpreted in various ways that could be relevant to a consideration of individual differences in beliefs and conceptions of reality, although I believe this will require extensions as well as restrictions on the class of models allowed by the theory. The purpose of the discussion is not to argue that linguists, philosophers, and psychologists should all tackle semantic issues by "doing" Montague grammar, but rather to suggest that if this way of looking at things is coherent and fruitful, it could help to facilitate discussion of theoretical issues involving quite different-looking theories.

2. Montague grammar and transformational modifications

2.1 Syntax and semantics

The version of Montague's framework that I will outline here, which comes from Montague (1973), gives a description of a natural language in three parts; a set of syntactic formation rules for the set of well-formed expressions of the language; a corresponding set of translation rules translating the set of natural language expressions into a language of intensional logic; and a model-theoretic semantics for the intensional logic, which gives for each sentence of the intensional logic a specification of the conditions under which it would be true with respect to any given possible world. Truth conditions for English sentences are obtained via the truth conditions for their translations, and entailment relations are definable in terms of the truth conditions. I will say more about the intensional models in section 2.3 below; here I will concentrate on the basic form of the syntactic rules and their relation to the translation rules.

The syntactic rules in a Montague grammar specify how to build up complex expressions of all of the categories of a given language, starting from a specification of the lexical items and their assignment to syntactic categories. Each rule specifies a way in which expressions of some given categories can be combined, and specifies the category

of the resulting combination. The form of a syntactic rule can be represented as in (1) below:²

- (1) If α is of category C_1 and β is of category C_2 , then γ is of category C_3 , where $\gamma = F_j(\alpha, \beta)$.

The syntactic details of how α and β are combined to give a new phrase γ are all contained in the function F_j ; the syntactic operations may be as simple as concatenation or as complex as a transformational operation. To a linguist familiar with transformational grammar, Montague's syntax gives the appearance of a bottom-up derivation, with phrase-structure-like rules and transformation-like rules freely interspersed in building up complex expressions from their constituents.

A main source of interest in Montague's system is the basic constraint he imposes on the relation between syntax and semantics. For each syntactic rule which combines expressions α and β to form a new expression γ , there must be a corresponding semantic rule³ which gives the interpretation of the complex expression γ as a function of the interpretations of the constituent expressions α and β . We can represent the form of a semantic rule as in (2) below:

- (2) If α translates into α' and β translates into β' , then $F_j(\alpha, \beta)$ translates into $G_k(\alpha', \beta')$.

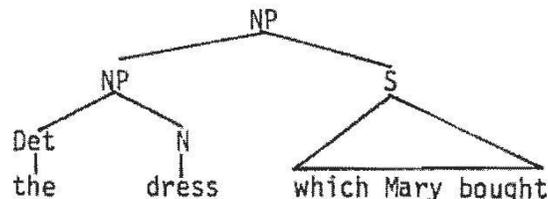
Here α' and β' are expressions in the intensional logic, and the function G_k will yield a new expression of the logic having α' and β' as subparts. One could think of the grammar then as a set of ordered pairs $\langle \text{syntactic rule}_i, \text{semantic rule}_i \rangle$; the syntactic rule will give the syntax-specific details of how the component phrases are to be combined, and the semantic rule will give the semantics-specific details of how the meaning of the whole is determined from the meanings of the parts.

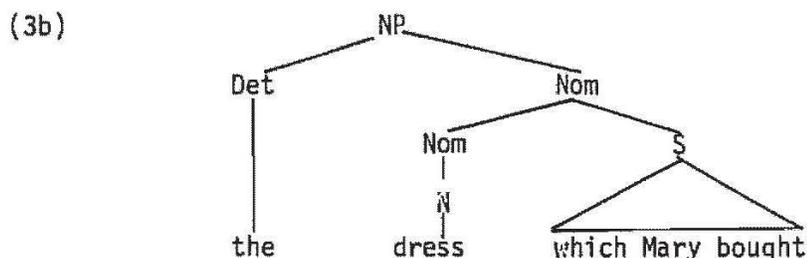
Let me illustrate both the rules and the constraint with a pair of examples. The first concerns noun phrases containing restrictive relative clauses, as in example (3).

- (3) The dress which Mary bought is blue.

Syntactically, linguists have debated about the following two possible structures for the noun phrase (ignoring for the moment the question of the underlying source of the relative pronoun which):

(3a)





I have argued elsewhere⁴ that Montague's constraint forces the choice of the second structure, with the relative clause modifying the common noun, if we wish to be able to give a uniform account of the semantics of the and other determiner and quantifier words. A first approximation to the semantics of the in singular noun phrases involves the condition that there is one and only one something-or-other in the given universe of discourse; (3) does not imply that there is only one dress, but that there is only one thing which both is a dress and is something Mary bought. Thus the relative clause should be part of the structure to which the determiner applies as it is in (3b) but is not in (3a). The structure diagrammed in (3b) can be represented in terms of Montague grammar by a pair of rules: one rule which combines a common noun phrase (NOM) with an open sentence to make a new common noun phrase, and another which combines a determiner with a common noun phrase (NOM) to make a term phrase (NP). Omitting details, the first of these can be written as follows:

- (4) If α is of category NOM and ϕ is of category S, then γ is of category NOM, where $\gamma = F_3(\alpha, \phi)$, and $F_3(\alpha, \gamma)$ is...

The specification of F_3 spells out the operations involved in relative clause formation, by which an input sentence ϕ of the form "Mary bought x_1 " is converted into the relative clause "which Mary bought", and the result is concatenated with the head NOM.

The corresponding semantic interpretation rule is basically just a conjunction of predicates, the two relevant predicates in this case amounting to (i) being a dress and (ii) being an x_1 such that Mary bought x_1 . The rule can be stated formally⁵ as follows:

- (5) If α translates into α' and β translates into β' , then $F_3(\alpha, \beta)$ translates into $\lambda x_1 (\alpha'(x_1) \& \beta')$

For this example, the interpretation of 'dress which Mary bought' would come out essentially as follows:

- (6) $\lambda x_1 (\text{dress}'(x_1) \& \text{Past}(\text{buy}'(\text{Mary}', x_1)))$

The lambda operator turns a sentence into a predicate; the resulting predicate in (6) can be thought of as expressing 'the property of being an x_1 such that x_1 is a dress and Mary bought x_1 '.

The details of the example are not of great interest *per se* to the issues at hand; I have included some of the details here in order to illustrate the force of the requirement that each syntactic formation rule be given a uniform semantic interpretation, and also to illustrate how having an explicit formalism for the semantic interpretation rules makes it quite explicit whether a proposed grammar meets that requirement. In my opinion, one of the big methodological advances made possible by Montague's work is to put semantics on as rigorous and explicit a footing as was done for syntax by Chomsky.

The second example concerns the distinction between predicative and attributive adjectives,⁶ which can be illustrated by the difference in validity of the inferences in (7) and (8).

- (7) (a) Susan is a red-haired doctor.
 (b) Susan is a violinist.
 (c) Therefore Susan is a red-haired violinist. [VALID]
- (8) (a) Susan is a skillful doctor.
 (b) Susan is a violinist.
 (c) Therefore Susan is a skillful violinist. [INVALID]

"Red-haired", a predicative adjective, simply picks out a set of individuals; the inference in (7) is valid, since (7a) involves simply the conjunction of the properties of being red-haired and being a doctor. "Skillful", an attributive adjective, does not simply express a property of individuals, however, as the invalidity of the inference in (8) shows. Parsons (1970) and others have proposed an elegant treatment of the semantics of attributive adjectives that represents them as functions which apply to the meanings or intensions of the common nouns with which they are in construction; informally, what it means to be a skillful doctor depends on what it means to be a doctor, and similarly for other nouns. (One corollary observation with potential practical utility for therapy is that any sentence which uses an attributive adjective alone in predicate position is semantically incomplete; a sentence like 'Susan is skillful' or 'I'm not successful' must be interpreted with respect to a further predicate to which 'skillful' or 'successful' can apply, and the choice of predicate may or may not be determinable from the context.)

This distinction between predicative and attributive adjectives can be captured in a Montague framework by providing two distinct syntactic derivations, each with its own semantic interpretation. Predicative adjectives start out in predicate positions, as in (9), and get into prenominal position by way of relative clauses, as in (10).

- (9) This block is cylindrical.
- (10) (a) Every block which is cylindrical is yellow.
 (b) Every cylindrical block is yellow.

Such a derivation corresponds to the standard transformational treatment of all adjectives. For attributive adjectives, however, their semantic

properties together with Montague's constraint rule out such a derivation. Those adjectives are best treated as combining directly with a common noun to form a new common noun phrase, semantically interpreted as a function applied to its argument. Many adjectives are ambiguous as between a predicative and attributive use, as is illustrated by sentence (11)⁷, for which both derivations are possible.

(11) Marya is a beautiful dancer.

Siegel (1976b) shows that there is a great deal of syntactic evidence in both Russian and English to support the double derivation of adjective-noun constructions and hence indirectly to support the requirement that each syntactic rule have a uniform semantic interpretation. It is this requirement, which is by no means uncontroversial, that accounts for a great deal of the current interest in Montague grammar; if it is borne out, it implies a strong connection between form and meaning in natural language; if it fails, it will be interesting to examine why and where syntax and semantics must diverge.

2.2 Adding further constraints for natural languages.

So far I have emphasized the constraint Montague's framework imposes on the relation between syntax and semantics. Now I must turn to some of the ways in which Montague's theory is much too unconstrained with respect to the interests of linguists. The syntactic and semantic rule forms given in (1) and (2) above involve syntactic operations F_i and semantic operations G_k , and Montague put essentially no constraints whatever on the form of these operations. If one is interested in characterizing the class of all and only possible human languages, it is imperative to look for constraints on the form of the operations allowed, and some of the recent linguistic work in the Montague framework is aimed at just that goal.⁸ I will not go into specifics here, but mention only that in this respect it seems reasonably natural to try to formulate a theory of possible grammars of natural languages as a restricted subtheory of Montague's more general theory, making use of the kinds of constraints on rules that have been formulated for transformational grammars.

A quite different sort of constraint which is to some extent implicit in Montague's own work is what I call the "well-formedness constraint."⁹ Since a Montague grammar works "bottom-up", constructing complex expressions from simpler ones, it is quite natural to impose the constraint that each expression built up in the course of a derivation be itself a well-formed expression of its respective syntactic category. This contrasts with the practice common in transformational grammar of positing underlying or intermediate stages in derivations which are not themselves well-formed expressions, but which are transformed into well-formed expressions by obligatory rules. To give just one illustration of the effect of this constraint, let us consider attributive adjectives again.

If all prenominal adjectives are derived from relative clauses, the postulated underlying sources are sometimes ungrammatical, as in

a transformation derivation of (12b) from (12a).

- (12) (a) *the owner ^{who}_{which} is rightful
 (b) the rightful owner

Since all of the cases in which the relative clause source is ungrammatical are also cases where the adjective is semantically attributive, the analysis which generates attributive adjectives directly in prenominal position avoids positing ill-formed underlying forms. Thus the well-formedness constraint in this case provides independent syntactic motivation for choosing the same analysis that is required by the constraint on the relation between syntactic and semantic rules.

The well-formedness constraint does not eliminate abstractness in syntax, but it essentially limits it to the operation of the rules which combine expressions to form new ones. By limiting the class of syntactic derivations, it limits the class of grammars compatible with the surface data of a given language, and thereby provides a potentially important restriction on the range of hypotheses a child might have to (unconsciously) consider in the process of mastering the adult language.

There are other kinds of constraints that could be sought within a Montague framework; for example, one might try to find correlations between the form of a given syntactic rule and the form of the corresponding semantic interpretation rule so as to strengthen still further the constraint on the relation between syntax and semantics. A great deal more descriptive analysis of particular constructions is required, however, since arguments about the correctness of various proposed constraints are invariably intertwined with arguments about the best analysis of particular constructions. But from the fragmentary evidence available so far, one can at least say that there appears to be no obstacle to pursuing the goal of characterizing the syntax and semantics of possible human language as a subtheory of Montague's theory by adding certain constraints motivated by empirical criteria of adequacy, criteria which may be of little interest to a philosopher or logician but of central concern to the linguist.

2.3 Model-theoretic semantics with possible worlds models

As I indicated earlier, semantic interpretation in the framework of Montague (1973) proceeds in two stages: the first stage, illustrated above, is a rule-by-rule translation of expressions of the natural language into expressions in a language of intensional logic; the second stage involves a model-theoretic interpretation of the intensional logic. It is at the second stage that the notion of possible worlds plays a critical role, and I want to try to sketch, non-technically, enough of what is involved in this second stage to raise certain issues about psychological reality.

The intensional logic has primitive constants of various types, such as names of individuals, one-place predicates, operators on predicates, etc., as well as variables of all these same types; the logic also contains various recursive means for forming complex expressions

of all the different types. An intensional model provides a domain of things outside of language for the expressions of the logic to refer to or be true of and a means of associating the expressions of the logic with particular entities or states-or-affairs in the model, so as to tie down the truth-conditions of the sentences of the logic and the references of the terms to appropriate non-linguistic anchors. The aspect of the models that I want to focus on is their inclusion of a set of non-actual possible worlds, treated as primitive (i.e., not defined or derived) within the theory. The motivation for the possible worlds part of possible worlds semantics has been its fruitfulness in providing an adequate account of the truth conditions in the actual world for sentences involving such "intensional" notions as logical possibility and necessity, contrary-to-fact conditionals, and psychological attitudes such as beliefs, hopes, and wishes, just to mention a small sample.

Let me illustrate the key notions of extension and intension and the role possible worlds play in their explication with some examples. Consider first the following inference:

- (13) (a) Alice wants to marry the richest man in town.
 (b) The richest man in town is the stingiest man in town.
 (c) Alice wants to marry the stingiest man in town.

The first and third sentences above are actually ambiguous; if we understand them as asserting of a particular individual that Alice wants to marry him, the inference is legitimate. But the reading I am interested in is one that could be brought out more sharply by appending "whoever he may be" to the final noun phrases of (13)(a) and (13)(c); on this reading, the inference is clearly invalid. The second premise asserts that the two descriptive noun phrases in fact pick out the same individual; this is what is meant by saying that the two expressions have the same extension in the actual world. The invalidity of the inference shows that the semantic contribution of the noun phrase to the truth-conditions of a sentence involving an intensional construction like "wants to" must involve more than just the extension of the noun phrase in the actual world. If we go further and consider how the same noun phrase, say, "the richest man in town," would pick out different individuals in different possible states of affairs, or possible worlds, we recognize that "the richest man in town" would not always pick out the same individual as "the stingiest man in town." The function which picks out the appropriate individual or extension in each possible world is in intension of the term; in an intensional context such as the "wants to" construction, it is the intension and not the extension of the terms that is crucial for the truth-conditions of the sentence.¹⁰ Substituting in (13)(a) another term with the same intension, e.g. "the man in town who has the most money" would preserve the truth-conditions of the sentence, but substituting a term with merely the same extension does not.

The attributive adjective plus common noun construction discussed above can also be seen to be intensional in many cases. First we must consider what extension and intension amount to for common nouns: the simplest treatment is to take the extension of a common noun like doctor

to be a set of individuals: the extension of doctor in the actual world is the set of individuals who are doctors in that world. The intension of doctor is the function which picks out the extension, i.e. the appropriate set of individuals, in each possible world. Now consider the attributive adjective construction good doctor. Let us suppose, just for the sake of examining our intuitions, that in fact all and only doctors were violinists. It would clearly not follow that the good doctors were the good violinists or conversely. Why? Because an attributive adjective like good operates on the intension, not the extension, of the common noun it is in construction with; its interpretation is a function which applies to a function to produce a new function.

Now perhaps we are in a position to say a little more formally what an intensional model looks like. In Montague's system, an intensional model is a quadruple $\mathcal{M} = \langle A, I, J, F \rangle$, where A is a set of individuals, I is a set of possible worlds, J is an ordered set of moments of time, and F is an interpretation function which assigns to each constant of the intensional logic an appropriate intension in the model. The extension of an individual constant in a given world will be some individual in the set A ; the intension will be a function from worlds to individuals. The extension of a predicate constant, like Man or Runs, in a given world, will be a set of individuals, and the intension will be a function from possible worlds to such sets of individuals. The function F which assigns an intension to the primitive constants of the logic thus determines for each constant its extension in each possible world. The rest of the model-theoretic semantics involves specifying the interpretation of the complex expressions of the logic; for example, one interpretation rule will say that a conjunction $[\phi \ \& \ \psi]$ is true in a given world if and only if each of the conjuncts is true in that world. The rule for sentences of the form $\Box \phi$ ("necessarily ϕ ") says that $\Box \phi$ is true in a given world w if and only if ϕ is true in every possible world.

It is impossible to convey an adequate picture of the workings of possible worlds semantics with such a brief and fragmentary sketch. One of its chief benefits has been in providing an account that is both rigorous and intuitive of semantic differences between sentences which superficially appear to be of the same grammatical and logical form. ~~As a final example before turning to psychological issues, consider~~ the following pair of inferences, one involving the intensional transitive verb look for and the other the extensional transitive verb talk to:

- (14) (a) Sam is looking for the president of the bank.
 (b) The president of the bank is the leader of the spy ring.
 (c) Sam is looking for the leader of the spy ring.
- (15) a. Sam is talking to the president of the bank.
 b. The president of the bank is the leader of the spy ring.
 c. Sam is talking to the leader of the spy ring.

The distinction between the extension and the intension of expressions such as "the president of the bank" or "the leader of the spy ring", which is formally captured by considering what such phrases would refer to in alternative possible worlds, makes it possible to explain neatly why

the inference in (14) is invalid while that in (15) is valid.

3. Abstract models and speakers' models.

3.1. Logically possible worlds

As I indicated in the last section, an intensional model is a quadruple $\mathcal{M} = \langle A, I, J, F \rangle$, with A a set of individuals, I a set of possible worlds, J a set of moments of time, and F an interpretation function. For the purposes of model theoretic semantics, which is centrally concerned with characterizing valid inferences, it is sufficient to consider such models abstractly, and one generally introduces a model by saying "Let A , I , and J be arbitrary non-empty sets, ...". Certain constraints on the possible models may be imposed by adding meaning postulates to the system; one can require, for example, that a proper name pick out the same individual in every possible world, or that the interpretations assigned by F to the predicates bachelor and male be such that the extension of bachelor is always a subset of the extensions of male. But the set of intensional models which count as potential interpretations of the intensional logic and hence indirectly of English is still very large, so there is a sense in which one has not by this means been given the interpretation of English. That does not matter for the task of characterizing valid inferences; a formula ϕ logically entails a formula ψ of the intensional logic of the conditional $[\phi \rightarrow \psi]$ is true in every intensional model consistent with the meaning postulates, with respect to all worlds and moments of time in that model.

But not only is it unnecessary and inappropriate to the logician's task to try to present particular intensional models in any concrete or explicit form, it is clearly impossible to do so with the resources of language. With the most conservative assumptions, the number of logically possible worlds in any realistic model must be at least non-denumerably infinite, which means that there is no way they could be listed or enumerated by a finite description or set of rules. By the same token, the interpretation function F must also be given abstractly one simply stipulates that to each constant of the logic there is assigned an intension of the appropriate type, i.e. a function from each possible world to an extension in that world. We can describe these functions by saying, e.g., that the intension assigned to the constant bachelor is the function which picks out in each world the set of individuals who are bachelors in that world. But we certainly cannot specify such a function by considering the possible worlds one by one and saying what the extension of bachelor is in each, since we cannot specify all the possible worlds individually. If we try to describe the intension assigned by F to some constant at all, we must do so by considering how the function applies to arbitrary possible worlds. To say that we must consider all possible worlds and cannot consider each possible world one by one is not as paradoxical as it may sound; we do the same when we consider functions defined on the real numbers, which are also a non-denumerably infinite set. But the nature of the psychological processes underlying such abstract reasoning ability is, I believe, a large and unexplored problem, and one to which I will return briefly below.

For the moment, however, I want to stay away from the psychological questions and consider further the logical models. The real work of the intensional model is performed not by the stipulation of the basic elements of the model, which is done in the abstract manner described above, but by the rules which define on the basis of the model how the interpretation of a complex expression is determined on the basis of the interpretation of its parts. In assessing the adequacy of a proposed semantics for English, say, we assume that we have a model in which the lexical elements have their intended interpretations, and we then judge the adequacy of the interpretations assigned to sentences by checking them against our intuitions about their truth conditions and entailment relations. It is easy to bring empirical evidence to bear on proposed semantic analyses at this level; the earlier-cited argument against the treatment of all adjectives as simple predicates is a case in point. The model may not specify anything about the difference between the intension of red-headed and that of blonde, nor about the difference between good and competent, but the difference in logical type between the two pairs has a great deal of explanatory force in accounting for the syntax and semantics of adjective constructions.

So let us assume we simply take as given a set of possible worlds. A sentence of the form ' $\diamond\phi$ ', where the diamond is the symbol for logical possibility, is interpreted as true in a given model if the sentence ϕ is true in the model at some possible world at some time. A model gives a 'correct' account of logical possibility if its possible worlds and the rest of the interpretation are such that each sentence that is in fact logically possible does come out true at some world and time in that model. But the attempt to determine which model or models are 'correct' in this sense is part of the enterprise of metaphysics, not of semantics. Similarly, we can characterize other notions of possibility, such as physical possibility, or possibility with respect to alternative future courses of events starting from the actual present. We can say, for instance, that 'Physically-possible ϕ ' is true in the actual world if ϕ is true in some logically possible world in which all the physical laws of the actual world hold; in this case, it is a matter for physics to give a characterization of the relevant class of possible worlds. Similarly, we can say that 'Future-possible ϕ ' is true at a given world and time if ϕ is true at some later time in some possible world which coincides with the given world up until the given time. There are many areas where the tools of possible worlds semantics can offer fruitful analysis of such modal constructions without giving more than an abstract specification of the relevant set of possible worlds.

3.4 Human conceptions of possibility

At this point, I want to turn away from the logician's conception of language as a formal system and back to the linguist's or psychologist's concern with the variously idealized competent speaker-hearer. Suppose we ask what it would take for a person to know a language as described in a Montague framework. Looking at the syntax, we find a finite number of primitive lexical items and a finite number of syntactic formation rules,¹² so there is no obstacle in principle to a speaker knowing that part. Similarly for the rules for translating the natural

language into the language of intensional logic.¹³ But when we get to the models described above, which form the basis of the interpretation of the intensional logic, we face some interesting problems. Let me focus on the possible worlds part of the model, since the sorts of problems I have in mind can all be illustrated there, although they arise for other parts of the model as well.

The first problem concerns finite representability. Suppose we accept the claim¹⁴ that any model adequate for getting the semantics of English right must have a non-denumerably infinite set of possible worlds. Then a human cannot have a finite representation of each of them any more than the logician can give a finite description of each of them. But it doesn't follow that we cannot have some internalized conception of 'all possible worlds', only that such a conception cannot consist of an itemized enumeration. It seems to me that we certainly do have conceptions of non-actual possible states of affairs, but not in the form of complete specifications of possible worlds. Among the clearest examples of conceptions of possibility are 'local' variations on actuality, as expressed in the sentences of (16), or the constructions of fantasy, which involve highly incomplete specifications of alternative possible worlds, as in (17).

- (16) a. Shirley could have been a linguist.
 b. If John had gotten home on time, ...
 c. Sam wishes that he had a son.
- (17) a. Suppose people walked on their heads and had eyes in their toes.
 b. Superman was born on the planet Krypton, where ...

We do not know all the facts about the actual world, but we presumably all share the notion that there is one, and from that basis we can generate partial specifications of many alternative possible worlds by the sorts of local variation illustrated in (16); we could think of these representations as generated by a formula like that of (18):

- (18) 'like the actual world except that ...'

Such specifications will be only partial, since a given change will usually entail that certain other things change as well, generally with considerable indeterminacy. (E.g. if we say "like the actual world except that Shirley becomes a linguist," we haven't specified whether Shirley goes to a different school, reads different books, meets different people, or who had the job Shirley now has, etc.)

The conceptual abilities involved in considering even these partial specifications of 'local' alternative worlds seem complex and powerful and worth a great deal of serious investigation; but especially interesting is the way in which we seem to be able to generate higher-order abstract possibilities. For instance, here is one simple way to arrive at a non-denumerably infinite subset of the set of possible worlds:

- (19) Consider some actual couple, say the Smiths, and consider some children they might have, and their children's possible children, and so on ad infinitum. For simplicity, fix the number of offspring of each couple and attach names to the individuals by generation and order of birth. This gives

an infinite domain A of possible individuals (incompletely specified). Now suppose we conceive that the property of having blue eyes may be distributed in any way among these individuals, so that for any subset of A, it could be the case that all and only the individuals in that subset have blue eyes.

The description in (19) is finite, but it is already enough to provide for the existence of a conception of non-denumerably many distinct possible worlds; all we need for that are the notion of an infinite set and the notion of subset.

So it seems to me that the problem of finite representability is potentially solvable; we can arrive by finite means at a conception of possible worlds such that there must be non-denumerably many of them, without anything like particular descriptions of each one, and in fact without having complete descriptions of any of them, and without having to have any explicit concept of non-denumerability.

But there are other interesting problems to consider as well. There are undoubtedly general limits imposed by our cognitive capacity on the kinds of alternative possible worlds we can consider. And since we are not able to process all the logical entailments of a given proposition, we are prone to fall into inconsistency both in our beliefs about the actual world and in our considerations of alternative possible worlds. Hence if we wanted to construct a theory of 'conceptually possible models,' we would probably need to allow for some sort of 'impossible worlds' as well as possible worlds, as Hintikka has urged for the semantics of belief-sentences.¹⁵

If we think about human conceptions of physical possibility or future possibility, we see that there are limitations imposed by our ignorance of facts as well as by our general cognitive apparatus. We may well have some notion of alternative possible worlds or states of affairs subject to the actual physical laws, but we clearly don't know enough facts to know what the physical laws are. And with respect to conceptions of alternative ways the future might turn out, we are limited not only in the foregoing ways but also by our ignorance of many present actual particular facts which could have a causal bearing on future events.

What I want to suggest is that the kind of model posited by possible worlds semantics may well have a kind of psychological reality, but where the logician simply stipulates the model in a highly abstract way, the psychologist and perhaps the linguist could profitably investigate the ways in which the individual's internalized model of both the actual world and of possible worlds is built up through the interaction of mind and reality. Different individuals' models undoubtedly share some properties and differ in others. What is shared presumably reflects properties of the human perceptual and cognitive apparatus plus the most universal features of human experience. Individual differences presumably result in part from differences in experience, which lead in many cases to differences in inductive generalizations and consequent differences in conceptions of what is possible, and probably in part from differential application of the cognitive apparatus (e.g., the creative imagination of great poets).

General cognitive psychologists, like most linguists, would presumably be more interested in the common part, raising questions of how our cognitive and perceptual apparatus operates on experience to generate conceptions of non-actual possibilities. The prelinguistic infant clearly has some notions of non-actual possibilities, as demonstrated by surprise reactions (presumably evinced by non-conformity of some actual state of affairs to some alternative expected state of affairs) and non-verbal expressions of wants (e.g. carrying a cup to an adult to ask for a drink). One might attempt a kind of generative model of conceptions of possibilities to try to show how conceptions of novel possibilities could be derived by basic cognitive operations applied to the objects of perception plus whatever innate non-perceptual ideas we may start out with. The nature of inductive generalization as a psychological process might be a central object of investigation. These suggestions are quite speculative and tentative, and I should apologize immediately for my lack of familiarity with most of the psychological literature; perhaps what I am suggesting is already underway.¹⁶

For the psychologist interested in individual differences, including those concerned with therapy, the relevant questions would presumably be how individual differences in internalized models arise, and how models can be changed. Perhaps a prior practical question to consider is how individual differences in models can be detected; here the notion of "language as a window on the mind" might take on a new importance. When two individuals have different models of what is possible--logically possible, physically possible, humanly possible, or whatever--they could utter the same sentence, and with the same meaning in one sense, and yet their conceptions of the truth conditions and entailments of that sentence could differ markedly. A typical case might be one in which a couple agree on the statement that she always gets angry at certain things he does; in his model of alternative possible worlds there are none in which he stops doing whatever it is (it's 'part of his nature') but there are worlds in which she doesn't get angry at it, while in her model of alternative possible worlds there are none in which that behavior of his fails to provoke anger (it's an 'involuntary reaction') but some in which he stops doing it. For some reason that therapists have probably figured out, the biggest holes in our conceptions of possibilities often seem to be in imagining ways we could ourselves change; we seem to have much less difficulty in imagining changes in others.

For the linguist or philosopher of language, a closely related issue is the question of how a public language is possible if all that speakers have to operate with are their internalized models.¹⁷ Perhaps the use of demonstratives is crucial here, as a link between language and reality that bypasses the need for descriptive accuracy.¹⁸ But that issue would take us too far afield here.

In conclusion, what I hope to have accomplished here is to describe enough of the framework of Montague grammar to show something of its suitability for describing the syntax and semantics of natural languages, and to show how issues that arise in considering the potential psychological reality of the possible worlds models that lie at its base may open up new areas for profitable joint exploration by philosophers, linguists, and psychologists.

Footnotes

I am grateful to Emmon Bach for much helpful discussion and support, to students and faculty at UC Berkeley and at Stanford University for opportunities to present and discuss early stages of my thoughts on this topic, particularly Julius Moravcsik, Patrick Suppes, Herbert Clark, Eve Clark, Jaakko Hintikka, and Alice ter Meulen.

I am also grateful to the Center for Advanced Study in the Behavioral Sciences for the fellowship which made the preparation of this paper possible; part of the funds for the Center fellowship came from the National Science Foundation and the Andrew W. Mellon Foundation. I also wish to express my gratitude to Raphael Stern for inviting me to participate in the conference for which the paper was prepared.

1. See Chomsky (1975), p. 4 and pp. 42-51.
2. For expository simplicity I am considering only binary constructions, but in fact the syntactic rules can combine any finite number of constituent expressions.
3. The rules I am here for simplicity calling 'semantic rules' are just the first stage of the two-stage semantics described above; their function is to provide the translation from English into the language of intensional logic.
4. Partee (1973); but see Bach and Cooper (forthcoming) for an argument that shows that a structure like (3a) is tenable if the is allowed to have more than one interpretation.
5. The formulations in (4) and (5) omit the important technical detail that both the syntactic and semantic rules in this case are actually schemata; F_3 is actually an infinite set of functions $F_{3,i}$, one for each choice of variable to relativize over syntactically, and perform lambda abstraction on semantically.
6. This discussion of adjectives draws heavily on the work of Terence Parsons (1970) and Muffy Siegel (1976a,b).
7. The example is from Siegel (1976b).
8. See Partee (forthcoming) for a discussion and some proposals.
9. See Partee (forthcoming).
10. Hintikka (1969) has interesting suggestions for not always considering the extension in all possible worlds, but rather in an example such as this one considering the extension in the set of all worlds compatible with Alice's wants. In this discussion I am ignoring the disputes and problems that still remain in the area of "propositional attitudes," and concentrating on the arguments for showing that something at least as powerful as the extension/intension distinction is needed.

11. More precisely, the predicate constants in the intensional logic into which the English words bachelor and male are translated.
12. In one sense, the schemata mentioned in footnote 5 result in there being an infinite number of syntactic formation rules, but since the schemata are finitely expressible, the distinction is not important here.
13. The same remarks apply here as in footnote 12.
14. See David Lewis (1973), p. 90, footnote. The chapter of which that is part is also an excellent statement of the "realist" position on possible worlds.
15. See Hintikka (1970a), (1970b), (1975).
16. Since the rest of this was written, I have become aware that there is indeed some apparent progress in this direction in cognitive psychology and artificial intelligence, including in particular a growing concern with the nature of mental representations and their acquisition. See, for example, Anderson and Bower (1973), Klahr and Wallace (1976), Miller and Johnson-Laird (1976), Kintsch (1974), Piaget (1970), Fodor (1975). I am grateful to Lauren Resnick for bringing several of these works to my attention.
17. In this connection, Hilary Putnam's ideas about the social division of linguistic labor and about the extent to which the nature of real-world paradigms determines the extension of our terms are of great potential interest. See Putnam (1975).
18. See Putnam (1975), Kripke (1972), Kaplan (1977).

Bibliography

- Anderson, John R. and Gordon Bower 1973. Human Associative Memory, V. H. Winston & Sons, Washington, D.C.
- Bach, Emmon and Robin Cooper forthcoming. "The NP-S analysis of relative clauses and compositional semantics," to appear in Linguistics and Philosophy.
- Bandler, R. and J. Grinder 1975. The Structure of Magic I. A Book About Language and Therapy, Science and Behavior Books, Palo Alto, Calif.
- Chomsky, Noam 1975. Reflections on Language, Pantheon Books, New York.
- Fodor, Jerry A. 1975. The Language of Thought, Thomas Y. Crowell Co., New York.
- Hintikka, Jaakko 1969. Models for Modalities: Selected Essays, D. Reidel, Dordrecht.
- _____ 1970a. "Knowledge, belief, and logical consequence," Ajatus, 32, 32-47.
- _____ 1970b. "Surface information and depth information," in J. Hintikka and P. Suppes, eds., Information and Inference, D. Reidel, Dordrecht.
- _____ 1975. "Impossible possible worlds vindicated," Journal of Philosophical Logic 4, 475-484.

- Kaplan, David. 1977. "Demonstratives: an essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals," read in part at the March 1977 meeting of the Pacific Division of the American Philosophical Association; "Draft #2", ms., UCLA.
- Kintsch, Walter. 1974. The Representation of Meaning in Memory, Lawrence Erlbaum Associates, Hillsdale, N.J.
- Klahr, David and J.G. Wallace. 1976. Cognitive Development: An Information-Processing View, Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Kripke, Saul. 1972. "Naming and necessity" in D. Davidson and G. Harman, eds., Semantics of Natural Language, D. Reidel, Dordrecht.
- Lewis, David. 1973. Counterfactuals, Harvard University Press, Cambridge, Mass.
- Miller, George A. and Philip N. Johnson-Laird. 1976. Language and Perception The Belknap Press of Harvard Univ. Press, Cambridge, Mass.
- Montague, Richard. 1973. "The proper treatment of quantification in ordinary English," reprinted in Montague (1974).
- _____. 1974. Formal Philosophy: Selected Papers of Richard Montague, edited and with an introduction by Richmond Thomason, Yale University Press, New Haven.
- Parsons, Terence. 1970. "Some problems concerning the logic of grammatical modifiers," Synthese 21, 320-334.
- Partee, Barbara. 1973. "Some transformational extensions of Montague grammar," Journal of Philosophical Logic 2, 509-534. Reprinted in Partee (1976).
- _____. 1975. "Montague grammar and transformational grammar," Linguistic Inquiry 6, 203-300.
- _____. (ed.). 1976. Montague Grammar, Academic Press, New York.
- _____. (forthcoming), "Montague grammar and the well-formedness constraint," to appear in a collection edited by F. Heny and H. Schnelle of papers from the Third Gronigen Round Table of 1976.
- Piaget, Jean. 1970. Genetic Epistemology, Columbia University Press; 1971 paper edition, W.W. Norton & Co., Inc., New York.
- Putnam, Hilary. 1975. "The meaning of 'meaning'," in K. Gunderson (ed.), Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science VII, University of Minnesota Press, Minneapolis; reprinted in Putnam (1975a), Mind, Language, and Reality: Philosophical Papers, Vol. 2, Cambridge University Press, Cambridge.
- Siegel, Muffy E. A. 1976a. "Capturing the Russian adjective," in B. Partee, ed. (1976), pp. 293-309.
- _____. 1976b. Capturing the Adjective, unpublished Ph.D. dissertation, University of Massachusetts, Amherst.