

Semantics - Mathematics or Psychology?

Barbara Hall Partee
University of Massachusetts, Amherst

University of Konstanz Colloquium "Semantics from Different Points
of View" September, 1978

1. Introduction. My goal in this paper is not to answer the question in the title, but to argue that it needs to be answered. I believe that we will not be able to get the semantics of propositional attitudes right until we face the question squarely and work toward a theory which reconciles the conflicting demands of the two kinds of views of what semantics is.

Where do these two views come from? The view that semantics is a branch of psychology is a part of the Chomskyan view that linguistics as a whole is a branch of psychology. This view, which is shared by many linguists, derives from taking the central goal of linguistic theory to be an account of what the native speaker of a language knows when he or she knows a language, and of how such knowledge is acquired. It is taken for granted by most linguists that people know their language; that in fact it is what is in people's heads that determines what their language is.

The contrasting view is ascribed to Montague (and endorsed) by Thomason in his introduction to Montague (1974): "Many linguists may not realize at first glance how fundamentally Montague's approach differs from current linguistic conceptions. ... According to Montague the syntax, semantics, and pragmatics of natural languages are branches of mathematics, not of psychology." (p. 2.) Thomason holds that lexicography is an empirical science that demands considerable knowledge of the world, but is

not part of semantics proper. I will return to that issue below.) The view that semantics is not psychology can also be reasonably ascribed to Frege (1918), and seems to be either implicit or explicit in the work of many logicians and philosophers.

It might seem from the recent growth of cooperation among linguists, philosophers, and logicians that the question really doesn't matter. And for many purposes I believe it doesn't. To put some perspective on what I take to be the real problem, let me list some respects in which the question is not a problem.

(1) There is no reason why a psychological theory can't be expressed in mathematical terms. In fact on a Chomskyan view it should be, since we are trying to discover something about the structure of a certain mental faculty, and mathematics is the best available tool for describing structure. So I am not suggesting that there is any incompatibility between mathematics in general and psychology in general.

(2) The fact that some logician is not interested in psychology does not preclude the possibility that he or she may develop a theory which can be taken as a serious candidate for a psychological theory. This was an argument I used in urging linguists to take Montague's theory seriously (Partee 1975); it was my hope that one might turn Montague's general theory of language (Montague 1970) into a theory of possible human languages simply by adding additional constraints, constraints designed to reflect human linguistic capacities. I still have hopes that such a program can be carried out for a theory bearing some resemblance to Montague's, since the progress that has been made in syntax and semantics working in constrained versions of that theory and similar ones seems too great to be an artifactual illusion. But I now believe that Montague's theory (and relevantly similar

ones) cannot be the basis of a linguistic theory without some radical revisions in the foundations of the semantics.

(3) As far as I can see, there are no problems with logicians' treatment of syntax analogous to the problems in the semantics. But I can't say any more about the difference until I have described the semantic problems. I will return to this later.

Having said what I think the problems are not, let me turn to the discussion of what they are.

2. Idealizations. Idealizations are of course indispensable for making headway in any science. Linguists are accustomed to making a distinction between competence and performance, and Cresswell (1978) shows that the same distinction can be used to argue for the reasonableness of truth-conditional semantics as an account of semantic competence. But the arguments of Putnam (1975b, 1975c), Chomsky (1975), and Linsky (1977) can be used to show that some of Montague's idealizations are incompatible with the view of semantics as psychology. In an earlier paper (Partee forthcoming), I discussed the Putnam arguments and concluded that speakers of a language do not in fact know their language; but Linsky's arguments and further reflection on the semantics of propositional attitudes have convinced me that it is Montague's idealizations that must be given up.

The crucial idealizations that Montague makes are the following:

- (1) The objects of propositional attitudes are propositions.
- (2) Propositions are functions from possible worlds to truth values, i.e. intensions of sentences.
- (3) Words have intensions, which are functions from possible worlds to extensions.

The crucial linguistic idealizations which I believe are incompatible

with those are the following:

- (4) People know their language.
- (5) The brain is finite.

A typical Putnam example is the intension of natural kind terms such as gold, water, or tiger. Putnam has shown that a speaker can, by all reasonable standards, be in command of a word like water without being able to command the intension that would represent the word in possible worlds semantics.

We might attribute this difficulty to the more general impossibility of fitting a Montague model into a speaker's head, since by any reasonable assumptions there must be non-denumerably many possible worlds, and hence the possible worlds could not all be represented distinctly within a finite brain. I don't think this is the source of the problem, however. For one thing, you don't need to represent all of the possible worlds distinctly in order to know a function which has them as domain. We know the function for adding arbitrary real numbers without being able to represent all the real numbers distinctly. For another thing, it makes sense independently to assume that our knowledge of meanings of lexical items should be represented by partial functions rather than total functions from possible worlds to extensions. The finite brain could be just a special sort of "performance" limitation, just as finite memory span is often assumed to be in syntax.

If we didn't worry about propositional attitudes, I think we might be able to achieve a reconciliation of psychological semantics with the possible worlds approach along the following lines. Take the Fregean or Montague semantics as representing a kind of super-competence: what we would be like if not limited by finite brains and finite experience (e.g. if we were God.) Finiteness restricts us to constructing partial models, and in place

of complete intensions of words we construct imperfect algorithms which yield partial functions on these partial models. Different individuals will have different partial models and different algorithms, since our brains and our real-world experience are not identical. Communication will be possible as long as there is sufficient similarity in our partial models and our imperfect semantics. Viewed in this way, there may be no problem in principle in regarding a theory like Montague's as a kind of competence model. It could be telling correctly how we can determine the meaning of a sentence from the meanings of its parts, with "performance" factors like finite brains explaining why we don't have complete meanings of lexical items to begin with and why we sometimes make logical mistakes along the way.

But such a story remains plausible only so long as we ignore the propositional attitudes.

3. Propositional attitudes. The trouble is that we know that we have these limitations, and this knowledge is reflected in propositional attitude sentences. The difficulty of formulating an appropriate semantics for belief-sentences and other sentences about propositional attitudes is well-known, and I would certainly not want to suggest abandoning any semantic theory out of hand simply because that theory did not so far seem to allow any adequate treatment of the propositional attitudes. But I believe that some of the idealizations that we take as fundamental in possible worlds semantics are the source of some of the problems we have in dealing with the propositional attitudes, and that an attempt to make our theories more psychologically realistic may be essential for solving those problems.

Among the problems associated with propositional attitude sentences, I believe that possible worlds semantics works relatively well for the problem of "quantifying in" and the problems associated with demonstratives and other

indexical words inside propositional attitude contexts. More "quotational" or "linguistic" approaches have serious difficulties with those problems. But there are two major problems that are not dealt with at all in Montague's sort of theory.

The first is the well-known non-substitutivity of logical equivalents in propositional attitudes. If P and Q are logically equivalent, we cannot validly make an inference from (1) to (2).

(1) Irena believes that P.

(2) Irena believes that Q.

This problem is widely admitted, but seldom confronted within Montague semantics, since it results directly from the assumption that propositions are the intensions of sentences and are functions from possible worlds to truth values. Montague's semantics requires that logical equivalents be intersubstitutable everywhere, and it will take a major modification to remove that requirement. To describe what I see as the source of the problem, let me refer informally to a language with no propositional attitude expressions as a level 0 language, one with a single layer of propositional attitudes as a level 1 language, and so on. (This is an informal borrowing from Russell's theory of ramified types.) And suppose we were to accept something like the view of Montague's theory as a theory of competence for speakers of a level 0 language, as I sketched in the previous section. Then we could express the argument of Linsky (1977) by saying that the performance limitations of speakers of a level 0 language must be acknowledged at the competence level of speakers of a level 1 language, and in general the level i performance limitations must be acknowledged at the level i+1 competence level. That is, as Linsky argues, even if an idealized speaker will always recognize the logical equivalence of P and Q, he should not make the inference from (1) to (2), since part of his competence would be the knowledge that holders of propositional attitudes can make

logical mistakes. I see this as one deep-rooted connection between the demands of a psychological theory of semantics and the demands of an adequate account of the semantics of propositional attitudes.

Another way of putting the point is this. Suppose we view Montague's semantics as a super-competence model: a semantics for English as spoken by God. Then the semantics works perfectly well for the level 0 parts of the language, but it still fails for the propositional attitudes, since God would not make the inference from (1) to (2).

Note that we cannot get around this problem by assigning the objects of propositional attitudes to some other semantic type within a Montague-like system that retains his intensional logic, since substitution of logical equivalents is valid in such a system for every semantic type. Nor can we solve it by allowing non-standard interpretations for the logical connectives if, as I believe, the typical case of the failure of the inference from (1) to (2) arises simply because holders of propositional attitudes don't always recognize logical equivalence, independently of what logic they are using.

The second major inadequacy of Montague's treatment of propositional attitudes has to do with the nature of lexical meaning. Let me add some background before stating it. Within the non-psychological semantic tradition, there are excellent arguments for why certain words such as proper names and perhaps natural kind terms should be viewed as rigid designators, that is, intensions which pick out the same extension in every possible world. As Putnam and others have persuasively argued, such intensions cannot be identified with psychological states narrowly defined. The mystery of how people can use such words at all can be solved by the causal chain story (see Kripke (1972), Donnellan (1974), Evans / (1973)). On such a story, a person who knows nothing about Frege can use the name Frege to refer to Frege, for instance in asking the

questions "Who is Frege?", simply by intending to use the name in the way that others before him have used it, so long as there is an appropriate causal chain tracing back to a situation that makes an appropriate connection to the individual Frege himself. Similar remarks apply to certain words that designate natural kind terms such as names of species, chemical elements, etc.

In these cases the gap between the "mathematical" view and the psychological view seems much wider, perhaps unbridgeable. What is in a speaker's head in association with a proper name bears almost no resemblance to the intension. The intension is a rigid designator, while the psychological representation is probably more like an incomplete and possibly incorrect definite description, or a partial algorithm for picking out the referent across times and worlds by qualitative characteristics. One complicating factor is that if the rigid designator theory is correct, then people in effect intend to use proper names as rigid designators, and therefore do not regard their associated descriptions and identifying procedures as constituting the meaning or intension of the name. Thus on a kind of metalevel (all this is unconscious, of course), our psychological states may be quite compatible with the rigid designator analysis, but on the ground level the individual speaker's psychological state will not in general determine a rigid designator. As Mondadori (1978) says, the intension in the modal semantics sense is a function from really possible worlds to extensions therein, whereas the kind of concept that can be "grasped" (by us) is a function from epistemically possible worlds to extensions therein, and these are likely to diverge in the case of proper names and other rigid designators. Mondadori argues that natural language semantics does not need the epistemic or psychological

notion, and that failure to keep such psychological notions out of modal semantics will only lead to confusion.

Now we can see the second problem and where it comes from. The problem is that rigid designators do not remain rigid designators in propositional attitude contexts. Sentence (3) does not entail sentence (4), as is well known, nor does (5) entail (6).

- (3) The ancients did not know that Hesperus was Phosphorus.
- (4) The ancients did not know that Hesperus was Hesperus.
- (5) John wonders whether woodchucks are groundhogs.
- (6) John wonders whether woodchucks are woodchucks.

The standard treatment in intensional logic works well enough where either or both of the terms are non-rigid descriptions, but it fails in cases like these where the two terms are rigid designators which rigidly designate the same entity or species. Such cases are sometimes dismissed as being peculiarly metalinguistic; one sometimes hears it said that (5) describes a case in which John doesn't know the language, with the implied suggestion that a semantics for the language need not try to treat directly sentences which are about people who don't know the language.

But in this sense no speaker of English knows English, since the rigid designators are in general not in our heads. Therefore I believe that even a super-competence model (the 'God's language' model) cannot ignore the psychological view of semantics. The ancients may have used 'Hesperus' and 'Phosphorus' as rigid designators of the same planet, but the two names were not psychologically equivalent for them, and sentence (3) is distinct from sentence (4) because it involves a psychological modality and not a metaphysical modality.

This problem extends to other lexical items as well, and not just rigid designators. The line between what is true by definition and what is contingent fact undoubtedly varies from speaker to speaker. Quine's complete rejection of the analytic/synthetic distinction may be too strong, but he is certainly correct in emphasizing that the meanings of particular words for a given speaker are determined in part by what sentences he or she holds to be true, and these will not be the same for everyone. When Piaget observes that many children believe that clouds are alive, we can be sure that the children have some belief that differs from some belief of ours, but we do not know without further investigation just what belief or beliefs it is that differ - a belief about what clouds are, or about what it is to be alive, or something else.

It won't do to say that the children don't know the language yet, since at this level no two of us speak the same language, and the difference between the children's language and ours is just somewhat more striking than usual.

In giving a semantics for a level 0 language, there may be no harm in idealizing away from these differences. But for propositional attitudes they cannot be ignored. Yet it is not at all clear what kind of a semantics could handle them. A quotational approach does not seem to do any better. If we try to represent (7) as something like (8):

(7) Mary believes that clouds are alive.

(8) Mary believes this: "Clouds are alive".

and use a Davidsonian interpretation of (8), whereby we state that Mary believes something that I the speaker would express with the sentence "Clouds are alive", we are still in trouble, since I don't have any idea what I would express with the sentence "Clouds are alive". (I suspect that

I would be expressing something necessarily false, but I don't thereby believe that Mary believes something necessarily false.) There may be some other kind of treatment of propositional attitudes that can solve this problem, but I don't know of any.

Let me summarize the two problems that I have been discussing. The first is the non-substitutability of logical equivalents in propositional attitude contexts. My claim is that this problem arises in formal semantics because of the idealization away from psychological limitations on our capacity to do logic; but propositional attitudes are psychological, and it is just these psychological limitations that make substitutions of logical equivalents fail in such contexts. The second problem is that rigid designators are not always rigid in propositional attitude context and more generally, that words do not appear always to have their usual intensions in propositional attitude contexts. I believe this problem to have a similar source: the psychological semantic representation of a word is often very different from its intension, and properties of the psychological representation are often the crucial factor in propositional attitude contexts.

4. Is semantics possible? My general theme so far has been that the view of natural language semantics as psychology is not just a reflection of a Chomskyan approach to what linguistics is about. We can start from the logician's goal of giving a correct account of the entailment relations among sentences of a natural language, even with the idealization to a super-ideal speaker who was omniscient about all possible worlds and who never made logical mistakes. Montague's semantics, and any other semantics that has a similar treatment of intensions, will make false predictions about

entailment relations among the propositional attitude sentences for such a speaker.

So I don't see how we can get a correct account of propositional attitudes without bringing psychology into the picture, but I also don't see how we can get along with it. The relevant psychological factors are ones which vary from speaker to speaker and moment to moment. No one can infallibly recognize logical equivalence, but there is no general way of determining who will recognize which equivalence when. The psychological correlates of word intensions are similarly variable across speakers and times. These were the very reasons why Frege suggested that if we want propositions to stand in a close relation both to language and to truth, we must not equate them with ideas.

At this point, I think it can be made clear why I think syntax does not face the same difficulty that semantics does. In syntax, we can safely use the notion of an idiolect of a single speaker, determined completely by what's in his head. At no point do the syntactic rules of one speaker have to be sensitive to the syntactic rules of any other speaker. There are no syntactic analogs to the propositional attitude problems. (It might be thought that direct quotation could present a similar problem, but I have argued in Partee (1973) that quotations should not be regarded as linguistic parts of the sentences that contain them.)

But for semantics, I do not see how even a single idiolect can be given a coherent semantics. I see three possible kinds of approach, of which two seem wrong and one doesn't seem much like semantics, and seems very much like giving up.

(i) The idealized approach exemplified by Montague. This approach seems wrong, for the reasons I have been discussing. If there is some other way to explain away what I see as clearly false predictions about entailment relations among propositional attitude sentences, then this approach could be maintained. But if, as it seems to me, the errors result from the basically non-psychologistic foundation of the theory, radical modification will be necessary.

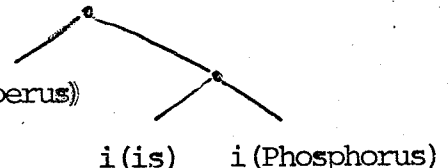
(ii) A realistic-egocentric approach. If I could figure out what my own psychological performance factors on making errors in logic were and what partial algorithms I associated with each lexical item, I could perhaps build them into a semantics for my own idiolect. But then in propositional attitude sentences I would be attributing those same properties to everyone else, and that would be just as wrong as attributing "super-competence" to everyone.

(iii) A realistic-nonegocentric approach. Add a variable m over psychological mechanisms incorporating performance factors; add a variable i over (psychological) interpretations of individual lexical items. Then try to devise a sensible way of representing a sentence like (9) as something like (10) or (11).

(9) Charles believes that Hesperus is Phosphorus.

(10) $(\exists i) (\exists m)$ Believe (Charles, m ($i(\text{Hesperus}) = i(\text{Phosphorus})$))

(11) $(\exists i) (\exists m)$ Believe (Charles, m ($i(\text{Hesperus})$)



This approach is attempting to incorporate Charles's own psychological performance factors and his idiosyncratic interpretations of lexical items

into the object of his propositional attitude; since we don't know what they are, we just put existential quantifiers for them. (This can be viewed as an extension of David Lewis's suggestion in Lewis(1970) that the objects of propositional attitudes might be meanings rather than intensions.) But not only do I not know how to make any concrete sense out of such an approach, I don't see how to begin to apply it to propositional attitude sentences with quantified subjects, like (12), or to propositional attitudes without subjects, like (13).

(12) No adult believes that clouds are alive.

(13) It is foolish to believe that clouds are alive.

Furthermore, such an approach seems not to allow any inferences at all from propositional attitude sentences. Perhaps there aren't any that follow without additional premises, but that seems too much like giving up. It seems to me that we do want to count inferences like (14) as valid, and I don't think they would be in a system of this sort.

(14) Mary believes that clouds are alive.

Joe believes everything that Mary believes,

therefore Joe believes that clouds are alive.

I should mention here that Hintikka and Saarinen's "urn model" is a real attempt to construct a kind of model-theoretic semantics that builds in certain performance constraints. I don't understand it well enough to discuss it, however.

5. Conclusion. I certainly have not answered the question whether semantics is mathematics or psychology. There are undoubtedly too many different kinds of mathematics and psychology for a general answer to be possible. What I have tried to suggest is that the linguist's concern for psychological representation may be relevant to every semanticist's concern for an account of the semantics of propositional attitudes. So far I don't see how to achieve either goal; my only positive suggestion is that a good theory might be expected to achieve both at once.

REFERENCES

- Chomsky, Noam (1975), Reflections on Language, Pantheon Books, New York.
- Cresswell, M.J. (1978), "Semantic competence", in M. Guentner-Reuter and F. Guentner, eds., Meaning and Translation: Philosophical and Linguistic Approaches, Duckworth, London.
- Donnellan, Keith (1974), "Speaking of nothing", Philosophical Review 83, 3-32; reprinted in Schwartz (1977).
- Evans, Gareth (1973), "The causal theory of names", Aristotelian Society Supplementary Volume 47, 187-208; reprinted in Schwartz (1977).
- Frege, G. (1918), "The thought", in G. Iseminger, Logic and Philosophy, Appleton-Century-Crofts, New York, 1968.
- Hintikka, Jaakko (1975), "Impossible possible worlds vindicated," Journal of Philosophical Logic 4, 475-484.
- Kripke, Saul (1972), "Naming and necessity", in D. Davidson and G. Harman, eds., Semantics of Natural Language, D. Reidel, Dordrecht.
- Lewis, David (1970), "General Semantics", Synthese, v. 21 (1970).
- Linsky, L. (1977), "Believing and necessity", Proceedings and Addresses of the American Philosophical Association 50, 526-530.
- Mondadori, F. (1978), "Interpreting modal semantics" in F. Guentner and C. Rohrer, eds., Studies in Formal Semantics, North-Holland, Amsterdam.
- Montague, Richard (1970), "Universal grammar", Theoria 36, 373-98. Reprinted in Montague (1974).
- _____ (1974), Formal Philosophy: Selected Papers of Richard Montague, edited and with an introduction by Richmond Thomason, Yale Univ. Press, New Haven.
- Partee, B.H. (1973), "The syntax and semantics of quotation", in S.R. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle, Holt, Rinehart, and Winston, New York.
- _____ (1975), "Montague grammar and transformational grammar," Linguistic Inquiry 6, 203-300.
- _____ (forthcoming), "Montague grammar, mental representations, and reality", to be published in proceedings of the symposium "Philosophy and Grammar", Uppsala University, June 1977, Stig Kanger and Sven Ohman, eds.

Piaget, Jean (1929), The Child's Conception of the World, trans. J. and A. Tomlinson; 1972 edition, Littlefield, Adams, & Co., Totowa, N.J.

Putnam, Hilary (1975a), Mind, Language and Reality: Philosophical Papers, Vol. 2, Cambridge University Press, Cambridge.

_____ (1975b), "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).

_____ (1975c), "Language and reality", in Putnam (1975a).

Schwartz, Stephen R. (ed.) (1977), Naming, Necessity, and Natural Kinds, Cornell University Press, Ithaca.