Could Semantics Be Something Else? Philosophical Challenges for Formal Semantics

Martin Stokhof

April 1999

Abstract

Starting from an old, yet still relevant paper of Johan van Benthem, this paper explores the possible consequences of certain philosophical considerations for the concept of a formal semantics of natural language, and argues that this issue is important from the perspective of an overall theory of human information processing.

Contents

1 Introduction

When in 1980, on the Third Amsterdam Colloquium, Johan van Benthem read a paper with the title 'Why is Semantics What?' (cf. [1]), I was puzzled: Wasn't it obvious what semantics is? Why did our concept of it stand in need of justification? Later, much later, I came to appreciate what Van Benthem was doing in this paper (and in some others). Questioning the 'standard model', the assumptions on which the working semanticists silently agree, Van Benthem opened up a space of issues to be discussed, questions to be asked, routes to be explored, that had been hidden from view by the unreflective endorsement of just one possible, albeit fruitful way of doing semantics. History, by the way, has proven him right on many points: the monolithic approach that dominated formal semantics of natural language in the seventies, and which relied heavily on Montague's seminal papers, has given way to a multitude of different ways of tackling semantic issues, using different formal techniques. Some limitations, in particular the almost exclusive focus on sentences as the primary units of analysis, have been overcome. In another respect, however, I feel that the message of Van Benthem's paper has not caught on sufficiently. He urges semanticists to take more interest in the properties of their tools, arguing that such questions are important if we are to come to a real, deep understanding of what semantics is. Such 'meta-level' considerations, although certainly less scarce than they used to be, are still not an everyday concern of the working semanticist.

The present paper present a 'counterpoint' to Van Benthems considerations. His suggestions for putting semantics on a proper footing are primarily concerned with the logical and mathematical aspects of the trade. But there is also another way of viewing the question after the status of formal semantics: the philosophical one. Although Van Benthem expresses clear misgivings about the relevance of such philosophical considerations, and insists that 'if only philosophical background attitudes [are] at stake, then one need not bother' ([1, p. 30]), since such philosophical qualms about semantics are 'noble — and, therefore rather unexciting' ([1, p. 31]), I will in the remainder of this paper take up the challenge and try to argue that certain philosophical considerations might, in the end, lead to yet another perspective on what semantics is, or might become, which is both interesting and fruitful.

Modern philosophy of language has produced several conceptual challenges that, I think, formal semantics should meet if it is to be a viable theory of natural language meaning, one that will fit into an overall theory of (human) information processing. Examples are the 'thought experiment' of radical interpretation; questions concerning rule following and normativity; the phenomenon of direct reference and causal theories of reference. In what follows, I will sketch the idea of radical interpretation very briefly, indicate the nature of the challenge its presents, and argue that a serious attempt to meet that challenge will lead to different perspective on what semantics is about.

Before entering into that argument, however, I want to quickly survey and discuss some, I presume standard, reactions to the very question what formal semantics actually is.

2 What semantics is not

The most common initial reaction is that semantics is an *empirical science:* natural language and its meaning are an empirical phenomenon, and constitute its object of inquiry. However, this is a position that is more easily taken than defended. What exactly are the data, and how hard and independent are they? How do we gather them? And to what extent do they provide us with an independent testbed for the descriptions and theories we base on them? Here it is important to note that traditionally formal semantics uses a particular methodology to approach its subject. Its empirical domain is said to consist of intuitions, of the individual speaker of the language, about meanings of (classes of) expressions and about meaning relations (such as entailment, synonymy, etc.) between them. At the background of this view stands, I think, the Chomskyean revolution[1] in linguistics. Two main characteristics of this revolution can be found in formal semantics as well. First of all, there is the assumption that language exists as an object, i.e., as an entity in its own right, with its own structural properties and content, which can be studied independently of its actual use and its historical development. Secondly, the idea of an individual having complete 'knowledge of language' is concomitant to the idea of intuitions as the primary data of the linguist. These two features of the Chomskyean approach are then wedded to the descriptive kind of semantics that dates back to (at least) Fregean days. The result is an account of meaning that is individualistic, platonistic, and ahistorical. Usually this is justified by an appeal to the kind of abstraction that is supposed to be characteristic for every empirical science. Physics studies movement on a frictionless plane, or the properties of a complete vacuum; analogously, semantics deals with meaning and its properties on an equally abstract, yet empirical level. Or so it is assumed. But whereas physics, abstraction not withstanding, produces empirically testable predictions, it is far from clear that the same holds for semantics. Do we predict intuitions? What would that mean? Or are our formal theories intended as descriptions of the actual processes that occur in natural language users? If so, how are we to test them? And given the fact that these theories deal mainly with structural aspects, and not with lexical content, are they not hopelessly incomplete?

A less naive reaction we might dub the *engineering view:* here we take semantics to be a much more modest enterprise. Instead of viewing it as a theory (in a strict sense) about what meaning is and how language users operate with it, we refrain from such far-reaching claims, and merely require that semantics provide a formally and empirically adequate specification of the 'input – output' conditions. Whatever the actual underlying processes turn out to be, an account of them is to be in accordance with the results of a formal semantic theory, but the latter is supposed to be neutral with respect to the former. This is not unlike one of the ways in which semantics plays a role in practical applications, such as in natural language query systems, translation systems, and the like,

[1]Though not the Chomskyean *paradigm:* for that is quite at odds with a formal, modeltheoretic account of natural language meaning.

where semantics itself is not at stake in the actual systems, but functions on the meta-level as a kind of correctness criterion. Obviously, this is a more modest stance, and it saves semantics from some of the more nagging questions. But it does so at a cost, since this point of view seems to rob semantics of quite a lot of its empirical content. If natural language meaning is in some sense of the word a systematic phenomenon, it can most probably be captured by formal (logical, mathematical) means: that is not a very exciting claim.[2] Moreover, this line of defense betrays the actual practice of semanticists. In reality there is a lot of debate about the empirical aspects of formal semantics,[3] so presumably the goals of the people debating these issues are less modest than this view of its status can account for.

The engineering view leads quite naturally to an even more radical position, which is that semantics is a *deductive science* . Its apparent empirical contents are more like the intuitions (about space, for example, or numbers) that put some constraints on a mathematical theory. This way of conceiving semantics has a certain attractiveness. (And I must confess that for quite a long time this was my favorite answer to the question what semantics is.) It lends it a kind of elegance and puts it aloof of some of the difficult questions raised by the empirical point of view. However, I feel that in the end it will not do. The question of justification can be avoided for some time, and it may even be quite reasonable and justifiable to ignore it in a certain stage of the development of a discipline, but in the end it will simply not go away. Even mathematics struggles with this problem, albeit not in the same way and to the same extent. On this view semantics is concerned with the discovery and study of abstract structures underlying an empirical domain. However, there is structure everywhere, and not all of it is equally significant. So even on this detached view there remains a question to be answered, or, equivalently, a demand to be met: How do we decide which structures are important, significant, or, as we however reluctantly are inclined to say, real? 'Reality bites': significance as a criterion of adequacy is necessarily transcendent to a theory, so the empirical question, it seems, can be forestalled, but not avoided.

But if that is true, as I believe it is, neither can we forego the task of examining the philosophical assumptions on which formal semantics rests. This involves not just identifying such assumptions, but also asking the question of their justification. Above, I indicated two such assumptions, viz., that language is an entity and that linguistic (semantic) competence is individual. These assumptions are under heavy attack, and this means that we also need to investigate whether they can be changed. Can formal semantics be embedded in a different, more justified set of assumptions about the nature of language and

[2]This is not to imply, of course, that actually giving such an account is by any means easy. It involves a lot of ingenuity and some hard thinking as any practitioner of the discipline can testify.

[3]A case in point being the debate about compositionality and representationalism, which obviously rests on the (implicit) assumption that the 'inner apparatus' of a semantic theory makes empirical predictions. One question that immediately rises is: Which features of our theories is it that we consider to be 'realistic'? The concepts used? Just the mechanisms and rules proposed? Clearly, this is an issue that needs to be addressed in a systematic fashion, but which hardly ever is discussed at all.

competence? Or can such assumptions be dropped altogether, i.e., can formal semantics be done in a philosophical 'void'?

3 Radical interpretation

Radical interpretation poses a simple but real problem: How do we understand the utterances of a speaker in a tongue that is completely foreign to us? We know nothing of her language, her convictions and beliefs, her customs and other ways of behavior. At first we do not even know how to segment the sounds she utters, how to determine the boundaries between the units of her speech. Is this an exclamation, a word, one word or two? A sentence, or a series of such? Does she make a statement, ask a question, express bewilderment or firm conviction? A lot needs to be established, and anyone who has actually been confronted with speech in a completely foreign tongue knows how insurmountable the difficulties seem at first, how impossible the task at the outset appears to be. And a lot may go wrong right from the beginning. When Cook and his company landed in Australia and were confronted with some specimens of one of the characteristic local species, they asked the aborigines what they were called. The answer was 'Kangaroo', and by that name they go today. But the aborigines call the species in question by an entirely different name, and in the aboriginal language 'Kangaroo' is not even a name, but a sentence. Such stories are well-known, and although the misunderstandings seem harmless and actually quite funny, the questions they raise are deep and important. Of course, these are misunderstandings we have discovered, cases in which we have found out that and where things went wrong: But how did we manage to do so? And is there any reason to think that such procedures will always work?

And the problem is not just empirical: philosophical issues arise at every step, and it is these philosophical questions that take the problem of radical interpretation right into the heart of our own speech. Sure enough, when conversing in our own tongue with our fellow speakers we take it for granted that we all mean the same things with the same words and sentences, that we communicate thoughts and conjectures, desires and queries, using a common tool: a shared language with shared meanings. But on what assumptions is this conviction based, and how justified are they? Is there any reason to think that meanings are shared the way we think they are? And what exactly is it that we share? What kind of object is a meaning that we can share it, within a language and across languages? Or is it not objects that are shared, and is meaning to be located on an entirely different ontological plane? And how exactly is meaning tied to linguistic entities, to our words and phrases? Are they merely conventional signs for something that is in essence not of a linguistic nature? Does an interior intentional act bestow meaning upon the words and phrases that we use? And how is it that some inner mental process or action is able to link two such utterly different entities as a meaning and an expression? For expressions are physical entities, written signs or spoken sounds, and meanings are not like that at all. Or is the meaning somehow located in our minds and fused into the sign as we produce it, the way we slip a letter into an envelope?

It is questions such as these that have been the central concern of much of twentieth century philosophy of language, and the various answers that have been proposed to them, have shaped, in one way or the other, the semantic theories that make up the confusing multiplicity that the field exhibits, even today. The fact that the problem does not seem to have been solved (and are philosophical problems ever? Here Van Benthem (cf. above) obviously has a point) should not blind us to the progress that has been made in the interim. Sure enough, several positions on these issues are still held today, and positions which are quite incompatible at that. But as for the ways in which they are defended, quite a lot has changed there. So it seems that we have come some way to a better understanding of the issues involved, at least in this sense that the questions we ask are deepened, and the arguments we bring to bear upon them are sharpened.

The most celebrated form in which the problem has been presented, is probably that of Quine's 'thought experiment' of radical translation (cf., [6, chapter 2]). Here we are invited to imagine ourselves trying to come up with a translation of a language of an unknown people, whose culture is completely unfamiliar, and whose tongue bears no affinity to our own. Starting from premisses as parsimonious as possible Quine allows us information about observable happenings in the immediate environment, the verbal behavior of this people displayed in reaction to these events, and a little introspection, into our own verbal reactions. Only a particular kind of utterances, those immediately tied to sensory input and more or less commonly agreed upon, are up for translations that can be verified. The rest, Quine tells, us, is 'theory': we construct the remainder of the translation manual on the basis of so-called 'analytical hypotheses', of which there will be many, mutually incompatible, yet all in accordance with whatever data we can gather.

Although Quine originally formulated the idea of radical translation as a philosophical thought experiment, we do well to notice that it has a real, empirical counterpart. For example, if we describe the case of (first language) learning without any pre-conceived ideas of the underlying mechanisms, it seems to fit Quine's bill quite accurately. All the child has to go on is its perceptual input from the environment, the verbal behavior of the adults surrounding it, and, of course, its own mental make-up. Now, it is important to note that we would be begging the question if we were to assume that the latter already contains so much content and structure that the problem is solved: that is exactly the issue that is at stake. Of course if we postulate innate mechanisms, introspective access to language independent meanings, and the like, the problem becomes solvable, — well, in principle. But precisely such postulates are not philosophically harmless, to say the least, and create problems of their own. Likewise, if we were to say that, of course, formal semantics can not account for the situation of radical translation, or language acquisition, but has to be looked upon as a theory that applies to the 'end stage', so to speak, we do indeed set aside these worries. But given that radical translation, language acquisition (and, as we shall argue shortly, interpretation in general) has priority, at least chronologically, how should one evaluate a theory about the result that has no way of accounting of how it came about in the first place?

But there is another, and perhaps brighter way to look at this predicament. For why exactly is it that the task seems so formidable and the result so hopelessly inadequate? I would like to suggest that some hidden premisses are at work here, in fact the very same assumptions we indentified earlier, viz., that language is an entity that is somehow surveyable in its entirety, and that competence is an individual matter.[4] Given these assumptions the results that Quine tells us are all we can get, certainly fail to make for an adequate theory, of translation, and ipso facto of meaning for the language translated. But we might turn the table and take the initiative by asking ourselves the following question: If that is all we can get, perhaps that is because that is all there is to it? Indeterminacy of translation, in Quinean terminology, may simply be due to indeterminacy of meaning, of their language, and of our own.

This point can be reinforced by considering the case, not of radical translation of one vernacular into another, but of radical interpretation within the same vernacular. If we can bring ourselves to refrain from the assumption that every individual language user enters the situation of interpretation equipped with a complete theory of the structure and content of the language used, it seems quite natural to conclude that whatever meaning we attach to utterances of others, and thus whatever theory of meaning we come to possess, is the result of (acts of) interpretation, and not a tool that we bring along right from the start. This is, I gather, what Davidson's idea about 'prior and passing theories' (of meaning and language) amounts to (cf., [2]). As users of a language we enter a situation of interpretation with some partial theory of meaning, but rather than the norm to which subsequent utterances have to conform and which determines their interpretation, it is a theory that is changed by the very act of interpretation. Of course, some more or less stable generalizations may emerge over time, but in principle these are all up for modification and they do not constitute anything like a theory of meaning in the traditional sense. There are some obvious parallels between this view and the hermeneutical approach to interpretation as it can be found for example in the work of Gadamer (cf., [3]), but space does not permit me to go into that here.

4 What semantics should be

It is obvious that if we take the very idea of radical interpretation seriously our entire conception of what semantics is about, will change quite radically. The idea of language and its meaning as an entity which consists of a potentially infinite number of sentences each of which has a definite meaning, goes by the board. Consequently, the principle of compositionality loses its pivotal status, as does the idea of a strict dichotomy between structural operations and lexical content. After all, the principle of compositionality and the concomitant distinction between structural semantic operations and basic lexical contents were the primary (though perhaps not the only) conceptual tools for solving

[4]Another pertinent assumption seems to be that of a particular kind of fit between language and world, but, however dubitable, we will leave that aside as it seems not equally relevant for the present argument.

the problem of giving an complete description of an a priori given, infinite language. But if language and meaning are not entities like that at all, much of the intuitive appeal of these notions is lost.

Another aspect that comes to the fore is the role of the environment. Interpretation is essentially a situated process, and if meaning is the result of interpretation, this situatedness will be all-pervading. Of course, indexicality and context-dependency have been a concern of formal semanticists, but it seems the phenomenon has not been conceived of in the right way. The standard view is that by 'adding' values supplied by contextual parameters an essentially non-contextual meaning can be assigned to indexical expressions. However, this still gives conceptual priority to a priori, non-contextual meanings. A similar observation can be made vis à vis the traditional dichotomy between semantics and pragmatics. Here, too, the main approach has been, at least for a long time, to explain the use of language on the basis of its meaning, which was supposed to be somehow independent of it. The recent 'dynamic wave' has gone some way to reverse the order of conceptual priority, but a much more fundamental switch seems to be needed. Let me point to one, simple example. In [4] it is shown that in dynamic modal semantics some questions of identification can not be solved unless the language contains demonstratives. But rather than postulating a split between linguistic and non-linguistic sources of information, this dependence on the environment should be taken right into the very notion of meaning. Only if that is achieved can we hope to get an account of the seamless way in which information from different sources is integrated. A similar lesson can, of course, be learned from philosophical work on direct reference (Kripke, Putnam) and on causal theories of reference.

Here we touch on a very fundamental issue. Language is essentially a human phenomenon, and as such an integrated part of human information processing. The way we obtain, process, and transmit information is highly marked by our biological, psychological and sociological make-up. In particular, the 'embodied' nature of our various means of dealing with information seems to be severely underrated in current approaches to semantics. Taking this feature seriously would, I gather, lead to semantic theories that will be much easier to integrate with other 'modules' of human information processing, such as vision.

One might think that this a plea to transform semantics into a study of the underlying (neuro)physiology of information processing.[5] But, at least as far as language is concerned, I think that would be a mistake. Let me point out very briefly two reasons why.

First of all, language is not only embodied, it is also essentially normative. The rules, of whatever formal nature they turn out to be, that govern our use of language do not simply describe our actual practices, they constrain them in a normative fashion. Any theory of meaning that does not account for this (and formal semantics as we know it certainly has a hard time doing so) is incomplete at the very least. Normativity shows that language and meaning are the result of a subtle and very complex interplay between our physical, psychological

[5] Note that such a move would be quite in line with some of the fundamental tenets of the Chomskyean view.

and social nature.[6] Of course, one might take a reductionist stance here and argue that normativity is merely an 'epiphenomenon': a concern for us, yet not characteristic for what language and meaning really are.

However, there is a second observation to make which makes it doubtful that this is a viable position. For this view runs into a problem which is in fact characteristic for much of the theories that have dominated linguistics, semantics and cognitive science since the sixties. It is what Jackendoff (cf., [5]) has called 'the mind–mind problem'. Where Descartes saddled us with the problem of accounting for the relation between the material world, including our bodies, and our disembodied minds, the cognitive revolution has doubled our worries: now we also have to account for the relation between our 'computational minds' and our 'conscious minds', between the 'automated', non-introspectable mechanisms that are supposed to constitute our cognitive abilities, and our phenomenological awareness of the world and of our minds.

Notice that the fact that certain processes (even some cognitive ones) *are* not introspectable as such (vision might be an example, respiration certainly is) is not a problem. But vis à vis meaning things are different. Here we have a realm that *does* have conscious 'counterparts'. We reflect on meaning, we worry and argue about it, we try to figure out what someone's utterances mean, and so on. How is this to be accounted for when meaning (language) would be completely sealed off from our conscious minds?

As we indicated above, one consequence of taking the idea of radical interpretation seriously is that meaning is the result, and not the toolkit, of interpretation. This reversal of the traditional order at least provides us with a handle on tackling this mind–mind problem. First of all, interpretation is a conscious process, at least in principle. That much of our everyday use of language takes place 'at the back of our minds' is not an issue. We employ our abilities consciously only if there is a need to do so: if we do not understand, if a misunderstanding arises, if we want to mislead (or try to be funny). Secondly, the need to postulate a complete description of an infinite language is dropped, and, along with it, the need to appeal to a mysterious mechanistic realm of the mind that completely accounts for our linguistic abilities.

An obvious rejoinder to the argument outlined above, is that it may (on a sympathetic reading) show some shortcomings, perhaps even rather serious ones, of formal semantics as we know it, but that it does not provide us with a better model. And that much is true. But my suggestion would be not to leave it at that and continue doing what we know rests on questionable assumptions, but to try and find out if we can do better. And one way of exploring what semantics could be is to put radical interpretation to the test. But how should we do that? A small step in the right direction, and one that is increasingly being taken, is to broaden our empirical basis, and operate on real data, such as large corpora, and to try to extract a semantic description from them. But this is certainly not enough to really model radical interpretation. One of the problems is that the nature of the target language determines the success con-

[6] For a very lucid statement of this argument, and a fundamental criticism of the Chomskyean paradigm along these lines, cf., [7].

ditions of the description to be obtained. This way of doing semantics always operates from within a fixed framework that makes certain basic assumptions about the source language, and, by that very fact, pre-judges the entire issue. And another serious shortcoming is the restricted nature of the data. As we argued above, the situatedness and the embodied nature of language and its users have to be taken into account if we are to come up with an adequate account.

So a better approach might be to try to build a robot which lives in a real environment, i.e., which has a variety of sensory inputs, among which speech from real language users, and which is forced to use (some approximation of) natural language as one of its means of communication. We could use different sets of 'hard wired' initial programs and datasets, and by studying the robots success in acquiring and using language under various circumstances get a clearer picture of the balance we have to strike between what we have to assume as 'given' and what is the actual result of interpretation by an adaptive, self-modifying system. Such an approach, I venture, would learn us a lot about the nature of meaning.

What role does formal semantics as we now practice it have to play here? I am not sure, but I do not doubt that it will lead to a transformation of the discipline. Which explains the title of this little exercise in philosophical speculation: I think we can be fairly sure what a theory of meaning should be, but I am not (yet) sure that formal semantics could be exactly that.

References

[1] Johan van Benthem. Why is semantics what? In Jeroen Groenendijk, Theo M.V. Janssen, and Martin Stokhof, editors, *Formal Methods in the Study of Language*, Mathematical Centre Tracts 135, pages 29–49. Mathematical Centre, Amsterdam, 1981.

[2] Donald Davidson. A nice derangement of epitaphs. In Ernest LePore, editor, *Truth and Interpretation*, pages 433–46. Blackwell, Oxford, 1986.

[3] Hans-Georg Gadamer. *Wahrheit und Methode*. Mohr, Tübingen, 1960.

[4] Jeroen Groenendijk, Martin Stokhof, and Frank Veltman. Coreference and modality. In Shalom Lappin, editor, *Handbook of Contemporary Semantic Theory*, pages 179–213. Blackwell, Oxford, 1996.

[5] Ray Jackendoff. *Consciousness and the Computational Mind*. MIT Press, Cambridge, Mass., 1987.

[6] Willard Van Orman Quine. *Word and Object*. MIT Press, Cambridge, Mass., 1960.

[7] Harry Stein. *The Fibre and the Fabric*. PhD thesis, ILLC/Faculteit der Wijsbegeerte, Universiteit van Amsterdam, Amsterdam, 1997.