Formal semantics and Wittgenstein: An alternative?

Stokhof, M.

Published in:
The Monist

DOI:
10.5840/monist20139629

Citation for published version (APA):
FORMAL SEMANTICS AND WITTGENSTEIN: 
AN ALTERNATIVE?

ABSTRACT

This paper discusses a number of methodological issues with mainstream formal semantics and then investigates whether Wittgenstein's later work provides an alternative approach that is able to avoid these issues.

1. Introduction

Formal semantics is an example of a relatively young, but very successful enterprise. It originated in the late sixties and early seventies of the previous century from the efforts of philosophers, linguists, and logicians who shared an interest in the semantics of natural language and wanted to explore the possibilities of applying the methods of logic to this area. Spurred in part by the success of the generative paradigm developed by Chomsky, which had revolutionised linguistics in the fifties and sixties, the concept of a formally rigorous study of natural language meaning, based on an equally rigorous analysis of its syntax, seemed a promising enterprise. Thus the work of philosophers and logicians such as Davidson, Montague, Lewis, and Hintikka, combined with that of linguists such as Partee, Bartsch, Keenan and others, started to define a paradigm that in its basic features still stands today. Of course, formal semantics has seen a lot of further developments and it has grown into a many-varied discipline in which a number of theoretical frameworks are being explored and in which, next to the more standard type of descriptive and analytic work, there is an increasing interest in experimental work as well. Yet some basic theoretical principles and methodologies still stand, and they are what characterises the discipline of formal semantics as an intellectual unity.

But how does the success of formal semantics sit with other approaches to natural language meaning that have been explored in the past, and that are still claimed as relevant today, such as that explored by
Wittgenstein in his later work? The question is not just of importance from an external perspective, it is also relevant from within formal semantics itself. For the variety of approaches that characterises present-day formal semantics does raise a number of questions concerning some of its core concepts, in particular those of meaning and semantic competence, and the methodologies that can be used to study them.

To explore all these issues in full detail is of course beyond the scope of a single paper, so in what follows we will focus on just one particular issue, viz., the question whether a Wittgensteinian concept of meaning can be a relevant alternative for formal semantics.\(^2\) We feel this question is relevant for a number of reasons. First of all, with its emphasis on the role of the (social) context and its focus on linguistic behaviour, the Wittgensteinian conception seems to be an almost natural alternative to the logical conception of meaning and the strictly individual concept of semantic competence that formal semantics endorses. Second, apart from considering the Wittgensteinian conception as a straight alternative, a closer look at its status might help to gain a better understanding of the status of formal semantics.

We will proceed as follows. In section 2 we will first outline three core characteristics of formal semantics that have shaped the standard model of formal semantics and that are still present across the variety of theoretical frameworks that are around today. In section 3 we will discuss some core questions concerning conceptual and methodological issues with formal semantics that are raised by these characteristic features. Then, in section 4 we will look at the possibility of a Wittgensteinian alternative for formal semantics. Finally, in section 5 we will evaluate the differences and outline conclusions concerning the status of formal semantics to which they might lead.

Before proceeding, one caveat might be in order. The aim of what follows is not to ‘show’ (or ‘prove’) that formal semantics is ‘wrong’ or ‘misguided’, or ‘uninteresting’, or anything like that. Rather it is to investigate what formal semantics is, and how some of its problems might be solved or avoided by creating a better picture of what it is and does.

2. Formal Semantics: Some Core Characteristics

When one looks at the work that is being done in formal semantics today, one is struck by what at first sight looks like a substantial theoreti-
cal and conceptual diversity. Thus, some frameworks are firmly rooted in a sentence-based outlook on grammar, whereas others focus on discourse (texts) as the main unit of analysis, and yet others, such as game-theoretical approaches, focus on the linguistic exchange (e.g., a question-answer sequence) as the primary entity. Also, there does not seem to be a shared set of methodological principles: intuition-based description and analysis still account for a major part of the work that is being done, but other methodologies are on the rise. Corpus-based studies have been around for a while, and recently computational modelling and experimental studies have been added to the mix.

Usually, this diversity is not regarded as particularly problematic, and is often explained by pointing out that they are merely different ways of addressing the phenomena that semanticists are interested in. Be that as it may, what does seem puzzling to us is that there is no firm consensus on what constitutes a proper conceptualisation of the core phenomena. Thus we find meaning described in terms of truth-conditions (intensionally or extensionally conceived), as constituted by assertability conditions, characterised in terms of update conditions or context-change potentials, analysed in terms of inference potential, and so on. And then there is the added dimension of speaker’s meaning and conversational implicature, and the concomitant discussions about the dependence between such notions and literal meaning (if such is acknowledged as a bona fide entity to begin with).

Especially the latter kind of diversity creates confusion: if two semantic theories, or two semanticists, do not agree on what meaning is, their results threaten to become substantially incomparable. If there is not sufficient agreement about the nature of the core phenomena, we are not dealing with ‘simple’ theoretical diversity, but with something much more complicated and confusing. In its turn this confusion also leads to uncertainty about the way in which formal semantics relates to other disciplines. One example, which will be discussed in more detail later on, relates to the competence-performance distinction. Theories that subscribe to this distinction and that start from the assumption that it is linguistic (semantic) competence, and not actual performance, that is their proper subject matter, have a hard time coming up with clear predictions that are testable by experimental methods, e.g., in psycholinguistics or in neuroimaging studies, since the latter deal with actual performance (albeit under definitely artificial restrictions).
What has created this situation? In order to answer that question, we have to trace the frames of thought from within which formal semanticists work, i.e., we have to uncover the basic concepts and principles, investigate their origins and justifications (if any), and look at the stated goals of the disciplines and the methods that are used to achieve them. In doing so, we focus on textbooks, surveys, and lecture notes as our main sources. The reason is that in the ongoing descriptive and analytic work, the tricks of the trade are hardly, if ever, discussed explicitly. And for good reason: in reporting on actual research these are being applied, not investigated. Systematic methodological reflection is rare, and occurs usually only at the 'edges', either historically, or systematically. But in the introductory textbooks, overviews, and sundry materials, it is the framework itself that is being introduced, explained and illustrated. And it is here that we might expect to find the concepts and principles to be formulated most explicitly, and their justification and their relations with goals and method explained most distinctly.

So what are these basic concepts and principles that have shaped formal semantics and that continue to be relevant for an assessment of its present status? Among the core characteristics of what we might call the 'Standard Model', we count methodological individualism, the distinction between grammatical form and logical form, and methodological psychologism. Space does not permit us to discuss these features in very much detail, but a quick overview is necessary for a discussion of the possibility of a Wittgensteinian alternative.

'Methodological individualism' refers to the individualistic nature of some of the core concepts in formal semantics. A good example is the central role of compositionality as a generally accepted constraint on semantic descriptions and analyses. One way of reconstructing its importance starts with an assumption about the nature of language that is characteristic of both the generative tradition in syntax as well as formal semantics, namely that languages are potentially infinite objects. The concept of a language as an infinite set of strings generated by a finite set of recursive production rules over a finite lexicon is one that Chomsky took over from the work of Post and others in mathematical logic. The assumed recursivity of (some) rules in the grammar is often regarded as one of the defining characteristics of human languages. The same assumption was, of course, a natural one to make for logicians who started to apply their tools to natural languages.
In terms of language users the infinite nature of language reappears in the form of their so-called ‘creativity’. As speakers, competent users of a language, it is assumed, are in principle able to generate an infinite number of well-formed expressions and to assign them a meaning, and, conversely, as hearers they are able to recognise (parse and interpret) these expressions and their meanings for what they are. On the assumption that it is individuals that are competent languages users, this creates a problem, which is that (interpreted) languages don’t fit inside the head (brain, mind), which is accommodated by a shift from language to grammar. After all, mastery of a finitely representable set of grammar rules is a viable candidate for an individual property, but it will serve its purpose only if the grammar satisfies the constraint of being compositional. Thus we observe that a central and characteristic property of descriptions and analyses in formal semantics derives from an assumption that is inspired by work done in an unrelated discipline.

The distinction between grammatical form and logical form tells a similar story. The idea that for philosophical and scientific purposes we need a rigorously defined language that is precise and unambiguous, and that provides us with the means to formulate arguments and positions in a manner that will allow them to be scrutinised in an objective and decisive manner, has a long history. For formal semanticists in the early day, the main reference point presumably was the work of Frege, who in the preface of his *Begriffsschrift* complained that: 7

... I found the inadequacy of language to be an obstacle; no matter how unwieldy the expressions I was ready to accept, I was less and less able, as the relations became more complex, to attain the precision that my purpose required. This deficiency led me to the idea of the present ideography.

A distinction between grammatical form and logical form that is motivated along these lines is basically a philosophical, not a linguistic distinction, not a distinction that is supported by empirical observations. Yet, it is a concept that has survived in some form or other in formal semantics as it is so congenial to the use of unambiguous formal languages as models for interpreted natural languages and also presents itself as a prerequisite for the application of model-theoretic semantics. It does come with some quite particular assumptions on the nature and accessibility of natural language meanings. Elsewhere,8 we have argued that the distinction between grammatical form and logical form rests on what we called the ‘Availability
Assumption', which holds that meanings are available independently of their being expressed, in a natural language or in a formal language. Only on that assumption does it make sense to judge the adequacy of expression, to compare two different expressions, and to use one expression as a formal representation of the meaning of another. And especially the latter is the daily work of the formal semanticist.

The content of the Availability Assumption points in the direction of the third characteristic mentioned above, viz., that of methodological psychologism. This is a feature of the standard model of semantics that it shares with many other approaches in linguistics: the reliance on intuitions of competent (native) speakers of a language as data and test bed for description and analysis. The origins of methodological psychologism reside in a particular construction of a core concept of modern linguistics, viz., linguistic competence. This concept is constructed as the competence displayed by:

[... ] an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest and errors (random or characteristic) in applying his knowledge of the language in actual performance.

From this perspective, the intuitions of a native speaker can count as reliable manifestations of his or her underlying competence, and hence as the primary method to be used in description and analysis. Intuitions being both individual and mental, this put a particular form of introspection firmly at the core of the methodology that semantics has relied on almost exclusively in its first two or three decades. And it continues to form the basis of much empirical work, despite the fact that the use of experimental and corpus-based methods has increased.

3. Formal Semantics: Some Core Questions

The three characteristics outlined in the previous section are not independent: the individualism that turns compositionality into such a central constraint also informs the psychologism that is a feature of the main methodology that formal semanticists employ. And the distinction between grammatical form and logical form really makes sense only from an internalist perspective that turns meanings into an independent, mental type of entity. And together they all illustrate the multidisciplinary origins of formal semantics, which has its roots in generative linguistics, in logic, and in philosophy.
The last observation gives rise to a perhaps surprising yet genuine question: what kind of discipline is formal semantics? Is it an empirical discipline, as presumably most semanticists would claim? Or should it be considered a formal-conceptual enterprise, as two of its godfathers, Davidson and Montague, conceived of it? Or is it perhaps more of an engineering type of enterprise, one that is used in application-oriented work, such as machine translation, as a formal framework in which one can state system requirements?

At first sight it seems odd that the question can be raised in the first place, since the answer seems so obvious: language being an eminently empirical phenomenon, how could linguistics, including semantics, the discipline that studies this phenomenon, be anything but an empirical enterprise? However, things might not be quite as obvious as that. (Which is not to suggest that any of the other answers is obviously the right one.) In what follows we briefly introduce two considerations that suggest that things might not be as straightforward as they seem. The first issue concerns the way in which the core concepts of semantics are constructed: is this a matter of idealisation or abstraction? The second issue relates to the reliance on intuitions as a core methodological tool: can an empirical semantics be based on a ‘methodology of intuitions’?

Any empirical theory deals with the phenomena that constitute its object by means of constructions: core concepts that the theory is about and that are taken to correlate with the actual phenomena. Such constructions are both needed and desirable. No theory can deal with the flux of events as they actually happen, since these do not constitute in any way a sufficiently conceptualised, coherent whole. And no theory wants to deal with actual events in this sense, since theories are not supposed to mimic reality in the sense of event tokens, but rather to explain it in terms of general, preferably nomological, connections between event types. Thus, in semantics we deal not with the myriad of actual utterances which we encounter in real time but with ‘objects’ such as language, meaning, and competence.

In dealing with such constructions we can distinguish between two types: abstractions and idealisations. Abstractions are constructions that abstract over some quantitative parameter by setting it to some numerical value that does not come from observation or experiment. Examples from physics are a frictionless plane, a perfect vacuum, a perfectly rigid rod. Such objects don’t exist in reality, but in many circumstances we can construct them because the feature in question is either not relevant (e.g.,
for a particular application), or intractable, or otherwise not ready to be incorporated in the theory. Idealisations, on the other hand, are concerned with qualitative features of a phenomenon that are left out, i.e., that do not enter into the constructed concept. A relevant example for our concerns is the concept of linguistic competence. As the quotation from Chomsky that was given earlier illustrates, the concept of an ‘ideal speaker-listener’ leaves out a number of features that are characteristic for actual speakers-listeners, such as their interactions with their environment, their goals, their embodied nature, and so on.

One way in which the difference between an abstraction and an idealisation makes itself known is by the effect it has on the resulting theory. In the case of an abstraction we can in principle always go back from the theory with the abstracted concept to the actual phenomena: the theory will make the wrong predictions as far as the abstracted feature is concerned, but this will then be apparent from subsequent observation and experiment, and they can, at least in principle, be incorporated as characteristic features of the object in question in a more encompassing theory. In the case of an idealisation, however, there is no such ‘back-and-forth’ between the theory with the abstracted concept and the phenomenon as it appears in observation and experiment because the constructed object lacks the feature in question, and hence the theory containing it does not state or imply anything about the feature that is ‘idealised away’.

The consequence can also be stated thus: construction by means of abstraction is wilfully ignoring a feature that is however still considered part of the phenomenon; construction by idealisation is quite literally ‘changing the subject’.

The word ‘language’ has highly divergent meaning in different contexts and disciplines. In informal usage, language is understood as a culturally specific communication system. In the varieties of modern linguistics that concern us here, the term ‘language’ is used quite differently to refer to an internal component of the mind/brain. We assume that this is the primary object of interest for the study of the evolution and function of the language faculty.

Other examples of concepts in the construction of which some kind of idealisation seems to be operative are that of language as an infinite object, that of a language user as a disembodied individual, the focus on written
language and the relative neglect of speech, and, of course, the core concept of semantics, meaning. Thinking about the wide variety of conceptions of meaning that are around as the results of idealisation, rather than abstraction, goes some way into explaining the fundamental diversity, including the incomparability of results obtained in various frameworks that seems to have become a mark of semantics today.

Let us now turn to the second question, that of the possibility of basing an empirical semantics on a methodology of intuitions. The standard model of semantics draws on three central conceptions: intuitions, semantic facts, and semantic competence. The intuitions that are deemed relevant for the description and analysis of semantic phenomena are those of a native speaker of the language under consideration. In many cases the semanticist himself or herself may play that role, but in other cases he or she has to rely on native speakers that act as informants. The semantic facts concern such properties and relations as entailment, synonymy, analyticity, ambiguity, and so on. These are considered to be properties of and relations between the expressions of a natural language. Semantic competence, finally, can be found referred to in a number of ways. It may appear in a characterisation of whose intuitions count: a native speaker is often also referred to as a competent speaker. Or it may be part of a characterisation of the object under investigation: what semantics is supposed to describe and explain is semantic competence.

What is important to note here is that all three concepts—facts, intuitions, and competence—are conceptually related. And that has important consequences. For one thing, it means that competent speakers can not be wrong: after all it is their intuitions that define the object of study, and that is what it is, there can not be a wrong or right about that. Another consequence is that for a semanticist who describes and analyses his or her own language there can not be any discoveries. And it also follows that in every case of conflicting intuitions (and anybody with a working knowledge of the semantics literature knows that there are many such cases) in the end can only be resolved by an appeal to idiolectal variation. But that is a sure sign that there is something wrong with the empirical status of the data that this methodology of intuitions is concerned with.

A key concept in all this is that of competence. This idealisation plays a crucial role in determining both what semantics is about and how
semantics should be conducted. Note that if one maintains the distinction between competence and performance, then, by definition, competence has priority. After all, it is introduced to define the object of study, so whatever facts one takes to be established about competence, they will have to have priority over what one may actually observe about linguistic behaviour: in case of conflict, competence trumps performance. On the other hand, due to the conceptual nature of the links between competence and other central concepts, such as meaning (semantic facts), should one decide to abolish the competence-performance distinction and focus on actual linguistic behaviour as the main phenomenon to be investigated, it is not just competence that goes, but other idealised concepts, such as meaning, go with it, too. For similar reasons the central role of intuitions as the main tool would be undermined, as well. We may have ‘intuitions’ about actual behaviour, but these are of a completely different kind. An intuition about behaviour is a hunch, a lead that one may follow up on, but always something that stands in need of corroboration by observation. That is quite unlike the intuitions that derive from competence: those are the data, not hunches about what the data could turn out to be.

So the idealised concept of competence is crucial for the methodology of intuitions. But the role it plays in cementing the various concepts together in the end turns out to be something of a conjuring trick. The crux of the matter is that intuitions are supposed to play two distinct and mutually incompatible roles at the same time. On the one hand, intuitions, being reliable, even incontestable manifestations of competence, are what a semantic theory is about. This we can call ‘intuitions-as-objects’. But on the other hand intuitions are also the primary access to what the theory aims to describe and analyse, viz., semantic facts. This we can call ‘intuitions-as-data’. Obviously, an apparent circularity is just one substitution away. . . . What the standard view requires, then, is a genuine Von Münchhausen trick. However, needed as that may be to bootstrap the theory, the result can not be a stable configuration.

So there really is a conundrum here, and the source of it resides in the core characteristics discussed in section 2. The combination of individualism and psychologism (mediated by the assumption of availability of meanings) leads to the idealisation of a homogeneous, individually situated competence that can be accessed by consulting equally homogeneous and individually situated intuitions about available semantic facts.
This strong form of methodological psychologism has to be weakened, and that will involve changes in the way in which core concepts are constructed. Here are some reasons to think so.

First of all, as we noticed above, any difference between intuitions of native speakers has to be accounted for in terms of idiolectal variation. That may seem just a matter of the methodology of intuitions as such, but it is not: it also concerns the concepts involved. For if we apply more empirical data gathering methods but still consider the concepts as strictly homogenous, we end up, theoretically at least, in the same situation. Suppose we examine the judgments of a randomly selected group of native speakers, by means of a questionnaire or via an experiment, on the purported ambiguity of some (type of) expression, and find that 80 percent of them consider it to be ambiguous, but 20 percent do not. Then what is it? Ambiguous? Nonambiguous? We could resort to an appeal to dialectal variation, of course, but then we run the risk of having to postulate more and more such variants with each new difference in judgements that we find, with in the limit as many dialects as there are competent speakers.

A second kind of case concerns the role of the grammar that is being described and analysed. If some construction that is predicted by the grammar does not occur in any corpus of actual data (text, speech), is it part of the language? That is, do we construct the core concept of language in terms of its grammar— itself a highly theoretical construct— or in terms of what actual linguistic material we can observe to be produced ‘out there’? An even starker contrast arises when we note that in a great many cases actual communication is not hampered at all by the lack of well-formedness—syntactic, semantic, and otherwise—of the utterances exchanged. It is only by completely dissociating language from what people use in actual communicative exchanges (witness the move made by Hauser, et. al., in the passage quoted above) that can rescue the notion of well-formedness from irrelevance.

The first observation questions the presumed homogeneity of concepts such as meaning, ambiguity, and the like, that is at the root of the methodology of intuitions, and suggests that we should take its observational heterogeneity seriously. The second one goes further and questions the very usefulness of concepts such as well-formedness in the first place, and suggests that as a concept it is not even a construction of a phenomenon, but a by-product of the theoretical framework.
What should be noted is that problems such as these as such do not favour a thoroughly externalistic and socially situated approach. But what they do show is that it minimally has to be acknowledged that beside whatever cognitive aspects of language and meaning there are that are rooted in individual psychology, there are other aspects that need to be taken into account as well. And that means a shift, not just in the object of study, the heterogeneity of which has to be faced, but also in the methodological tools that can and must be applied for studying the phenomenon in full.

So, let’s now turn to the second question we want to address in this paper: What about a Wittgensteinian alternative for formal semantics?

4. A Wittgensteinian Alternative?

The question whether a Wittgensteinian approach to language and meaning—with its emphasis on the behavioural aspects of language use, on externalism with respect to meaning, and on the role of the social context—might constitute an alternative framework that is able to deal with the observed heterogeneity on a more principled level, seems a natural one. As a matter of fact, generative linguistics and formal semantics have been criticised from a broadly Wittgensteinian perspective by a number of authors. And in philosophy of language, though not in empirical linguistic semantics, attempts have been made to develop Wittgenstein’s ideas about meaning and use into an explicit theoretical framework. Both are not without problems, so before trying to address the question of the possibility of a ‘straight’ Wittgensteinian alternative head-on, we would first like to very briefly discuss one or two examples of these attempts, so as to get a better grasp of some of the potential pitfalls of such an enterprise.

4.1 Preliminary concerns

Early criticisms of generative linguistics and formal semantics have been put forward, for example, by Gordon Baker and Peter Hacker in their 1984 book *Language: Sense and Nonsense*, and by Bede Rundle in his *Wittgenstein and Contemporary Philosophy of Language*, from 1990. The main characteristics (for our purposes, anyway) of these critical engagements is that they are object-level comparisons and wholesale rejections, rather than metalevel investigations into assumptions and presuppositions.
To start with the dismissiveness, the following passages hopefully illustrate what is at stake here. First, Baker and Hacker: 14

We have argued that the basic problems [that modern theories of meaning seek to address, MS] are all bogus [. . . ] Each of these questions [. . . ] makes no sense [. . . ] Truth-conditional semantics suffers from a dire disease: it is at a loss to find any genuine problems.

This leaves little room for doubt about the appreciation Baker and Hacker have for the work done in this field. Next up, Rundle: 15

[T]he more that comes from the pens of those who derive their inspiration from Frege and subsequent work in formal semantics, the more the inadequacies in this approach are exposed.

Obviously, not much love is lost between these authors and those working in generative linguistics and formal semantics. It would not be difficult to find equally dismissive statements from generative linguists and formal semanticists about broadly speaking Wittgensteinian ideas. But that's not the discussion we want to get into here.

What makes this line of criticism uninteresting for our purposes is not the sweeping and dismissive conclusions, but the way in which these conclusions are arrived at. The discussions in the books of Baker and Hacker, and of Rundle, centre around certain descriptive or theoretical problems, and usually take the form of critiquing in detail the descriptions and analyses that linguists and semanticists have provided of them by arguing that they lead to all kinds of metaphysically loaded conclusions, that from a Wittgensteinian perspective should be considered as mistaken or misguided. Now, the various claims may or may not have a point from the perspective of what these authors take to be Wittgenstein's stance on metaphysics, but that is not what is of concern to us here. We are not interested in the critical potential of the Wittgensteinian enterprise as far as formal semantics is concerned, but in its positive potential of providing an alternative framework that is able to avoid certain of the problems that formal semantics runs into, quite on its own, without needing the help of Wittgensteinian critics such as Baker and Hacker, and Rundle.

From our perspective the second reason why these critical engagements are less interesting is that no attempt is made by these authors to investigate the assumptions of formal semantics from a historic perspec-
tive or to analyse them from a systematic point view. Actually, by ‘bashing’ the descriptions and analyses of formal semantics simply by assuming that whatever Wittgenstein had to say about what appears to be the same phenomenon is right, presupposes, rather than argues, that formal semantics and Wittgenstein are engaged in the same kind of enterprise, and that the latter simply has the better insights. However, as we shall argue shortly, that formal semantics and Wittgenstein are indeed pursuing the same goals appears not to be the case. By proceeding in this way, these critical studies, despite the fact that they contain sharp observations and justified criticisms as well, fall short of providing room for any kind of alternative, Wittgensteinian view of semantics quite generally seem to be not in accordance with the main gist of Wittgenstein’s work.

So what about approaches that explicitly do try to distill a theoretical framework from some central insights in Wittgenstein’s work? One prominent example is Horwich’s ‘Use Theory of Meaning’ (‘UTM’ for short). The main goal of UTM is described as follows: 16

[... ] to define a core notion of meaning of a word (distinct from speaker’s meaning, truth-conditional meaning, conversational implicature, etc.) as an ‘idealised law governing the use of a word’ in terms of acceptance conditions.

While this hits a number of right notes, UTM does have two characteristics that disqualify it as a genuine ‘Wittgensteinian alternative’ (which, to be sure, is explicitly not how UTM was intended by Horwich.) So what follows is not intended as a criticism of UTM as a viable approach to natural language meaning, it is only meant to discourage taking UTM as somehow embodying the alternative to formal semantics ‘that Wittgenstein would have endorsed’. The two features of UTM that we think make it not a plausible candidate for a Wittgensteinian alternative are its explicitly aiming to be a theory, and its subscribing to methodological individualism.

To start with the latter, we have argued above that methodological individualism in the end is not a viable position to take. Leading as it does to a form of methodological psychologism, it leads to an unrealistic construction of central concepts that simply do not do justice to the heterogeneity that we can observe. But it also seems quite at odds with Wittgenstein’s own take on what linguistic competence is. For Wittgenstein emphasises the externalistic influences and sources that go into constituting meaning,
and that, hence, lead to a view on competence that combines individual capabilities with external, socially determined constraints. Meaning, and hence competence, in Wittgenstein's approach is never a purely individualistic affair: it always also factors in environmental and social contributions.

As for UTM aiming to be a theory, that, too, is a problem from the Wittgensteinian perspective, of course. We will have more to say on this issue later on in section 5, so for now, suffice it to observe that the very idea of a theory of meaning that satisfies constraints that we usually associate with a theory, viz., a characterisation of a field of empirical phenomena in terms of law-like generalisations that allow for explanations and predictions of individually observed events, is quite alien to Wittgenstein's philosophical aims. Philosophy, he famously stated, "simply puts everything before us, and neither explains nor deduces anything." So from this perspective, too, an approach like Horwich's UTM, its many positive contributions not withstanding, is at odds with what we could acknowledge as 'a Wittgenstein alternative'.

4.2 Wittgenstein on meaning and use

But, one might ask, did not Wittgenstein himself provide us with a 'meaning is use' theory? Is this not what he is famous for, and is 'meaning is use' not one of his main contributions to our understanding of meaning? This is quite a commonplace conception, not just among the general public in philosophy, but also among some of those Wittgensteinians that have expressed sharp disagreements with formal semantics as a viable enterprise to natural language meaning.

This raises at least two questions: What exactly is the relationship between meaning and use that Wittgenstein intended? And can that relationship form the basis of a theory, in a sense that is minimally comparable to the sense in which formal semantics is a theory?

To start with the last point, in Wittgenstein's later work the issue of theory is complex and involves both the status of science and its explanatory theories, and the possibility of theory in philosophy. The latter Wittgenstein rejects explicitly: there can be no theories in philosophy, since philosophy, at least the kind that Wittgenstein deems viable, does not aim at explanation, nor is it involved in the kind of generalities and essences of phenomena that the theories of science are after. Whether that
also means that philosophy is not concerned with providing new ways of understanding phenomena is an issue that has been, and still is, intensely debated. On the one hand there are those who see only a therapeutic function for philosophy in Wittgenstein's later work, with philosophy focussing exclusively on the clearing up of misunderstandings that lead to philosophical muddles. But there are also authors who, while not denying that philosophy has a strong therapeutic element, claim that Wittgenstein does leave room for a kind of philosophy that makes more substantial contributions to our understanding of phenomena. This is not the place to go into the merits of each of these positions.  

But such contributions would still be of a different nature than the results that the empirical sciences provide. As for the latter, it appears that where Wittgenstein was quite explicitly opposed to various forms of scientism, and was genuinely concerned with the exclusivity that some claim for the sciences as providers of understanding, this does not lead him to reject scientific inquiry as such. As long as the sciences are aware of their inherent limitations, as regards to both the kind of understanding they provide and the kind of phenomenon that they can provide understanding of, Wittgenstein seems content to let science and philosophy go their own way. Which does not mean that there are no interactions possible, but which does maintain that the two are not in direct competition.

From that perspective, it would seem that formal semantics, and linguistics in general, are 'safe', i.e., that whatever a philosophical analysis of the same phenomena would come up with, in principle could neither disqualify, nor substantiate, the results of formal semantics. The two enterprises, though engaged with the same empirical phenomena, simply do not provide comparable results, but rather different types of understanding altogether. Note that this also would imply that the kind of criticisms that Baker and Hacker and Rundle bring against formal semantics from a Wittgensteinian perspective would be 'out of order'. But, and this is a but that directly relates to the discussion in section 3, this would hold only if formal semantics is a properly empirical discipline. If the objects of formal semantics would be the natural and homogeneous objects they are claimed to be, semantics would be 'in the clear'. But as we have seen, there is ample reason to consider that not quite as straightforward a position as it would seem at first sight.

So let's now see how things stand with the second question identified above, viz., the relationship between meaning and use that Wittgenstein
indicated. In the original German text, the central passage from *Philosophical Investigations*, section 43, reads as follows:

Man kann für eine große Klasse von Fällen der Benützung des Wortes ‘Bedeutung’— wenn auch nicht für alle Fälle seiner Benützung—dieses Wort so erklären: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.

This is the passage that many authors will refer to if they want to ascribe to Wittgenstein a ‘meaning is use’ theory. Presumably, they do so partly because they are acquainted not with the German original but with the English translation by Anscombe, which is this:

For a large class of cases—though not for all—in which we employ the word ‘meaning’ it can be defined thus: the meaning of a word is its use in the language. [italics added]

The point is, of course, that the German ‘erklären’ does not mean ‘to define’, but ‘to explain’ (and ‘to declare’, as in ‘declare war’, or ‘declare one’s love’). And an explanation of a phenomenon is quite something else than a definition. The latter aims to give a set of necessary and sufficient conditions that characterise the phenomenon in question uniquely, but an explanation can be something much looser and much more contextual. The mistake is odd also because Wittgenstein, with his introduction of family concepts and his insistence that philosophy should not look for essences, clearly holds that a definition of meaning as use, or any other definition for that matter, is not within the bounds of proper philosophy.

But odd as it may have been, and despite that fact that it has been noticed by many authors over the years,21 this mistranslation did stick around for a long time (it was in all the editions of *Philosophical Investigations* up until very recently), and has been instrumental in forming the idea among the general philosophical public that Wittgenstein held a ‘meaning is use’ theory.

It was only with the revision of the Anscombe translation by Hacker and Schulte that the mistake was rectified. Their translation reads as follows:

For a large class of cases of the employment of the word ‘meaning’—though not for all—this word can be explained in this way: the meaning of a word is its use in the language.

So the original Anscombe translation suggests that the connection between meaning and use that Wittgenstein makes amounts to an essentialistic characterisation of meaning. Such an interpretation is at odds with
other features of Wittgenstein’s work, and has to be ruled out on those grounds. The revised translation is more appropriate because, as was already pointed, it is closer to the German original, and because it appears to be in line with Wittgenstein’s practice.

Throughout the *Philosophical Investigations* (and also in other work of the later period), Wittgenstein often explains what a word means by describing instances of its use (along with, e.g., observations about how the word is learned). Such explanations, as was also already suggested, are more like descriptions than explanations in a scientific sense. They do not appeal to nomological causal connections between event-types, but rather make their point by describing a particular situation in which the expression is used that serves to illustrate a certain position, counter a misunderstanding, and so on.

But as such stating Wittgenstein’s position as holding that ‘meaning can be explained by looking at use’ still underdetermines the connection that is being postulated. We can interpret it maximally, i.e., read it as claiming that ‘the use is the phenomenon’. Or we can give it a minimalistic interpretation and read it as a methodological slogan: ‘look at the use’, i.e., don’t rely introspection, or experiment, or something like it.

The maximal interpretation takes the use of language to exhaust the phenomenon of meaning. In that sense it tends towards a form of ontological behaviourism: if we are to study meaning we can only, and need only, look at verbal behaviour, at the use that is being made of expressions. Such an interpretation is not quite like the ‘meaning is defined as use’ in that it does not suggest that the two are really the same entity. But it does come awfully close and, in a sense, this maximal interpretation of what it means to explain meaning by looking at use is too much a ‘return to theory’ to be a plausible reading of what Wittgenstein means here.

Apart from that, the behaviourism it implies is also at odds both with Wittgenstein’s practice and with his claims. That he is a behaviourist Wittgenstein explicitly denies, and if we look at his intricate descriptions and analyses of, e.g., our mental vocabulary, it is clear that Wittgenstein never identifies the corresponding phenomena with behaviour. He will argue that such phenomena ‘stand in need of outward criteria’, and that such criteria are often closely connected with characteristic forms of behaviour is clear. But a reduction to, and identification with, behaviour is never what Wittgenstein argues for or suggests.
The minimalistic interpretation of ‘explaining meaning by looking at the use’ reads it as a purely methodological statement. On this view it is not so much a connection between meaning and use that is made, but a shift of attention that is effected. It invites us to stop looking for some ‘thing’ that we can call meaning, and focus instead on the way expressions are used: that should suffice. The minimalism resides in this: that there is not only no characterisation of an underlying entity, but effectively no identification either of the phenomenon that we access by looking at use.

But that seems rather paradoxical. If looking at the use is just the methodology, then what is it that we study using that methodology? On the minimalist interpretation there is no phenomenon distinct from the methodology, there really is nothing that we apply the methodology to. But we do want an entity of sorts that is being accessed, if only indirectly, by looking at the use, in the sense that when observing the various ways expressions are being used we gain an understanding of what the meanings of these expressions are. Without resorting to an unwarranted reification we do expect phenomenon and methodology to be distinct.

After all, the justification of any particular methodology has to include a reference to the specific nature of the phenomenon that is being studied by that methodology. Also, the relevant aspects of that nature have to be accessible in ways that are independent of the methodology that we are trying to justify, otherwise a justification will never get off the ground. So we need some ‘pretheoretic’, common sense-like access to the phenomenon in question that we can use as a grounding. Of course, other factors may be involved in the justification of a particular methodology, as well, such as considerations that are derived from what we want from the investigation, in terms of special applications that may call for deliberate restriction to just certain aspects, and so on. But apart from and prior to that, a distinction between phenomenon and methodology seems required in any case.

Perhaps we should steer a middle way between the maximal and the minimal interpretations of the meaning-use connection. What could that look like? There are presumably a number of different alternatives that could be explored here, and it is quite likely that none of them on its own will be ‘the right one’. We would like to suggest though that it is imperative in all alternatives that we take the heterogeneous nature of the phenomenon of meaning very seriously. This means acknowledging that
what we call ‘meaning’ is both individual and social; internal and external; natural and socio-cultural; and so on. If we follow the close association between meaning and use that Wittgenstein’s work suggests, we can not but conclude that some aspects of meaning reside in the individual whereas others are determined by the community (or communities) to which the individual belongs; that there are aspects of meaning that are closely connected with mental content in the narrow sense, whereas others are intrinsically related to facts about the external environment; that there biological and psychological determinants of meaning, but also defining influences from the socio-cultural environment.

Acknowledging heterogeneity also means foregoing attempts to postulate one particular aspect as somehow basic or fundamental and trying to explain all others as epiphenomena that are the results of interactions of ‘meaning proper’ with context, application, and so on. Such attempts to postulate a homogeneous core beneath the heterogeneous surface would, from a Wittgensteinian perspective, be a lapse into the essentialistic thinking of old. But what does all this mean for formal semantics?

5. Semantics as Theory

So let’s finally look at what this Wittgensteinian perspective means for formal semantics and its aspirations to be a theory, in the sense of a systematic inquiry into a field of related phenomena. The question how formal semantics can deal with the heterogeneity of meaning that we discussed above will be central here.

We propose to make use of a distinction introduced by David Marr between two kinds of theories that reflects a distinction between the phenomena that are their subject matter. Marr is famous for his ‘three-levels’ hypothesis concerning the ways in which cognitive information processing systems can and should be understood.22 The first level is called the ‘computational level’: here the input-output conditions of a cognitive process are described, often in functionalist terms. Next is the ‘algorithmic level’, which consists of a specification of the procedure or procedures that derive the output of the system from its input, preferably deterministically. The third level, finally, is the ‘physiological level’, where the ‘wetware’ implementation of the process is accounted for. This is the general format that theories that deal with cognitive processes should take.

However, analyses of cognitive processes at all three levels are possible only if certain conditions are met. Let us illustrate this in terms
of the heterogeneity of the meaning phenomenon. Acceptance of the heterogeneity still leaves two possibilities open. The first is that in each concrete case of meaning all aspects of the phenomenon are relevant, albeit in different proportions in different cases. The second is that different aspects are relevant for different cases.

This is a subtle but nevertheless fundamental distinction. In the first case we are dealing with a heterogeneity of features that are still sufficiently bound together to allow for a unifying theory. It will be a complicated and internally heterogeneous theory, but it would lend itself to a complete computational specification and to subsequent analyses at the other two levels. This is what Marr calls a ‘Type-1 theory’. Type-1 theories thus deal with phenomena that are multifarious yet coherent and that lend themselves to the three-level methodology that Marr devised.

Per contrast, in the second case, where certain cases display some aspects that are missing in other cases and where there is no justification for holding that each aspect is somehow relevant in each case, we are dealing with a phenomenon that is truly heterogeneous and that rules out a unifying theory. Here the various accounts that can be given of the various aspects do not lend themselves to one computational specification, and thus also not to analyses on the algorithmic and/or physiological level. This is what Marr calls a ‘Type-2 theory’. Where the three levels methodology works for Type 1 theories, it fails for Type 2 theories. Here, what we get is much more piecemeal and descriptive, rather than unificatory and explanatory.

That sounds familiar. In addition, notice that if we are dealing with a phenomenon that allows only for Type-2 theories, but we still insist on having a Type-1 theory, we are more likely to construct the central concepts of our frameworks in terms of idealisation, rather than abstractions, and thus will be very likely to ignore crucial aspects of the phenomenon in question.

This seems to diagnose the problem with formal semantics and the relevance of the Wittgensteinian perspective, more specifically: it is a failure to acknowledge the kind of heterogeneous phenomenon that meaning is, and the insistence of having a Type-1 theory that leads formal semanticists (like other linguists, such as those working in the generative tradition) to work with idealised constructions and to base a methodology on those that, in the end, will lead to insurmountable conceptual problems. And it is here that the Wittgensteinian perspective provides a much needed counterpart. Wittgenstein once diagnosed problems in philosophy as follows:
A main cause of philosophical disease—a one-sided diet: one nourishes one's thinking with only one kind of example.  

That, it seems, can be repeated with regard to formal semantics, in a way. It, too, suffers from one-sidedness, perhaps not so much in completely disregarding the variety within the phenomenon of meaning, but in insisting that this variety can be 'homogenised' by means of idealised constructions. 

But facing the music of heterogeneity does not mean defeat, for it opens up a new way of looking at what it is that formal semantics does. For one thing, taking the fact that meaning has intrinsically socio-cultural aspects means that it is potentially a performative concept. Our reflection on meaning is (in part) constitutive of what meaning is. Some such reflection may be on the permeability of the division between the natural and the socio-cultural aspects of meaning, which means that investigation of natural aspects may come to be reflected in the socio-cultural aspects. What we find out about the natural mechanisms that underlie our meaning practices may come to be reflected in what we take meaning to be, and in what meaning actually is. And on another score, taking the heterogeneous nature of meaning seriously makes the distinction between linguistic and non-linguistic meaning both much more fluid and more productive. 

But this change in perspective does come with a change in how we can view formal semantics. As we already saw above, the Wittgensteinian perspective rejects the idea that philosophy and science are on the same plane, that they pursue the same goals (broadly conceived) but with different means. Rather, although often, though not always, concerned with the same phenomena, the two strive for a different kind of understanding. In such a conception there is also no real role for philosophy as a 'provider of conceptual systems', as for example Hacker conceives of it. So with science and philosophy being different in this way, what alternative view of semantics does a Wittgensteinian perspective like this suggest? 

One way to answer this question is to look for an answer to a different, but related one: How much of the framework of the *Tractatus* can be accounted for in the perspective of the *Philosophical Investigations*? Arguably, formal semantics shares a number of important assumptions with views on language, meaning, and reality, and the role logic plays that Wittgenstein developed in the *Tractatus*. The distinction between the surface, grammatical form of an expression and its logical form, the all-pervading referentialism, including the defining role of truth conditions, the assump-
tion that meaning is not only homogeneous but also universal in the sense that there can be one characterisation that applies to all (possible) languages, are some of the most important features that formal semantics shares with the Tractarian framework. This allows us to discuss the later Wittgenstein’s ‘criticisms on formal semantics’ without actually being anachronistic. For many of the criticisms that Wittgenstein vents in the *Philosophical Investigations* against his own earlier views in the *Tractatus*, either directly or indirectly via his critique of the Augustinian picture, can be considered as criticisms of formal semantics as well, provided they are related to the assumptions that the *Tractatus* and formal semantics share.

So, the question of what from the Wittgensteinian perspective the status of formal semantics might be, can, at least in part, be answered by addressing the question of what is the core objection that the *Philosophical Investigations* makes against the Tractarian framework. Does Wittgenstein mean that the *Tractatus* is ‘wrong-as-an-account-of-how-things-actually-are’? Or is his point rather that it is ‘wrong-as-a-general-theory’? The issue is complex and presumably does not allow for a univocal, straight answer. Yet we see at least three reasons for preferring the latter option, viz., that Wittgenstein’s main gripe with the Tractarian framework relates to its aiming to be one, all-encompassing theory. The first one is simple: if Wittgenstein’s problem with the Tractarian picture is that it gives a wrong account of empirical phenomena, then his criticisms should be empirical themselves, and his alternative view should be an alternative empirical account. But that is emphatically not what Wittgenstein is after, as we have seen. The second reason is that there is nothing in the extensive discussions of language games, rule following, and forms of life in the *Philosophical Investigations* that excludes moves in a language game, or perhaps even entire language games, for which the systematic elements of the Augustinian picture give an adequate account. (The learning part of the Augustinian picture is a different issue, that we don’t need to go into here.) So, it’s not that sometimes expressions do not have meaning because of what they refer to, or that truth does not play a role in some practices, and so on. In that sense, the Tractarian framework, and hence formal semantics, does not lack relevance. What is problematic, and this relates to the third reason for preferring the second option, is the purported generality of the Tractarian framework, its ‘universalism’. As is evident from the discussion in §89 ff. of the *Philosophical Investigations*, Wittgenstein’s main problem with logic-inspired approaches is that they
claim to unveil the essence of a phenomenon, i.e., give an all-encompassing and necessary characterisation of it. The point is that such investigations on the one hand constitute themselves on the model of the sciences, yet on the other hand strive for a philosophical, rather than an empirical answer. It is the combination of these two that leads them astray.

As for formal semantics, finally, the resulting view then becomes something like the following. What formal semantics delivers is a systematic account of broadly ‘referential’ aspects of meaning. As such that is an essential ingredient of an overall account, since in certain circumstances, as part of certain practices, these are the relevant features around which our use of language turns. And, as we read him, Wittgenstein does not exclude such accounts, provided we keep in mind their contextual and partial nature. From this perspective, then, formal semantics is one methodology that deals with one particular aspect of the heterogeneous phenomenon of meaning. Its contribution to our understanding consists of systematic, conceptual reconstructions of certain aspects of meaning at the idealised level of competence. By itself that is only marginally an empirical enterprise in the scientific sense of the word. But it does suggest further empirical investigations of actual performance, and as such is subject to indirect testing via such empirical investigations.

From this Wittgensteinian perspective, then, formal semantics is more like a specific type of ‘perspicuous representation’: a systematic laying out of observations concerning certain aspects of meaning, a description of particular ways of using expressions in certain language games in a specific vocabulary. It leads to a specific way of understanding those aspects, one that can be incorporated in a more encompassing one. And this also answers the question that forms the title of this paper. No, the Wittgensteinian perspective on meaning can not be regarded as an alternative to formal semantics, at least not in the sense of it constituting a rival theory that aims to replace the existing framework. In that sense, Wittgenstein’s work remains faithful to his claim that philosophy ‘leaves everything as it is’. But it does suggest an alternative self-image for formal semantics.

Martin Stokhof

ILLC/Department of Philosophy,
Universiteit van Amsterdam
FORMAL SEMANTICS AND WITTGENSTEIN

NOTES

1. Cf., Kamp and Stokhof (2008) for an overview of the various theoretical and descriptive changes that have taken place over the years.

2. Where by ‘Wittgensteinian’ we will mean what can be based on Wittgenstein’s major works of the later period, i.e., mainly, Philosophical Investigations (1953) and On Certainty (1969).

3. Cf., the extensive debate between contextualists and semantic minimalists. What is interesting to note is that, despite the many sophisticated arguments, the crucial question is not decided.

4. The competence-performance distinction has played a crucial role in establishing modern linguistics as a distinct discipline with a subject matter of its own. However, the concomitant difficulties that one encounters when one tries to explain the relevance of the results thus obtained in terms of testable hypotheses and predictions about actual, observable behaviour are severe. And note that linguistics is not on its own here: similar observations pertain to other humanities disciplines that, like linguistics, rely on certain idealisations regarding their subject matter that make their results difficult, if not impossible, to connect with empirical approaches.


6. Cf., e.g., Hauser, et al. (2002).


10. What follows is dealt with in more detail in Stokhof and van Lambalgen (2011a, b) (abstraction versus idealisation) and Stokhof (2011a) (methodology of intuitions).

11. A terminological remark: the distinction indicated goes by a variety of names in the literature. For example, Cartwright (Cartwright 1983) uses ‘idealisation’ for what we call abstraction, and ‘fiction’ for what we call idealisation; Jackendoff speaks of ‘soft’ versus ‘hard’ idealisations (Jackendoff 2002).


13. Note that since the standard theory considers the connections between intuitions, competence, and semantic facts to be conceptual connections, it does not help to rephrase the first step as: ‘competence is what a semantic theory is about’.


18. Despite the fact that many Wittgenstein scholars have convincingly argued against the attribution of a ‘meaning-is-use’ theory to Wittgenstein, this characterisation of his views can still be found in surveys, encyclopaedias, and the like.

19. Cf., e.g., the various contributions to Ammereller and Fischer (2004).

20. Cf., Stokhof (2011b) for more discussion.

21. The earliest example I could find is (Binkley 1973, 183).


25. Cf., Giddens’s conception of ‘double hermeneutics’, or Hacking’s ‘looping concepts’.
28. Under ‘referential’ we include all those aspects that are analysed in terms of a determinate relationship between expressions and extralinguistic reality, be it direct (extensional) or indirect (intensional), contextual or dynamic.
30. Part of this material was presented at the Department of Philosophy of the University of East Anglia in Norwich. I would like to thank the members of the audience for their very useful critical comments.

REFERENCES


