[WLG]
WIENER LINGUISTISCHE GAZETTE

A fresh look at compositionality

Paul Horwich

Sonderdruck aus: *Wiener Linguistische Gazette* (WLG) 82 (2017): 137–143

Themenheft *11-11-17. Festschrift für Martin Prinzhorn*
Hg. v. Clemens Mayr und Edwin Williams

Universität Wien · Institut für Sprachwissenschaft · 2017

Eigentümer, Herausgeber und Verleger:
Universität Wien, Institut für Sprachwissenschaft
Sensengase 3a
1090 Wien
Österreich

Redaktion: Christian Bendl, Mi-Cha Flubacher (Angewandte Sprachwissenschaft),
Stefan Schumacher (Allgemeine und Historische Sprachwissenschaft)
Kontakt: wlg@univie.ac.at
Homepage: *http://wlg.univie.ac.at*

ISSN: 2224-1876
NBN: BI,078,1063

Die *Wiener Linguistische Gazette* erscheint in loser Folge im Open-Access-Format.
Alle Ausgaben ab Nr. 72 (2005) sind online verfügbar.

A fresh look at compositionality[*]

Paul Horwich

New York University

1. Introduction

My discussion will have three parts. I'll begin with a thesis:– roughly, that a sentence's meaning what it does is simply its property of having a certain structure and having words with certain meanings. Then I'll mention some of the implications of that thesis. And finally I'll consider a few objections to it.

This last and defensive part will be the longest. For I know from bitter experience that the thesis will strike most readers as clearly wrong – even bizarre. I think that this is because it goes against a long tradition of theoretical work in semantics. It seems to me, however, that if one can loosen oneself a bit from the grip of all that tradition, the thesis can be seen to possess some quite attractive features. In the first place, it has considerable intuitive plausibility. In the second place, it is wonderfully simple. And in the third place, it enables us to slice through a number of thorny difficulties. So it's certainly worth an airing.

2. My thesis

So much for the advertising. What exactly is the thesis? Well it's a claim – an *a priori* claim – about *what it is* for a non-idiomatic complex expression of a language (e.g. a sentence-type of English) to mean what it does. It specifies how the *meaning*-properties of such expressions are *constituted*. The heart of it is just that the meaning-what-it-does of a non-idiomatic complex expression reduces to what I call the expression's "*construction*-property" – its property of *being the result of imposing such-and-such structure on words with such-and-such meanings*. (In what follows, "complexes" are restricted to "non-idiomatic complexes").

For example, according to this thesis, the *meaning*-property, 'x means MARS ROTATES' – which is possessed by the English "Mars rotates", the Italian "Marte gira", the German

---

[*] My profound thanks to Martin Prinzhorn for the suggestive conversations that we had in the mid-nineties, when I was beginning to think about the problems of compositionality. I'm also grateful to Jim Pryor and to Guillermo Del Pinal for their incisive comments on the penultimate draft of these remarks.

"Mars rotiert", etc. – is constituted by the *construction*-property, 'x results from applying a word meaning _ROTATES to a word meaning MARS'. (Capitalized expressions will be used to name the meanings of the corresponding lower case expressions).

To put it more generally. Suppose an expression, e, is the result of imposing structure P on words, ⟨w1, w2, . . . , wk⟩. Then e's meaning property is constituted by, 'x results from imposing P on words whose meanings are ⟨W1, W2, . . . , Wk⟩'.

OK. That's the claim in a nutshell. Of course it's highly cryptic as it stands, and the reader has every right to expect further clarification. But I'm hoping that what I will now go on to say about its implications, and then about various objections, will provide at least some of that clarification.

3. Import of the thesis

I'd like to emphasize two significant implications of this thesis.

First: it becomes very easy to see *why* compositionality holds (*when* it does) – to see *how* the meaning of a non-idiomatic complex is determined by the meanings of its elements and their arrangement. After all – on this view – the fact that a complex means what it does simply *is* the fact that its words have the meanings they do and are arranged as they are.

Thus, what we might call "Davidson's Problem" receives a trivial solution. He raised the question of how there *could* be (as there surely *must* be) logical deductions – hence explanations – of facts about the meanings of sentences on the basis of facts about the meanings of their words. And he suggested that the only way of achieving such deductions was to construe sentence-meanings as truth conditions and word-meanings as reference conditions. – For Tarski showed us (at least in the case of certain simple languages) how the former could be deduced from the latter.

But now we find that there is an *alternative* way of achieving these explanatory deductions – a much easier way. And this is potentially quite liberating, given the various difficulties that plague the Davidsonian approach and its Montegovian descendants. To begin with, those approaches make dubious assumptions about sentences-meanings – namely, that they are *truth conditions* or *sets of possible worlds*. (Dubious because, for example, "Water is falling from the sky" and "H_2O is falling from the sky", have different meaning but are true in the same possible worlds!). And in addition, *any* such initial assumption imposes a hard-to-satisfy constraint on a decent compositional account. – For many sentences (e.g. counterfactual conditionals, belief attributions, and epistemic modals), it can be extremely difficult to identify (i) a logico-semantic structure, (ii) semantic values for the words, and (iii) semantic combinatorial rules, that, taken together, will deliver the 'right' truth condition. In contrast, the approach I'm recommending involves no assumptions at all about the meaning-properties of non-idiomatic complex expressions, except that they are constituted by the expressions' construction-properties. So these problems simply don't arise.

A second important implication of my thesis is that compositionality *per se* imposes no constraint whatsoever on how *word*-meanings are constituted. One may be inclined, for one reason or another, to think that a word's meaning-property is engendered by its *use*, perhaps; or, alternatively, by an associated *stereotype*; or, alternatively, by a *recognitional capacity*. Or one might have another specific suggestion. No doubt some of these proposals are more plausible than others (– indeed, I myself strongly favoring a neo-Wittgensteinian use-theoretic view). But a consequence of my thesis is that *all* of them cohere perfectly with compositionality. So compositionality can't be a basis for deciding between them.

Suppose, for example, that the meaning of "Mars" is engendered by its having underlying property, U1(w), and that the meaning of "_rotates" is engendered by its having U2(w). Then – in light of my thesis – the meaning of "Mars rotates" will be engendered by its having the property, 'x results from applying a word with U2(w) to a word with U1(w)'. Thus, no matter what the U-properties are, compositionality will be accommodated.

But what about the famous series of papers by Fodor and Lepore – the papers in which they argue that none of the theories I have just listed is compatible with compositionality, and that none, therefore, can be correct? (See for example, Fodor & Lepore 2002). From my point of view their reasoning goes wrong in making a certain Uniformity Assumption. They presuppose that whichever kind of property provides a *word* with its meaning will also provide a *complex* with its meaning. They claim, for example, that if a *word*'s meaning derives from its *basic use*, then a *sentence*'s meaning would have to derive from its basic use too. But, as they go on to argue, the basic use of a sentence is not determined by the basic uses of the words in it. And similar reasoning is taken to disqualify many other candidates for what constitutes word-meanings.

However, my thesis implies that their Uniformity Assumption is false. If we advocate a use-theory of *word*-meanings, we should not advocate a use-theory of *sentence*-meanings. We should say, rather, that a sentence means what it does in virtue of its *construction* – that is, in virtue of its words having certain use-properties, and being combined as they are.

Returning to the above example, and supposing that the meaning constituting properties of words are *use*-properties: the idea is *not* that U2(w) is some sort of *function* or *operator* that applies to U1(w), yielding whatever constitutes the sentential meaning-property, 'x means MARS ROTATES'. So we have no reason to suspect that the meaning-property of any other predicate applied to "Mars" (e.g. the *complex* predicate, "_rotates slowly") must *also* be constituted by a use – a suspicion that would push us back towards a problematic Uniformity Assumption. No. The construction-property (= meaning-property) of "Mars rotates slowly" is its being the result of applying (A) to (B) – where (A) is a complex expression whose construction–property is the result of applying a word with U3(w) (in virtue of which it means SLOWLY) to a word with U2(w); and where (B) is a word with U1(w).

4. Objections to the thesis

My thesis about compositionality is *deflationary* in the following sense. The proposed explanation of compositionality is much *thinner* – much more superficial and un-theoretical – than what has standardly been thought to be needed. Consequently, most of the objections take the form: "It's *too* deflationary. – Here's an aspect of the phenomenon that cannot be accommodated!".

Well, perhaps that's right. But let's see. I'll only have space here to consider a small proportion of the many such objections that have been thrown at me. (Others are addressed in Horwich 1998, 2005).

> *Objection 1: Your claim conflicts with the explanatory significance of compositionality.– Intuitively, the compositionality of a language is a very significant fact about it – explaining (as Davidson emphasized) the quasi-infinite capacity that we are able achieve with respect to understanding such languages. But how can this explanatory significance possibly be squared with the sort of triviality attributed to compositionality by the deflationary thesis?*

But I am not denying the great causal/explanatory import of a given language being compositional. So I am not attributing triviality to the fact that English (say) is predominantly compositional. What is trivially easy, according to me, is to see *how* compositionality occurs when it occurs – to see how the meaning-properties of the sentences of a compositional language are determined by the meaning-properties of its words.

So there is no real tension here. On the contrary, it seems to me that a point in favor of my proposal is that the explanatory significance of compositionality becomes easier to understand. For it becomes easier to see how the meaning of a sentence issues in its overall use (including the circumstances in which it tends to be accepted, and what other sentences tend to be inferred from it). In other words, it becomes easier to see the role of compositionality in explaining the torrent of linguistic activity of which we are capable.

To help appreciate this point, let me suggest a certain analogy. The overall behavior of a *physical system* is the product of its having a certain construction property – i.e. a property specifying which laws govern the system's basic components and specifying how those components are arranged with respect to one another. Therefore, if we combine my main thesis (that a complex expression's construction-property is what grounds its having the meaning it has) with the further view that each word-meaning derives from its distinctive laws of use, it is unsurprising that we can explain the overall use of each sentence in terms of its construction-property – in term of its being the result of embedding, within a certain structure, words governed by a specific laws of use. So, we see how compositional meaning properties are causally significant, and therefore why it's explanatorily important that we are able to attach them to such a vast body of sentences, i.e. why the predominant compositionality of our language is important.

> *Objection 2: Your claim conflicts with holism:– The suggested model delivers a form of compositionality that is too crude, too extreme, too building-blockish. What we surely need is a more subtle, nuanced form – one that can incorporate a degree of holism, whereby word-meanings are abstracted out of the meanings of sentences, and are not individually identifiable.*

In the first place, this objector appears to be on the verge of *rejecting* compositionality, rather than wanting a subtle version of it. Do we understand sentences because we understand the words in them – or is it the other way around? I don't see how one can embrace the second of these options without denying compositionality. And then we'd have to revisit the question of what explains our understanding of an unlimited number of sentences!

Still, the imagined objector is perfectly right, it seems to me, in maintaining that meaning is "holistic", *in a different sense*. In particular, I believe we must grant that there are collections of fundamental word-meanings that are mutually inter-dependent in the following way: – it is not possible for a language to have a word with one of those meanings unless it also has words with all the others as well. Or, to put this inter-connectedness point more theoretically: – the law of use for any one of these words will make reference to its deployments in relation to other words with such-and-such laws of use.

But this form of holism is perfectly consistent with supposing (a) that each of the words has its own peculiar meaning-property, specifiably different from those of the others, and (b) that the meaning of each sentence is constructed, in the way I am proposing, from these individual meanings.

A second concession we can happily make to the present objection is that sentences are, in a certain respect, the *fundamental* units of a language. For it is plausible, I think, that the fact about a word that engenders its meaning (i.e. the word's 'law of use') is a fact that explicitly concerns certain *sentences* containing the word – a fact of the form "such-and-such sentences containing it are accepted in such-and-such circumstances". E.g. for "bachelor" it might be that "The bachelors are the unmarried men" is *unconditionally* accepted; and for "red" it might be (very roughly speaking) that "That's red" is accepted in the presence of red things.

It is a mistake, however, to think that this concession involves some retreat from the crude, building-block view of sentence-meanings. What we have allowed is merely that certain facts about the *usage* of certain sentences constitute the meanings of words. But the *meanings* of those sentences (amongst others) are explained by the word-meanings thus constituted.

> *Objection 3: Your claim conflicts with contemporary truth-theoretic semantics. – Presumably, we ought to take matter to be what the chemists and physicists tell us it is; for they have looked into the question much more carefully and competently than we have. Similarly, we should take meaning to be what the relevant scientists – the linguists who focus on semantics – tell us it is. And what they tell us – the framework for at least 99% of work in the field – is that the meaning of a sentence is its truth condition, that the meaning of a word is some referential characteristic, and that the compositionality of a sentence resides in the fact that its truth condition is logically deducible from the referential characteristics of its words in light of its structure. Thus, the deflationary proposal, in ignoring what science has to say, is on a par with an insistence that matter is continuous, that surely every bit of a bit of copper is a bit of copper, etc. – something only a crank could take seriously!*

This certainly is a formidable objection. For I agree that there is indeed a clash between mainstream formal semantics and my proposal. But it seems to me that mainstream formal semantics is not as successful and unimpeachable as its popularity might suggest. I venture this opinion with considerable trepidation and embarrassment. Who am I – who is any philosopher, or bunch of philosophers – to judge that some entire field of science is radically defective? However, in the end we mustn't allow ourselves to be intimidated simply by the weight of tradition. We just have to let the skeptical arguments speak for themselves. So here they are, in a nutshell.

In the first place - and as I've already mentioned - the meaning of a sentence appears to be more fine-grained than its truth condition: non-synonymous sentences can have the same truth condition; similarly, two words can have the same reference without having the same meaning. This is an old Fregean point. It involves a conception of "meaning" in which the meanings of terms (i) are supposed to help explain how those terms are deployed (e.g. that one might accept "Hesperus = Hesperus" without accepting "Hesperus = Phospherus"), and (ii) are supposed to be preserved in a good translation manual. Davidson worried about this divergence between meaning, in that rich sense, and reference/truth conditions; and he tried to fix up his theory in order to accommodate it. But no satisfactory way of doing that has ever been found.

Of course, this objection to truth-theoretic semantics does not entirely discredit the project of trying to find referents for various parts of speech such that the truth conditions of the sentences containing these parts of speech will be deducible. Such projects may well be perfectly coherent, legitimate, challenging, and fascinating. But – if the objection is correct – they do not yield a complete theory of *meaning*.

In the second place (and an implication of the preceding point) we want – or *should* want – a science of semantics that stands shoulder to shoulder with the other empirical sciences in providing causal explanations of observable linguistic phenomena. But mainstream truth-theoretic semantics doesn't do that. Its primary aspiration is to tell us what referents would have to be assigned to words, and what logical forms would have to be assigned to sentences, in order that the known truth conditions of sentences be logically derivable. And its next aspiration is to explain to us – in light of these results - why these sentences bear the logical relations to one another that they do bear. Relatively little attention is given to explaining linguistic *activity* – facts about our *acceptance* of sentences, about the *inferences* we draw amongst them, etc. And this shouldn't be surprising. For, it's hard to see how the derivations of truth-theoretic semantics could make any contribution to explanations of concrete empirical phenomena such as these (see Horwich 2010).

To cut a long story short, what instead needs to be looked at are those *use*-theoretic properties of words that underlie their possession of meanings and referents. These are the sorts of properties that, taken together and in conjunction with structural assumptions, stand a reasonable chance of accounting for the patterns of activity characteristic of different sentences. And, from this point of view, it is quite natural to suppose that a *sentence's* meaning what it does is simply the fact that it has a certain structure and has words whose basic use properties are this, that, and the other.

Finally – my third response to the objection – it's worth remembering how semantics got pushed into the truth-theoretic direction in the first place. An important factor was Davidson's hugely influential argument – in "Truth and Meaning" – to the effect that this was the only way to make sense of compositionality. My speculation is that if the deflationary alternative had been appreciated way back in the 70's, then natural language semantics would perhaps not have got off on the wrong foot, as I think it did, and might now be in a much healthier state than I think it is actually in.

References

Fodor, Jerry, & Ernie Lepore. 2002. Why compositionality won't go away: Reflections on Horwich's 'deflationary' theory. In *Meaning and representations*, ed. Emma Borg. Oxford: Blackwell.

Horwich, Paul. 1998. *Meaning*. Oxford: Oxford University Press.

Horwich, Paul. 2005. *Reflections on meaning*. Oxford: Oxford University Press.

Horwich, Paul. 2010. *Truth – meaning – reality*. Oxford: Oxford University Press.

Paul Horwich
ph42@nyu.edu