Structure and Conventions

Zoltán Gendler Szabó
Cornell University

The slogan "words mean what people use them to mean" has a democratic aura. Typically it is used to fend off authoritarian attempts at linguistic legislation. But it also has a secondary use in curbing the anarchist excesses of the linguistic non-conformist. It announces that in matters of meaning, it is we the people who are in charge. So politically the slogan has strong credentials. But can you build a theory of meaning on it?

Part II Wayne Davis's *Meaning, Expression and Thought* attempts to do just that. His first pass at a definition of linguistic meaning is a straightforward generalization of the democratic slogan:[1]

(1) e means μ iff people use e to mean μ (191).

Davis devotes over a hundred pages to successive refinements to this account, which he dubs the *neo-Gricean analysis* because it preserves the Grice's insight that linguistic meaning is defined by speaker meaning, while abandoning Grice's commitment that the latter is invariably propositional.

I think the neo-Gricean analysis represents one of the best attempts to articulate a credible use theory of meaning, and as such, it deserves serious attention from philosophers of language.[2] While there are important aspects of the analysis with which I disagree, I endorse many of its insights. I begin by discussing some of Davis's reasons for being a heterodox Gricean. In section II, I rehearse what I take to be the two fundamental difficulties with (1) and explain how Davis proposes to sidestep them. The key to his proposed solutions involves expanding our vague commitment to the conventionality of linguistic meaning beyond its ordinary scope. In the final part of this paper I raise an objection against this maneuver.

I

When it comes to the neo-Gricean analysis, the definiendum probably raises more eyebrows than the definiens. When one fills in the suppressed universal quantifiers in (1) it is clear that the schema not only ascribes meanings to linguistic expressions, but also quantifies over them. This runs counter to recent orthodoxy. Philosophy of language in

[1] Linguistic meaning is relative to languages. I will suppress reference to language relativity throughout my discussion whenever that leads to no confusion.

[2] As Davis notes (167), the analysis of linguistic meaning in terms of speaker meaning is often viewed as a rival of any analysis in terms of use. This is largely because the use theory of meaning has been a longtime ally of behaviorism. Wittgenstein and Austin presupposed that the use of an expression can be identified without appeal to mental states. But the presupposition is suspect and can be dispensed with. Grice's own theory can be viewed as an attempt to identify a particular illocutionary act – *saying* – that is simultaneously definable in terms of speaker meaning and defines linguistic meaning.

the twentieth century has, for the most part, eschewed talk of meanings, and to reify them was supposed to be a very bad thing. Fifty years ago, during the heyday of behaviorism, nominalism, and verificationism, this was understandable. But nowadays, when appeals to mental states, abstract entities, and unrestricted bivalence pass without remark, there are few arguments left that could justify the practice of keeping meanings out of the theory of meaning. I know of two, and like Davis, I find neither convincing.[3]

The first argument, due to Davidson, is that meanings have no job to do in semantics. Davidson proposed that a theory of meaning should be a theory of knowledge of meaning, that knowledge of meaning consists of knowledge of an interpretative theory of truth, and that an interpretative theory of truth does not assign meanings to anything whatsoever.

I think Davidson is right that systematic assignment of truth-conditions should be part of an adequate theory of meaning. But that shouldn't be all that a theory of meaning includes. There are meaningful expressions with null truth-conditional content – 'by the way', 'moreover', 'indeed', or 'needless to say' are fairly uncontroversial examples. And there seem to be pairs of linguistic expressions differing in meaning but not in truth-conditional content – such as 'and' and 'but', or 'woodchuck' and 'groundhog'. In addition, as Davis emphasizes, truth-conditional theories leave open the fundamental question about *why* sentences have the truth-conditions they do (171).[4] Any answer to this question is likely to appeal either directly to linguistic meanings, or to facts that could provide the ground for a reductive analysis of linguistic meanings. An appeal to truth-conditions does not alleviate the need for meanings in semantics.

There is a second reason for not assigning meanings to linguistic expressions, although this reason does not generalize fully. As Dummett has frequently emphasized, we lack non-contrived ways to say what words mean. Although most of us know that no means no, we are hesitant about the correct typography: there should certainly be quotation marks around the first occurrence of 'no' but it is unclear what to do with the second. These days it is customary to use capital letters to signal that the expression on the right is used to refer to its own meaning:[5]

(2) 'no' means NO
(3) 'dog' means DOG

This looks acceptable, although the analogous clauses for prepositions or bound morphemes are somewhat outlandish:

[3] I neglect concerns about synonymy. Most of these are specific instances of completely general worries about vagueness that have not much to do with meaning *per se.*

[4] Standard truth-conditional semantics delivers a partial answer. 'Dogs bark' has the truth-conditions it does because 'dogs' is true of dogs and 'bark' is true of things that bark and because the sentence is composed from these constituents in a certain fashion. But when the question is raised why 'dogs' is true of dogs and why 'bark' is true of things that bark, truth-conditional semantics remains silent. While explanations must undoubtedly end somewhere, it does not seem promising to declare that here we have already hit rock bottom.

[5] Davis uses double quotes, so he would have ' "no" means "no" ' and ' "dog" means "dog" ' instead. It is important that the double quotes are interpreted differently in their two occurrences. The claim should certainly not misconstrued as saying that 'dog' is synonymous with itself (146).

(4) ‘of’ means OF
(5) ‘-s’ means -S

All this is in a stark contrast to the case of declarative sentences, where the use of the appropriate complementizer makes the meaning-ascriptions sound perfectly acceptable:

(6) ‘The sky is blue’ means that the sky is blue.

It is unclear whether such linguistic facts should be thought to have deep significance.[6] We do, after all, think that ‘no’, ‘dog’, ‘of’ and ‘-s’ each *have* meanings, just as we think ‘The sky is blue’ has one. What could be wrong with assigning to an expression a meaning it has?[7] There might, of course, be all sorts of legitimate reasons for wanting to eschew meanings altogether – their existence is, after all, incompatible with many austere ‘ism’s. But it is odd to think that some meanings (i.e. those of sentences) are less objectionable than others (i.e. those of prepositions and bound morphemes) simply because we have a colloquial way to designate them.

Even if we grant (as we should) that an adequate theory of meaning for a language assigns meanings to all the meaningful expressions of that language, it remains an open question whether sentences have a distinguished semantic role to play. There are excellent reasons for thinking that they must. Most of the data of semantic theorizing come at the sentence level. Most of the force indicators attach to sentences only. There is a powerful intuition that the unit of communication is the sentence, and another that one should not be credited with full understanding of a word unless one understands a reasonably wide range of sentences in which that word occurs. Such facts have been used by many to justify the *primacy of sentence meaning*, the thesis that the meanings of sentences are basic and the meanings of words somehow derivative.

Although the primacy thesis is announced at many places,[8] what exactly it amounts to remains a bit obscure. Since it is supposed to assert the primacy of something over something else it should be interpreted as a claim that is incompatible with its own converse. So, to believe in the primacy of sentence meaning it is not enough to accept that the meanings of words are *determined by* the meanings of sentences, for that does not rule out that the meanings of sentences are in turn determined by the meanings of words. For similar reasons, it is also not enough to say, as Davis sometimes does, that the meanings of words *can be defined in terms of* (176) or *derived from* (175) the meanings

[6] For a forceful recent statement that they do, see Schiffer (2003): 100-5.
[7] Dummett (1974) says such considerations are superficial (7). He mentions that although a successful investigation should tell us the identity of the murderer, we do not expect it to deliver a true sentences of the form ‘The identity of the murderer is …’. This is a nice example, but the analogy between identity-talk and meaning-talk breaks down when it comes to quantification. ‘Some words have more than one meaning’ is true if interpreted in a straightforward fashion. ‘Some murderers have more than one identity’ is also true, but it cannot be read as saying that some murderers are identical to more than one murderer. My claim is that ‘meaning’ – unlike ‘identity’ in the sense in which it is true that a successful murder investigation must tell us the identity of the murderer – is an ordinary count noun. If we accept this much, and if we go along with the truth of sentences like ‘Some words have more than one meaning’ we should accept that ‘meaning’ has a non-empty extension.
[8] See Davis (2003): 175, fn. 16 for a representative list.

of sentences in which they occur.[9] The formulations that guarantee asymmetry are "Words *get their meanings* entirely from the fact that they are parts of meaningful sentences" (182) and perhaps somewhat less contentiously "Words have the meanings they do *because* sentences in which they occur mean what they do."[10] I will assume that this last formulation is adequate.

Davis points out that none of the standard considerations for the special status of sentence meaning proves the primacy of sentence meaning. In my view, these arguments are thoroughly convincing. When it comes to his arguments *against* sentential primacy, however, the results are more ambiguous. Davis's central point is that sentential primacy is incompatible with compositionality (181). He thinks this, presumably, because he assumes that compositionality is itself a primacy claim, something along the lines of "Sentences have the meanings they do *because* they are built from words with certain meanings." I submit that this is not a consequence of any standard reading of the principle of compositionality. If compositionality were an asymmetric principle, there couldn't be a serious debate whether *in addition* to the principle of compositionality we should also adopt the principle of reverse compositionality.[11] But there is such a debate.[12]

In my view, Davis's best arguments against sentential primacy involve toy languages where there are fewer sentences than words, and so the meanings of sentences cannot fix the meanings of words (187). Such examples show that sentential primacy is false in some conceivable languages. Whether the thesis holds in richer languages that more fully resemble the languages we actually speak strikes me as an open question.

To sum up, I think Davis makes the right theoretical decision in turning against a long tradition that tries to do away with meanings. Sentential meanings may have a special role to play is semantics, but that is feeble justification for thinking that words are meaningful only because they occur in meaningful sentences, and no justification for assigning meanings to sentences only. Davis is asking the right question: What is it that asymmetrically determines the meaning of an arbitrary linguistic expression? Let's see whether his answer works.

II

The neo-Gricean analysis is a definitional equivalence. As such, it claims the primacy of speaker meaning over linguistic meaning: (1) tells us, for example, that 'rabbit' means what it does *because* people use 'rabbit' to mean just that. This is the source of the first

[9] If we adopt the Aristotelian conception of a *real definition*, we may say that if A can be defined in terms of B then B cannot be defined in terms of A. But I don't think it would be easy to defend the idea of real definitions when it comes to meanings.

[10] There is an extreme view in this neighborhood associated with Frege's dictum that only in the context of a sentence does a word have a meaning. Davis calls this the *sentential confinement* thesis (176) and points out that it seems plainly false, at least when it concerns *Sinn* rather than *Bedeutung*. (If words are meaningless in isolation but some sentences are meaningful, it must be that when we put together a bunch of meaningless things in the right way each of them somehow becomes meaningful. I don't think this is incoherent, but it sure sounds like something straight out of a Borges story!)

[11] E.g. Fodor and Lepore (2002) and Pagin (2003).

[12] Szabó (2004a).

important objection against the analysis. For why it is that people use 'rabbit' to mean what they do? Embarrassingly, the natural response seems to be this: people use 'rabbit' to mean what they do *because* 'rabbit' means just that. (Had 'rabbit' meant something different, people would surely not use it to mean what they do.) If we combine our two explanations (and assume that explanation is transitive) we get that 'rabbit' means what it does because it means what it does – an unacceptable result. Call this the *circularity objection* against the neo-Gricean analysis.[13]

Davis responds by appealing to *conventions*. He calls the following the *basic neo-Gricean analysis*:

(7) e means μ iff it is conventional for people to use e to mean μ (192).

If we combine Davis's own expression theory about speaker meaning (developed in Part I of *Meaning, Expression and Thought*) with the basic neo-Gricean analysis we get the following equivalence:

(8) e means μ_i iff it is conventional for people to use e to directly express i (193),

where i is the idea expressed by μ_i. It sounds plausible that 'rabbit' means what it does because it is conventional for people to use it to directly express the idea of a rabbit, while the converse is surely false. Conventions ground linguistic meanings not the other way around; so the circularity objection has been circumvented.

Western philosophical tradition has been deeply ambivalent about the conventionality of linguistic meaning. Consider the two most influential ancient accounts on the origin of language. There is the Biblical story that tells how Adam gave names to the birds and the beasts when he first encountered them (Gen. 2.19). Both Hebrew and Christian commentary on the passage has contended that the original naming was guided by divine suggestion, and that the original names were in some way singularly adequate to their bearers. Then there is Lucretius's account from *De rerum natura* (5.1028-90). It says that language emerged from our innate ability to indicate objects by gestures and emotions by cries. The two accounts differ markedly in most crucial regards: whether language is ultimately divine or a human invention, whether it originates in the intentions of a single name-giver or in social interactions, and whether its history is that of decline or ascent. But they agree in one thing: although language started as a system of natural signs, somehow or other it ended up as a system of conventional signification. How this happened is a mystery. According to the Bible it was God's punishment for trying to build the tower of Babel – Lucretius remains conveniently silent on the question.

The conventionality of linguistic meaning is puzzling because we do not learn our native tongue in the way we learn about social conventions. Learning our first language requires little or no explicit instruction, it happens at a certain developmental stage, and tends to yield a remarkably uniform level of competence. None of this is true when it comes to learning about traffic signs, dress codes, or proper table manners. Learning a language comes naturally to us in a way that learning conventions tend not to. This strongly suggests that language is not entirely arbitrary, and hence, not entirely

[13] Davis calls it the *explanation problem* (195).

conventional.[14] Of course, it is beyond dispute that our linguistic expressions *could* have different meanings, and that most of them actually *did.* But that is compatible with the possibility that there is a robust non-conventional component to linguistic meaning. Such an assumption may help us account for the remarkable facts of first language acquisition.

This leads to the second important objection to the neo-Gricean analysis. This one is even more straightforward than the first. Meaning isn't any kind of use (including conventional use to express ideas), for most meaningful linguistic expressions have never been used. The simplest way to show this is to point out that the number of linguistic expressions that have ever been used is finite while the number of sentences in English is not.[15] Call this the *productivity objection*.

As a quick fix, Davis introduces a new clause in his definition and calls the result the *recursive neo-Gricean analysis*:

(9) e means μ_i iff (i) it is conventional for people to use e to directly express i, or (ii) people conventionally use $E[x_1, x_2, \ldots, x_n]$ to directly express $I[i(x_1), i(x_2), \ldots, i(x_n)]$, where $e = E[x_1, x_2, \ldots, x_n]$ and $i = I[i(x_1), i(x_2), \ldots, i(x_n)]$, and $i(e_i)$ is the idea expressed by e_i for all i from 1 to n (235).

The recursive clause addresses the productivity problem by associating conventions of use not only with linguistic expressions but also with the modes in which such expressions can combine into larger expressions. Consider Davis's example, the complex noun phrase 'aardvark lover.' The phrase is hardly ever used, so there is no convention for its use. It means what it does because in addition to the *lexical conventions* fixing the meaning of 'aardvark' and 'lover' there are also *constructive conventions* fixing the meaning of the schema *N V-er*. The constructive convention is that an *N V-er* is something that *V*s *N*s (and, say, not something that is an *N* that *V*s). Plugging in the meanings of 'aardvark' and 'lover', the meaning of the complex expression is thus determined.

Davis goes to some length to convince his readers that the appeal to constructive conventions is a common sense move. He writes: "given that the cognitive capacity to recognize not only particular expressions and ideas, but also their structures, the notion that there are conventions relating the latter is no more far-fetched than the platitude that there are conventions relating the former" (241). But I don't think the introduction of constructive conventions is a trivial matter. To say that the syntactic structure is a conventional component of its meaning is very much a substantive claim.

III

The principle of compositionality, as it is usually stated, says that the meaning of a complex expression is determined by its syntactic structure and the meanings of its

[14] Davis views conventions in a broadly Humean fashion as socially useful, self-perpetuating and arbitrary regularities (206).

[15] For example, any string of words consisting of a finite number of occurrences of 'buffalo' is a meaningful English sentence.

syntactic constituents. Given universally accepted assumptions about syntactic constituency, this entails that ambiguity in a complex expression has only two possible sources: the expression may have more than one syntactic structure or it may contain an ambiguous lexical item as a syntactic constituent. The existence of Davis's constructive conventions is incompatible with compositionality thus understood. If there are such conventions, then the contribution of syntactic structure to meaning is conventional, and hence arbitrary. This means that the very same syntactic structure could contribute in more than one way to the meanings of complex expressions – in other words, that there is a possible source of ambiguity beyond syntactic structure and lexical meaning.

Consider again Davis's example 'aardvark lover'. Besides its most natural interpretation in which it applies to things that love aardvarks, this complex noun phrase can also be interpreted to apply to lovers who are aardvarks. (As in 'The aardvark lover is particularly fierce after consuming a large number of termites.') For Davis, the obvious way to accommodate this fact would be to say that the schema *N V-er* is ambiguous: there are multiple constructive conventions for its use. But this would mean that the meaning of 'aardvark lover' is not determined by the syntactic structure of the phrase and the meanings of its syntactic constituents, and consequently, that English is not compositional.

Is there a credible compositional alternative? Definitely. The difference in meaning between the two interpretations of 'aardvark lover' is traceable to the thematic roles of 'love'. All we need to save compositionality is a syntactic structure where thematic roles are represented. To see how this could be done, consider a very similar case discussed by Richard Larson.[16] The sentence 'Olga is a beautiful dancer' is ambiguous between the readings where Olga is said to be beautiful and where her dancing is. Larson proposes the following logical forms:

(10) $\exists e$ (dancing (e) $\wedge$ Agent (e, Olga) $\wedge$ beautiful (Olga))
(11) $\exists e$ (dancing (e) $\wedge$ Agent (e, Olga) $\wedge$ beautiful (e))

Along the same lines, we can capture the ambiguity of 'Ivan is an aardvark lover' structurally as follows:

(12) $\exists e \exists x$ (loving (e) $\wedge$ Agent (e, Ivan) $\wedge$ Theme (e, x) $\wedge$ aardvark (Ivan))
(13) $\exists e \exists x$ (loving (e) $\wedge$ Agent (e, Ivan) $\wedge$ Theme (e, x) $\wedge$ aardvark (x))

This is not the place to review the question how these logical forms are derived, or whether they have sufficient syntactic plausibility.[17] My point is merely illustrative: we don't have to give up on compositionality because of the ambiguity of 'aardvark lover'.

I should emphasize that nothing in the neo-Gricean analysis speaks against adopting thematically enhanced syntactic structure. But if Davis wants to save compositionality *tout court* he must accept that syntactic structure contributes to meaning in uniform way. And this would be nothing short of a miraculous coincidence, if the

[16] Larson (1998).
[17] For some background and motivation for event-based semantics with thematic roles see Parsons (1990) and Higginbotham (2000).

contribution of syntactic structure to the meaning of a complex expression is a matter of arbitrary convention.

It is customary these days to regard compositionality as non-negotiable. I think this is a mistake. What is non-negotiable is the productivity of natural languages – it is simply a fact that competent speakers of Arabic or Zulu understand an unlimited number of complex expressions they have never encountered. Productivity must be explained somehow, and a theory that leaves no room for adequate explanation cannot be correct. But compositionality is not necessary for explaining productivity. What is needed is some sort of recursive procedure that generates meanings for complex expressions on the basis of the limited information competent speakers plausibly possess. The idea that the relevant information is exhausted by syntactic structure and lexical meanings is a bold empirical hypothesis.

Davis does have an explanation for productivity. He thinks there are conventions that associate idea-structures with syntactic structures just as there are conventions that associate ideas with lexical items. Productivity is explained by the assumption that people who understand a complex expression know how to make these associations. This is a fairly perspicuous explanation.[18] Still, it may well not be the *best* explanation. A better explanation might dispense with the assumption that in order to understand a complex expression people must know how to associate something or other with syntactic structure. If the principle of compositionality is true, it is enough if they know what the syntactic structure *is*. This explanation eliminates a source of ambiguity and simplifies the cognitive task of interpretation. Of course, it would be even better if we could eliminate all sources of ambiguity. But as long as we are dealing with a natural language, structural and lexical ambiguities are facts of life. Compositionality seems to be the strongest principle that has a chance of being true. If it really is, it provides the best explanation for productivity.[19]

Compositionality requires a modicum of anti-conventionalism about linguistic meaning. It may be a matter of convention what sort of syntactic structures are available in a language and what words and morphemes mean. But if the language is compositional, beyond these two components there is nothing arbitrary, and hence, nothing conventional about meaning in that language. I don't know whether natural languages are compositional. But I think it is sound methodology to assume that they are, until strong empirical evidence forces us to retreat to some weaker hypothesis. Until then, I remain skeptical about the existence of constructive conventions, and consequently, about the truth of the recursive neo-Gricean analysis as well.

---

[18] Davis says that he believes that compositional semantics provides the only explanation for productivity (239). I suspect that by compositional semantics he means nothing more than recursive semantics.

[19] There are plenty of reasons for doubting whether natural languages are compositional. (For some arguments pro and con, see Szabó (2004b).) One reason for doubting whether syntactic structure contributes to the meaning of complex expressions in a uniform manner in English is the fact that many (perhaps all) declarative sentences can be used with proper intonation to express questions. *Prima facie*, the sentence 'Hugo likes broccoli' pronounced in a particular way may mean what 'Does Hugo like broccoli?' does, although pronounced more naturally it means something else. Now, I happen to think that there is a good case to be made for intonation making a difference to syntactic structure. But if I am wrong about that we have a widespread violation of compositionality at our hands.

References

Davis, W. (2003) *Meaning, Expression, and Thought.* Cambridge: Cambridge University Press.

Dummett, M. (1974) 'What is a theory of meaning I.' In: *The Seas of Language.* Oxford: Clarendon Press, 1993, pp. 1 – 33.

Fodor, J. and E. Lepore (2002) *The Compositionality Papers.* Oxford: Blackwell.

Higginbotham, James (2000) 'On Events in Linguistic Semantics', in J. Higginbotham, F. Pianesi and A. Varzi eds. *Speaking of Events*, 49 – 80, Oxford: Oxford University Press.

Larson, R. (1998) 'Events and modification is nominals.' In D. Strolovitch and A. Lawson (eds.) *Proceedings from Semantics and Linguistic Theory (SALT) VIII.* Cornell University, Ithaca, NY.

Pagin, P. (2003) 'Communication and strong compositionality.' *Journal of Philosophical Logic* **32**: 287 – 322.

Parsons, Terence (1990) *Events in the Semantics of English.* Cambridge, MA: MIT Press.

Schiffer, S. (2003) *The Things We Mean.* Oxford: Clarendon Press.

Szabó, Z. G. (2004a) 'Review of J. Fodor and E. Lepore, *The Compositionality Papers.*' *Mind* **113**: 340 – 4.

Szabó, Z. G. (2004b) 'Compositionality.' *Stanford Encyclopedia of Philosophy.*