

BOOK REVIEW

Paul Horwich (ed.): *Meaning*, Clarendon Press, Oxford, 1998, ix + 241 pp.

Berkeley once compared philosophical problems to the situation when “we have first raised a dust, and then complain, we cannot see”; and this quote could be the motto underwriting Horwich’s attempt to dismantle the enigma of meaning in the way he attacked that of truth earlier (*viz.*, Horwich 1990). Horwich’s strategy is to show that meaning is in fact a simple and perspicuous concept and that it appears perplexing for philosophers only because they have piled it with unreasonable onuses. If the task of a theory of meaning is, to borrow David Lewis’ (1972, 173) often quoted dictum, “to find out what meaning does and then to find something that does this”, then Horwich’s suggestion is that once we discard distorted views of what meaning does, the task of finding something that does that becomes rather easy. In particular, Horwich claims we will then safely be able to say, together with Wittgenstein, that *meaning is use*.¹

Horwich claims that each meaningful expression has a property, *meaning property*, in virtue of which it means what it does. He introduces a ‘disquotational-capitalizing’ notation: he designates the meaning property currently possessed by the English word “dog” by the predicate “meaning DOG”, that possessed by “and” by “meaning AND” etc.; and he says that the meaning property of an expression consists in the way the expression is used. In particular, the meaning property of any word is constituted by the fact that certain sentences containing the word are being accepted. Thus, the word “and” means what it does in virtue of the fact that we tend to accept “*p* and *q*” if and only if we accept “*p* and *q*”; whereas the meaning of “red” consists in our having “the disposition to apply ‘red’ to an observed surface when and only when it is clearly red” (p. 45; let us note in passing that if meaning is to consist in an acceptance of *sentences*, then “to apply ‘red’ to an observed surface” has to be construed as something like “to accept ‘(This is) red’ when observing a surface”). The pervading idea is that the entire usage of a word is reducible to a basic regularity, which is thus constitutive of the meaning of the word.



The core of the book is devoted to showing that this simple theory fulfills all reasonable demands which can be made on a theory of meaning and defending this conclusion against all possible kinds of objections. Horwich counts with seven basic constraints which, according to him, a theory of meaning is generally expected; and he tries to show that if they are not misinterpreted, then the use theory of meaning fulfills them almost trivially. The discussion of the seven demands constitutes Chapter 2 of the book; some of them, however, are discussed in greater detail in subsequent chapters.

The constraints, in Horwich's view, are the following (p. 13):

1. *The Understanding Constraint.* The theory must explain how facts about meaning can be known; for understanding a language involves knowing what its words mean.
2. *The Relationality Constraint.* The theory must specify the general nature of the relation between terms and their meanings; that is, it must provide some analysis of the notion, 'x means y'.
3. *The Representation Constraint.* The theory must explain how language can represent reality – how, in virtue of their meanings, singular terms are used to refer to objects in the world, predicates pick out sets of things, and sentences express thoughts that are true or false.
4. *The Aprioricity Constraint.* The theory must accommodate the arbitrary, conventional, aprioristic character of the meanings of words. For the choice of which language to use to express one's empirical beliefs is not itself an empirical matter.
5. *The Compositionality Constraint.* It must explain how meanings are compositional – how it happens that the meanings of complex expressions are determined by the meanings of their parts.
6. *The Normativity Constraint.* It must explain how meanings can have normative consequences – how it is, for example, that given what the word "dog" means, one *ought* to apply it to this thing but not to that.
7. *The Use Constraint.* It must account for the relationship between the meaning of an expression and its use - the possibility of explaining the way words are used on the basis of how they are understood.

Horwich claims that if these constraints are understood properly, none of them poses a serious problem for a simple use theory of meaning. The problem with Constraint (1), according to him, is that it is often construed as amounting to *explicit* knowledge of meaning, whereas it should properly be understood as speaking merely about an *implicit* one. Thus, understanding is "a skill or a practical ability with no propositional content" (p. 18) – to understand an expression is to be able to use the expression in question in an appropriate way, not to have an explicit knowledge of an object.

In the case of (2), Horwich claims that although it is possible to construe “‘dog’ means DOG” as “the utterance of ‘dog’ indicates (i.e., justifies belief in) the presence (within some mental state of the speaker) of the concept, DOG”, this possibility concerns only the “superficial composition of meaning properties”, not their underlying nature, which is non-relational. Thus, we can say, meaning-talk is relational as such, whereas it becomes non-relational when translated into the language of physics.

Regarding (3), Horwich states that this constraint is again not controversial in itself; that it becomes problematic only if construed in an unwarrantedly strong way, namely as saying “that the meaning properties of terms must reduce to *relations* between those terms and the aspects of reality to which they apply” (p. 27). That “dog” refers to dogs or is true of dogs is a platitude explainable in a disquotational theory; it does not mean that there must be a real (e.g., causal) link between the word and something within the world.

Likewise, (4) presents Horwich with no difficulty – unless it is construed to imply that what thus comes to be known a priori are “substantive postulates”. Horwich claims that “the commitment to a substantive theory of *f*-ness, ‘#*f*’, is the product of two independent decisions: one of them is to hold true the existential thesis ‘ $\exists x(\#x)$ ’ (...), the other is to hold true the conditional ‘ $\exists x(\#x) \rightarrow \#f$ ’” (pp. 31–2). It is the latter conditional which constitutes the real implicit meaning of “*f*” – and this conditional is not, according to Horwich, “substantive”.

To satisfy (5), Horwich claims, it is enough “to find accounts of word meanings consistent with the fact that the meanings of words engender the meanings of sentences” – and to use conception of meaning, according to him, is precisely suited for this. Also, (6) is readily satisfied by the use theory – once we realize that “the normative import of meaning does not preclude a reduction of meaning properties in non-semantic, non-normative (...) terms”, for it is clear that “situations characterized in non-normative terms” may have “normative import” (pp. 38–9). And, (7) is obviously satisfied trivially.

In the following chapter, Horwich sketches the basic features of his use theory. His crucial theses are the following: (i) Meanings are concepts, “abstract entities from which beliefs, desires and other states of mind are composed”; (ii) the overall use of each word stems from its possessing a basic “acceptance property”; and (iii) two words express the same concept in virtue of having the same basic acceptance property (pp. 44–6). Hence, in Horwich’s view, according to Horwich, the meaning of any word is engendered, as noted above, by the fact that the competent users of the word accept certain sentences which contain the word. The remainder of

the chapter is devoted to countering any possible kinds of criticisms of this conception of meaning.

The ensuing five chapters elaborate on some of the points sketched in Chapter 2. In particular, Chapters 4 and 5 summarize Horwich's disquotational theory of truth and show how it engenders a deflationary theory of reference. Chapter 6 gives more consideration to the problem of implicit definitions and a priori knowledge, while Chapter 7 focuses on the problem of compositionality and Chapter 8 on that of normativity of meaning. The final two chapters of the book, Chapters 9 and 10, then present a discussion of controversial theses of Quine and Kripke, respectively.

It would seem that there are two basic ways to write a book expounding a question as general as that, about the nature of meaning. One is a 'scientific' way, which consists in discussing all the answers proposed by earlier inquirers, weighting their respective pros and cons, and then working towards an answer accommodating those aspects of the previous views which are taken to be warranted and correcting those which are not. The other way is chosen, e.g., by Wittgenstein for his *Tractatus*: consisting in paying less attention to what has been published about the theme so far, and more to the very exposition in retelling the whole story in one's own words. Horwich chooses the latter, 'Tractarian' way, and thus his approach resembles not only the late Wittgenstein in holding the conviction that solutions of philosophical problems are "open to view" when delusive prejudices and preconceptions are removed, but also the early Wittgenstein in trying to present a self-contained, 'crystal-clear' exposition.

I think this choice is understandable; and it has enabled Horwich to write an extremely readable and duly provocative book. His edifice is self-contained, perspicuous and 'seamless' to such an extent that it is difficult for an opponent to find a fissure to wedge it. (Which is not to predict that proponents of more 'substantial' theories of meaning will feel overwhelmed – I suspect that most of them will take the 'seamlessness' of Horwich's exposition to be engendered by the fact that the author did not appreciate 'the real depth' of the problem.)

As far as I can see, the most heterogeneous brick within the walls of Horwich's edifice is his argument concerning the constraint (2). There, his otherwise anti-psychologistic semantic theory adopts, what to me seems to be a discordantly psychological appendage. "Occurrences of the word 'dog'", says Horwich (p. 4), "provide reason to believe in the presence (within the thought of the speaker) of the concept DOG". However, what if the word "dog" occurs in a newspaper article – in whose thought should we believe the concept to be present then? The author of the article in question? But what if the article is a joke put together by a computer

program? And even if we dismiss doubts of this kind, it would seem that Horwich's proposal rests on an assumption which, I think, surely should not go without saying: that *to have meaning* is not only *to be used in a certain way*, but also *to be uniformly accompanied, when uttered, by a specific mental state or activity, 'presence of a concept within the mind'*. As far as I can see, this assumption is not only problematic (cf. Dummett 1988, 185), but also incongruous with the rest of Horwich's story. (After all, the author claims that understanding is a *skill*, not a possession of something.) But the truth is that, if we do not buy this mentalistic account, we will have troubles with understanding *why* there should exist the two levels of the meaning talk which play such an important role in Horwich's account: the 'surface' level on which meanings appear to be things associated with expressions, and the underlying level on which there are only expressions employed in certain ways.

Besides this, Horwich sometimes seems too hasty in dismissing opposing ways to view meanings. Thus, on p. 16, he writes: "It is not easy to identify the fact regarding, say, the word 'dog', the knowledge of which constitutes our understanding of that word. The obvious candidate is the fact that it means what it says, i.e., the fact that 'dog' means DOG. But there appears to be a decisive objection to this suggestion – as clearly correct as it may initially seem. For in light of how the meaning designator, 'DOG', has been introduced (...), it is a trivial matter to work out that 'dog' means DOG. ... So, it cannot be that understanding the word 'dog' consists in knowing that dog means DOG. But this is puzzling; for what other fact could possibly be relevant?" Now, suppose that somebody is of the opinion that the meaning of "dog" is, say, a kind of a mental content and that understanding the word, consists in possessing the very content. The reply of such a person to Horwich's rhetorical question would then probably be: "The fact that the word is associated with the relevant content, of course! – When you say that 'dog' means DOG, then your 'DOG' is bound to refer to the very content!" The point is that Horwich's argumentation appears to be on a par with the following: "For every country C , let ' CP_C ' denote the capital of C . Now the knowledge that the capital of C is a certain town (namely CP_C), cannot be substantial, to work out that the capital of C is CP_C (i.e., that 'the capital of C is CP_C ' is true) is a trivial matter".

The 'Tractarian' form of Horwich's exposition also engenders a certain lack of clarity in respect to the relationship of his views to those of other similarly-minded philosophers. Horwich pays some explicit attention to the relationship between his doctrine and that of Donald Davidson; however, he appears to almost overlook another prominent approach which

would seem even closer than Davidson's to his own stance, namely the approach introduced by Wilfrid Sellars (see, e.g., Sellars, 1963) and recently further developed by Robert Brandom (1994). Sellars' basic claim bears much similarity to Horwich's, namely that *to have meaning is to be used in a certain way*, especially to play a role within certain 'transitions', such as those from seeing a red surface to 'accepting' "This is red", or from 'accepting' "This is red" to accepting "This is not green".

However, the crucial issue between the Sellarsian approach and that proposed by Horwich is the possibility of 'naturalizing' the meaning talk, of reducing it to behaviorist talk (couched in physicalist terms). Horwich claims, as we saw, that the meaning of a word consists in our accepting certain sentences which contain it. Now what does it mean, according to him, *to accept* a sentence? Horwich gives a tentative behaviorist theory of 'accepting' (p. 96) based on postulates of the kind of

For each observable fact *O* there is a sentence type "*o*" such that:

$$O \text{ is instantiated in view of } S \leftrightarrow S \text{ accepts "o"}$$

The crucial question is, whether such a theory is actually feasible at all – even if we invested it with more sophistication. The trouble is that it does not seem to be capable of allowing for the notion of an *error* – and any notion of accepting which does not provide for the possibility of *misaccepting* does not seem to fit, to underlie meaning talk. At least, so Sellars would insist.

In fact, as far as I can see, the situation is very similar in respect to the relationship between the approach of Horwich and that of Davidson. Davidson's claim, that we have to base semantics on the *irreducible* concept of truth (and hence construe meanings not as uses, but rather as truth conditions) stems precisely from his conviction that the 'acceptance' relation holding between speakers and sentences, if it is to underlie semantics, cannot be reduced to behaviorist terms. Thus, both Davidson and Sellars argue at length that to take an 'accepting' is crucial for semantics as reducible is to commit a naturalistic fallacy (for Davidson such an 'accepting' would consist in the irreducibly idiosyncratic 'holding for true', for Sellars in the irreducibly normative 'holding for correctly assertible') – and I am not sure whether Horwich's gesture towards a behaviorist reduction can be taken as a substantial counterargument.

Another point which seems to me somewhat unclear and which is crucial for assessing the depth of the disagreement between Horwich and Sellars concerns the question of holism. Horwich claims that his approach is only moderately holistic (for it embraces only those "mild" varieties of holism which are engendered by the fact that the sentences whose accept-

ance establishes the meaning of a word may contain also other words). However, it is not clear whether this should be understood as saying that a word means what it says, *exclusively* thanks to its acceptance property and *independently* of the presence of any words which are not needed to articulate its property. Imagine a language containing a single sentence, namely “Red(!)”, governed by the same acceptance property as the homophonic English sentence (namely “the disposition to apply ‘red’ to an observed surface when and only when it is clearly red”). Would it follow that “red” in this language means the same as in English? Or would this be somehow dependent on whether the language in question contains also some other words? From the Sellarsian viewpoint, it is crucial that “red” could not express our usual concept of redness, nor indeed any concept at all, if it were not able to figure in judgments – and nothing can be a judgment unless it can be negated, conjoined with other judgments etc. Horwich seems to reject this – however, if this is the case, does he not have to grant a sensor reacting to red color, the *concept* of red?

So much for the viewpoint of a basically sympathetic reader. As for unsympathetic ones, i.e., those who take for granted that meanings are some real entities labeled by corresponding expressions, as I said earlier, I doubt that they would be swayed by Horwich’s book. However, I am convinced that even the unsympathetic reader – one who hastens to dismiss the ‘crystal-clarity’ of the author’s exposition as simply a form of superficiality – should ask herself: Is it *really* the superficiality of the exposition which is at fault, might it not be that the illusion of depth which apparently escapes the author is itself the product of a certain picture which “holds us captive”? This is not to say that some criticism of this kind cannot be substantiated, but it is to urge that Horwich’s arguments should not be judged by a first impression.

The founding fathers of analytic philosophy based their revolt against bad philosophy on the conviction that many (if not all) philosophical problems are solvable by ‘conceptual analysis’. Their conviction that this would lead to the ultimate solution or a complete dissolution of traditional philosophical problems subsequently proved itself too naive, and some of the current successors of analytic philosophy appear to take this to show that the whole analytic movement was a philosophical blind alley. I think that Horwich’s book is one of those which indicate that ‘conceptual analysis’ is still something to pursue, something which can strongly aid the understanding and solving of philosophical problems.

NOTE

¹ See Wittgenstein (1953, §43): “Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache”.

REFERENCES

- Brandom, R.: 1994, *Making It Explicit*, Harvard University Press, Cambridge, MA.
Dummett, M.: 1988 ‘The Origins of Analytical Philosophy, I–II’, *Lingua e Stile* **23**, 3–49, 171–210.
Horwich, P.: 1990, *Truth*, Blackwell, Oxford.
Lewis, D.: 1972, ‘General Semantics’, in D. Davidson and G. Harman (eds), *Semantics of Natural Language*, Reidel, Dordrecht.
Sellars, W.: 1963, *Science, Perception and Reality*, Routledge, New York.
Wittgenstein, L.: 1953, *Philosophische Untersuchungen*, Blackwell, Oxford.

Charles University
Prague
Czechia
E-mail: Jaroslav.peregrin@ff.cuni.cz

JAROSLAV PEREGRIN