

Review of *Reflections on Meaning*, by Paul Horwich

Eric Swanson, ericsw@umich.edu

Forthcoming in *The Philosophical Review*, January 2009

Reflections on Meaning refines Paul Horwich's use theory of meaning. Horwich holds that the meaning of a word is constituted by the non-semantic property that best explains a certain law. For a given word, the law to be explained governs that word's use by specifying the "acceptance conditions" of a privileged class of sentences containing the word (26). Horwich devotes considerable energy to details in *Reflections on Meaning*, and focuses on especially pressing problems for his use theory of meaning. As a result the book's topics run the gamut, and the connections between its chapters are not always strong. Rather than try to provide a synoptic overview, I'll discuss three areas where it seems further clarification and detail could be fruitful: the distinction between semantic and non-semantic properties, context sensitivity, and compositionality.

Horwich thinks ours is a "fundamentally non-semantic world" (27), making it crucial that meaning be explained in non-semantic terms. In particular, he insists that we "exclude from the analyzing-properties [of word meaning] anything that would itself require analysis in terms of meaning": we can't appeal to reference, belief, or intention, for example (37). But Horwich does not object to "accounts of meaning in *psychological* terms," and his own theory relies heavily on a psychological, non-semantic relation that Horwich calls 'acceptance' (37).

It's difficult to see a substantive difference between this technical notion and belief. Acceptance, for Horwich, is "...the psychological (but non-semantic) relation to a sentence that is manifested in our relying on it as a premise in theoretical and practical inference" (40–41). Belief, on the other hand, is a relation with these properties *except* that it is semantic. Horwich's other characterizations of acceptance don't sharpen the distinction very much: "S accepts a sentence just in case that sentence, or its mental correlate, is in S's belief box" (41); "*believing* a given proposition is nothing more than *accepting* some sentence that expresses it" (61). Horwich suggests that the reader refer to his presentation in *Meaning* (1998), where he writes that acceptance is largely characterized by the principle that "For each observable fact *O* there is a sentence type '*o*' such that: *O* is instantiated in view of *S* \leftrightarrow *S* accepts '*o*'" (96). The reader is left wanting details. How, for example, should the observable fact that *S* doesn't realize that ϕ be fit into this schema?

In characterizing the difference between non-semantic and semantic properties, Horwich gives only examples of such properties, and does not say very much about the examples themselves. (Horwich's affection for this kind of move is evident throughout; some may find it frustrating.) One might wonder whether an explicit statement of the distinction would clarify the differences between belief and Horwich-style acceptance. I'm not optimistic that it would, since Horwich thinks we must generalize from belief to degrees of belief and from acceptance to degrees of acceptance (110). And at least *prima facie*, a degreed acceptance relation *should* count as semantic. One can't accept to very high degrees both "S" and "It's not the case that S," for example, in part because of what those sentences *mean*. (Perhaps a similar point could be made about even

a non-degreed acceptance relation; it's not clear what sorts of constraints might be imposed by the 'belief box.')

Horwich's discussion of the norms governing belief is also difficult to square with the projected move to degrees of belief. He says that 'is true,' in the principle that our aim ought to be to believe only what is true, "is serving merely as a device of generalization" (112). But it's plausible that credence is something like an estimate of truth value, and that the aim of credence, like the aim of other estimates, is accuracy. Accuracy here is clearly not a mere device of generalization. But Horwich's treatment of the norms of belief and meaning leans very heavily on the deflationary thought that the truth predicate is *nothing* more than a device to form generalizations. So it's unclear how Horwich's story about the norms of belief could be adapted to cover the norms of 'degreed' belief.

Let us turn to context sensitivity. On Horwich's account, the acceptance conditions of a sentence are the "circumstances" under which the sentence is "regularly accepted" (48). Given the key distinction between circumstances of evaluation and contexts of utterance, it might seem that acceptance conditions are supposed to play a role in Horwich's semantics like that played by truth conditions (as opposed to Kaplanian 'character') in contemporary intensional semantics. And some of what Horwich says about "meaning-constituting facts" (26, 49) lends plausibility to this thought. But Horwich's target is in fact meaning in "the sense ... in which 'I' has a *single* meaning in English" (28): "the meaning of a word is the common factor in the explanations of its numerous occurrences" (26), whether or not the contexts of those occurrences affect the content borne by the word. It's thus crucial to distinguish between circumstances in Horwich's sense (I'll call them 'H-circumstances') and the circumstances of evaluation that are familiar from intensional semantics ('I-circumstances'). H-circumstances bundle together context and I-circumstances, since (again) the acceptance conditions of a sentence just are the H-circumstances in which the sentence is regularly accepted, and acceptance conditions yield all the meaning there is, even for context-sensitive expressions.

Horwich isn't clear, however, about *who* must "regularly accept" a sentence in a given H-circumstance. Suppose that 'I am hungry' and 'I am tall' are in the privileged class of sentences that fix the meaning of 'I.' Presumably, then, the relevant acceptance conditions would be partially characterized by principles like *Jones accepts 'I am hungry' iff Jones is the center and Jones is hungry; Jones accepts 'I am tall' iff Jones is the center and Jones is tall; Smith accepts 'I am hungry' iff Smith is the center and Smith is hungry*, and so on. It has to be Jones and not Smith who regularly accepts 'I am hungry' in the circumstance in which Jones is the center and is hungry. But Horwich says nothing to connect the condition of being "regularly accepted" to particular contextual centers.

Horwich also needs to say more about the semantics of modal operators (and other intensional expressions). It's widely accepted that 'necessarily' means, roughly, 'in all I-circumstances,' and that 'possibly' means, roughly, 'in some I-circumstance.' Because of the sharp distinction between I-circumstance and context in intensional semantics, these operators do not affect the

interpretation of indexicals that are rigid designators—the interpretation of such expressions varies under shifts in context, but not under mere shifts in I-circumstance. As we have seen, Horwich bypasses context and I-circumstances, calling on H-circumstances alone to play the roles they standardly play together. The interpretation of indexicals *will* vary under shifts in H-circumstances, and so ‘necessarily’ and ‘possibly’ can’t mean ‘in all H-circumstances’ or ‘in some H-circumstance’—both Kaplanian ‘monsters,’ and with really bizarre meanings. That leaves it unclear what Horwich should say modal operators mean. To my mind, his most promising course is to explain exactly how to reconstruct the ‘double-indexing’ techniques that have been standard since the 1970s in terms of acceptance conditions and H-circumstances.

Finally I want to turn to compositionality. Horwich recognizes (and convincingly argues) that a semantic theory can be compositional without trafficking in truth or reference. But the particular way that Horwich wants to save compositionality raises several questions. He posits, as lexical entries, “schemata containing ‘slots’ into which items of specified functional types may be inserted to yield terms that may in turn be inserted into other schemata” (181). For example, he associates “Dogs bark” with the semantic object that results from “applying” the lexical entry associated with ‘*ns v*’ to the lexical entries associated with ‘dog’ and ‘bark’ (182).

Horwich says, repeatedly, that the “application” of lexical entries to other lexical entries is on his view just familiar Fregean functional application (181, 183, 205). Why does ‘*ns v*’ need its own lexical entry? Horwich answers that “the combinatorial procedures can then be articulated in purely argument-functional terms” (183). But they can be articulated that way without lexical entries for particular phrase structures: consider the extremely simple combinatorial procedures of type-driven semantics since Montague. At any rate, the particular syntactic entries that Horwich proposes are extremely idealized—so much so that it isn’t easy to see how, exactly, they could be made more realistic. ‘Dogs bark’ is a generic sentence (on the reading that Horwich presumably intends); many sentences of the form ‘*ns v*’ are not; generic sentences plausibly do not have the same syntax as individual-referring sentences. One is left wondering how many syntactic entries Horwich will need, at the end of the day, and what they will look like.

Horwich’s assiduity in developing his theory of meaning produces much to admire, and much to learn from. There is also much left to be done. It would surely be illuminating to see Horwich develop a probabilistic account of acceptance that would facilitate non-circular analyses of meaning, and to see him develop analyses of the particular constructions that have proven the interest of contemporary truth-conditional semantics.