

Theories of Meaning, Part II: Philosophical Logic.

Lecture II, *Davidson on Theories of Meaning*, 26th January.

Christopher J. Masterman (cm789@cam.ac.uk, christophermasterman.com)

Last week, we looked at a variety of classical semantic theories—theories which assign a particular kind of object, i.e., the meaning of that expression, e.g., referents, intensions, sets of worlds, or fine-grained propositions. We ended by looking at Davidson's overarching criticism of such theories of meaning. This week, we look at Davidson's positive proposals about the structure and purpose of a theory of meaning.

1. Davidsonian Theories of Meaning

1.1. Davidson's approach combines two ideas—truth-conditional semantics and Tarski-style definitions of truth. A truth conditional semantics ties a sentence's meaning to its truth conditions and ties the meaning of subsentential parts of a sentence to its contribution to the truth conditions of sentences in which it occurs. An adequate Tarski-style definition of truth in some object language is one which entails all true sentences of the metalanguage of the form '*S* is true if and only if *p*'.

1.2. In one sense, Davidsonian theories of meaning (DT) have much in common with the kinds of classical semantic theories we discussed last week, since both are truth conditional theories of meaning. However, DTs are distinctive and radical in that they have a very particular form. That is, for a language \mathcal{L} , a DT is a Tarski-style truth theory for \mathcal{L} —finitely many axioms and inference rules which jointly determine as theorems of the system true biconditionals of the following form, for every sentence *S* of \mathcal{L} .

(T) *S* is true-in- \mathcal{L} if and only if *p*

('*S*' is a name or description of a sentence of \mathcal{L} and *p* is a sentence of the metalanguage.)

1.3. Crucially, Davidson's understanding and use of T-Schemas here departs in significant ways from Tarski's. The fundamental difference is that Davidson hopes to use a prior notion of truth to develop theories of meaning—Davidson 'inverts' Tarski. For Tarski, *p* in (T) is a translation of *S* in the metalanguage. This won't work for Davidson. Moreover, Tarski was only interested in formal languages. For Tarski, natural languages were problematically messy and imprecise and, worse still, semantically closed—they contained their own truth predicate. Davidson is optimistic about making progress, despite this. He suggests only considering fragments of natural language (no truth predicate, restricted quantifiers).

1.4. Recall that Davidson's chief objection to classical semantic theories was that they issued trivialities. Could the same charge not be raised against DTs? Consider a DT for English which use English as the metalanguage. We should expect the following to be a theorem of the theory:

(1) 'James is cool' is true-in-English iff James is cool

It is tempting to think that (1) is trivial and worry that Davidson's approach will not generate substantive claims about meaning. This would be a mistake, however. Importantly, the triviality here is illusory. In (1), the object language and metalanguage are the same. So, anyone who could *understand* (1), could already understand 'James is cool'. We can illustrate the substance of DTs in the case where the metalanguage and object language differ. Since *S* in (T) names a sentence in the object language and *p* is a sentence of the metalanguage, in a DT for Norwegian with English as the metalanguage, we have theorems like:

(2) 'Han elsker edderkoppene' is true-in-Norwegian iff he loves the spiders

Fundamentally: the left of (1) and (2) are about language, the right is about the world.

1.5. A natural question concerns the variety of semantic phenomena, broadly construed, which can be captured by this kind of truth theory. One version of this worry: what would a Davidsonian theory of meaning have to say about commands, questions, promises, etc.? For Davidson, such aspects of meaning are the purview of pragmatics—how language is *used*. We, first, settle questions of meaning for ordinary declarative sentences and, second, develop a pragmatics, see (Glüer, 2011).

1.6. Another version of this worry: what should we say about ambiguity? For instance,

(3) John went to the bank

(3) is ambiguous. What theorems should a good truth theory generate for (3)? In (Davidson, 1967), the answer is straightforward: in a homophonic theory of meaning, if a sentence s is ambiguous, the ambiguity is replicated in the metalanguage, i.e.,

(4) 'John went to the bank' is true-in-English iff John went to the bank

(4) is true, regardless of how we understand 'bank', i.e., as in river bank or financial institution. Another proposal would be to think that all so-called semantic ambiguity is in fact syntactic, i.e., we should distinguish between 'bank₁' and 'bank₂' and our theory of truth will contain distinct theorems for each.

2. Interpretive Davidsonian Theories

2.1. A Tarski-style theory of truth, as Davidson envisages it, is quite minimally constrained. But we want DTs which are *interpretive*. Roughly, we want to be able to know what the sentences of \mathcal{L} mean by knowing the relevant theory of meaning for \mathcal{L} . An initial worry is that simply constraining DTs to produce *true* biconditionals of the form (T) for \mathcal{L} will not result in only interpretive theories of meaning for \mathcal{L} . Consider:

(5) 'Snow is white' is true-in-English iff grass is green

(6) 'Snow is white' is true-in-English iff snow is white

(5) is true, but not interpretive: the sentence on the right is not synonymous with the sentence mentioned on the left. Which further constraints could narrow down the class of DTs to the interpretive ones?

2.2. One constraint would appeal to the compositionality. Languages are compositional and DTs are finitely axiomatised to reflect that. Axioms will specify what terms refer to, the conditions under which predicates are satisfied, how terms and predicates are concatenated, and the conditions under which such concatenations are true. The right axioms could plausibly differentiate (5) from (6). However, this would not eliminate *systematic* non-synonymy (cf. Quine and so-called proxy functions). See, (Fodor and Lepore, 1992).

2.3. Another constraint would require theorems to be law-like. DTs thus contain *law-like* biconditional theorems, i.e., biconditionals which support counterfactual. Thus, (5) and (6), if they are part of an adequate theory of meaning, must hold true even in situations where 'Snow is white' and 'Grass is green' diverge in their truth value (Davidson, 2005: 54). However, this would not eliminate non-synonymous, but necessarily extensionally equivalent biconditional theorems, see (Fodor and Lepore, 1992), e.g., replacing (4) with:

(5) 'Snow is white' is true-in-English iff snow is white and $2+2=4$

(Indeed, a related point concerns the underlying *logic* of our DT. Any T-sentence ' S ' is true-in- \mathcal{L} iff p entails ' S ' is true-in- \mathcal{L} iff $p \wedge \top$, provided we have a strong enough substitution principle in the logic. If the former is a theorem, the latter is. We, thus, must be careful in formulating the logic part of our DTs.)

2.4. Davidson also takes it that any adequate DT should be knowable under radical interpretation. Davidson argues that radical interpretation is possible given the Principle of Charity and the notion of *holding true*, i.e., a special mental state recognition of which does not presuppose *content*. Note that this would not get you interpretive DTs at the subsentential level, i.e., it does not eliminate indeterminacy at the level of reference.

3. Theories of Meaning and Linguistic Understanding

3.1. A central feature of the phenomenon that is language is that languages are learnt, mastered, and we possess sophisticated and rich linguistic knowledge. This learnability of language has featured in our discussion of theories of meaning so far *viz.* DTs are finitely axiomatised, since we are finite and capable of understanding an infinity of distinct sentences and theories of meaning attempt to model this understanding.

3.2. It's worth asking how a DT might fit into a broader project of explaining, or at least better understanding, our linguistic capacity. That is to ask, what can be explained about linguistic understanding and knowledge by modelling our linguistic capabilities as (implicit) knowledge of the relevant truth theory? To answer this, we must recognise that linguistic understanding is a heterogeneous phenomenon. At the very least, linguistic understanding, of course, involves knowledge what the semantically significant parts of our language mean. But it also at least involves a special kind of structured knowledge, and a knowledge/capacity to reason between language and the world and *vice versa*.

3.3. Next week, we'll look more closely at how a DT is supposed to ground knowledge of what parts of our language mean. But, to end, it's worthwhile noting how modelling linguistic capacity and tacit linguistic knowledge as a DT can offer explanations of other aspects of our linguistic understanding.

Structured Knowledge. Linguistic understanding does not consist solely in terms of knowledge of meaning, but also a knowledge of the structure of language, i.e., its grammar and how subsentential and sentential meaning relate. It's very plausible to posit that linguistic understanding is governed by the tacit acceptance of a finite set of rules governing the construction and interpretation of language. Theories of meaning as compositional, finitely axiomatized T-theories model this nicely. If our linguistic understanding was structurally similar to such T-theories, there is a nice explanation of how we systematically understand some sentences because we understand others.

Reason from Language to World and Back: Part of linguistic understanding licenses is us to reason correctly from language to the world. If I understand what 'snow is white' means and I know that it is true, then I know that snow is white. Theories of meaning as T-theories nicely capture this distinctive kind of reasoning. Indeed, this relation between language and the world which is part and parcel of linguistic understanding is at the very heart of the theorems of such theories of meaning. If I know what 'snow is white' means, then (at least tacitly), I know that 'snow is white' is true iff snow white. (Contrast *this* explanation with what could be said, if meanings were, e.g., Russellian propositions.)

References

- Davidson, Donald (1967). Truth and Meaning. *Synthese* 17, 304–323.
- Davidson, Donald (2005). *Truth and Predication*. Ed. by Donald Davidson. Cambridge, Mass.
- Fodor, Jerry A. and Ernest Lepore (1992). *Holism: A Shopper's Guide*. Ed. by Ernest LePore. Cambridge, Mass., USA: Blackwell.
- Glüer, Kathrin (2011). *Donald Davidson: A Short Introduction*. New York, US: Oup Usa.