Lepore and Ludwig on 'Explicit Meaning Theories'

Philosophical Studies, online first 2012. This is the penultimate version, please refer to the published version. Comments are welcome: miguel.hoeltje[ed]uni-due[dodd]de.

The fundamental problem proponents of truth conditional semantics must face is to specify what role a truth theory is supposed to play within a genuine meaning theory. This problem arises from the simple fact that truth theories are too weak to perform the job of a meaning theory. The most detailed proposal for tackling this problem is also the most recent one: the account developed in a series of publications by Ernest Lepore and Kirk Ludwig.[1] Lepore and Ludwig take the challenge head on, and consequently, their proposal goes well beyond a purely truth-theoretic approach. The effort is commendable and makes their suggestion interesting, for if they succeed they will have shown how to overcome a serious obstacle that has plagued the truth-theoretic project from its inception. Unfortunately, the account fails. As I will show in this paper, theories along the lines of Lepore and Ludwig do not suffice to put someone into the position to understand the objectlanguage, since they have to presuppose knowledge of the language of the truth theory. The fundamental problem of truth conditional semantics thus remains unsolved.

a. Stage setting

I will employ a minimal notion of a meaning theory: a meaning theory for a language is a finite theory knowledge of which would put someone into the position to understand the language in question.[2] One of the explanatory tasks of meaning theories in this sense is to show us how a finite being could be competent with an infinite language. Instead of talking of a subject *understanding* a sentence or a language, I will often talk of the subject *knowing* what the sentence means (or knowing what the sentences of the language mean). It is controversial what exactly the connection between linguistic understanding and knowledge of meaning is. But it will be clear that nothing in the following discussion hinges on possible differences between understanding and knowledge of meaning. Thus, we can also say: a meaning theory for a language is a finite theory knowledge of which would put someone into the position to know of every sentence of the language what it means. By a *theory* I simply mean a set of sentences. Importantly, the notion of *knowing a theory* should be understood such that one can know a theory without knowing the language of the theory: by saying that a subject α knows a theory {'p_1', …, 'p_n', …}, I mean simply that (α knows that p_1, and …, and α knows that p_n, …).[3]

The above characterization of meaning theories relies on a notion of a subject's being in a position to know something. Let me briefly indicate what this comes to. I will presuppose that sometimes knowing that *p* would in principle suffice to know that *q*. This seems to be the case if,

[1] Lepore & Ludwig (2005), (2006), (2007), (2011); Ludwig (2002).
[2] At least within the context of Davidsonian semantics, this conception is clearly central; cp. Davidson (1973: 125); (1976: 171); (1977: 215); (1990: 312).
[3] This corresponds to Davidson's use of 'knowing a theory'; cp. Davidson (1976: 175).

for example, that *p* has it as a logical or conceptual consequence that *q*; but there may be other cases. When knowing that *p* would in principle suffice to know that *q*, I will say that anyone who knows that *p* is in a position to know that *q*. Being in a position to know something does not entail knowing it; but whenever you are in a position to know that *p*, the information that *p* is within your epistemic reach. I will say that the information that *p* would *put* α into a position to know that *q*, if the following two conditions are fulfilled: *(i)* α is not already in a position to know that *q*, but *(ii)* if the information that *p* were added to α's stock of knowledge, α would be in a position to know that *q*. Thus a meaning theory for a language *L* must express something such that anyone who knows it is in a position to know the language *L*.

Proponents of truth-theoretic semantics claim (minimally), that a truth theory has an important role to play in a meaning theory. I will now introduce the truth theoretic apparatus in a way that prepares the discussion of Lepore & Ludwig's (L&L) proposal. A *T-sentence* is a sentence of the form '*S* is true iff *p*', where '*p*' does not contain semantic vocabulary. An *interpretive* T-sentence is one where '*p*' translates *S*. A *truth theory* for *L* is a theory which entails, for every *L*-sentence, an interpretive T-sentence. A truth-theoretic axiom for an objectlanguage expression *E* is *interpretive* iff the expression used to state *E*'s contribution to the truth or satisfaction conditions of sentences containing it translates *E*. Thus, since 'is wise' translates the German 'ist weise', but 'is wise and 2 = 2' does not, A_1 is interpretive, while A_2 is not:

A₁ $\forall t$ ($t^\frown$'ist weise' is true$_G$ iff Ref$_G$[t] is wise).
A₂ $\forall t$ ($t^\frown$'ist weise' is true$_G$ iff (Ref$_G$[t] is wise and 2 = 2)).

An *interpretive truth theory* is a truth theory all of whose (semantic) axioms are interpretive. Relative to the full deductive apparatus of, say, classical logic, every truth theory, even an interpretive one, will have non-interpretive T-sentences as theorems. It is thus helpful to employ only a restricted deductive machinery. In particular, the inference rules should, given a set of interpretive axioms, allow the derivation of a T-sentence only if it is interpretive. For if we have such a theory, the fact that it yields a T-sentence will be enough to ensure that the move to a genuine meaning-ascription is truth preserving.[4] Let *A* be a set of interpretive axioms for the expressions of a language *L* and *R* a set of inference rules such that: (*i*) for every sentence *S* of *L*, the rules in *R* allow the derivation of a T-sentence for *S* from the axioms in *A*, and (*ii*) every T-sentence derivable from *A* by the rules in *R* is interpretive. Following L&L, I will call a proof of a T-sentence from *A* using only the rules in *R* a *canonical proof*. Thus, an interpretive truth theory which employs only such a set of inference rules will have interpretive T-sentences for all the *L*-sentences among its theorems and it will not yield any non-interpretive T-sentences. In the terminology of Larson & Segal, such a theory is neither *over-* nor *underproductive*.[5]

Before turning to L&L's proposal, it is helpful to motivate their general approach by drawing attention to the following fact: Truth theories are too weak to perform the task of a meaning theory. By themselves, they provide information of a purely extensional kind. Thus, a truth theory may tell

[4] Cp. Larson & Segal (1995: 40 FN), Kölbel (2001), (2002: 77), Ludwig (2002: 159).
[5] Larson & Segal (1995: 34).

us that ‘Sokrates ist weise’ is true iff Socrates is wise; but this is a feature it shares with all true sentences, regardless of what they say. And it may tell us that the term ‘Sokrates’ refers to Socrates; but this is a feature it shares with all coreferring expressions, regardless of their meaning. Finally, the theory may tell us that the predicate ‘is weise’ is satisfied by something iff it is wise; but this is a feature it shares with all coextensional predicates, regardless of what they mean. Hence, the information provided by such a theory can hardly suffice to know that ‘Sokrates ist weise’ means that Socrates is wise.

b. Presupposing knowledge of a language

The fact that truth theories are too weak to serve as meaning-theories has frequently been pointed out.[6] But at least as frequently, authors are either unaware of or chose to ignore this basic point.[7] L&L are a welcome exception. I will lay out their proposal in the next section and I will argue that it presupposes knowledge of a language in a way which pre-empts the resulting theories from doing their explanatory work. Before this charge can be levelled, we need to be clear about what ‘presupposing knowledge of a language’ comes to. It will be important to be precise here, for two reasons. First, there are several notions one may associate with this locution and only one of them is relevant. Second, it should be noted that the diagnosed problem is distinctive to L&L’s proposal. As I will point out, other approaches to formulate meaning-theories from a Davidsonian perspective do not presuppose linguistic knowledge in the relevant sense (although they founder on different grounds).

First, we may want to say that a theory *T* presupposes knowledge of a *particular* language iff someone could know *T* only if she knew *that very language*. But this notion is empty: no theory presupposes knowledge of any language in this sense. Knowing a theory only requires knowing the theory’s content. Granted, the theory itself belongs to some language; but just as you need not know what ‘Schnee ist weiß’ means to know that snow is white, you need not know what the sentences of a theory mean to know the theory’s content. But if knowledge of the language of the theory is not necessary, then no knowledge of *any particular* language will be required. However, we might want to hold that knowing a theory requires knowledge of *some* language. In particular, we might claim that someone can only know that *p* if she had a language in which *that knowledge* could be expressed. If this is correct, then every theory will presuppose knowledge of a language. I can remain neutral on this issue, for this not the sense of presupposing knowledge of a language that I am interested in vis-à-vis L&L’s proposal.

Sometimes knowing something would be sufficient for knowing something else. The following theory provides information which would put someone into the position to know that there are infinitely many natural numbers: T_1 = {‘0 is a natural number’, ‘Every natural number has a successor, which is a natural number’, ‘If the successor of *x* is the successor of *y*, then *x* is *y*’, ‘0 is not the successor of anything’}. Contrast T_1 with T_2 = {‘T_1 is true’}. T_2 tells us that the sentences of

[6] E.g. Foster (1976), Soames (1992), (2008), Ludwig (2002).
[7] E.g. Davidson (1990: 312), Heim & Kratzer (1998: 1).

another theory are true. Since you can know that a sentence is true without knowing what it means, knowing T_2 does not suffice to know that there are infinitely many natural numbers. Hence, with respect to the task at hand, T_2 presupposes *linguistic* knowledge: T_2 puts someone α into the position to know that there are infinitely many natural numbers *only if* α knows what the sentences of T_1 mean. This, then, is the relevant sense of presupposing linguistic knowledge: relative to the task of putting someone into the position to know that *p*, a theory *T* presupposes knowledge of a language *L* iff (knowing *T* would put someone into the position to know that *p* only if she already *is* in a position to know *L*).

Before I address L&L's account, it should be noted that presupposing linguistic competence in the relevant sense is *not* a problem that other approaches in the Davidsonian tradition face. Let me give three examples. First, take the proposal that truth theories of the right kind simply *are* meaning theories. According to this proposal, a meaning theory will tell us things like the following: 'Sokrates ist weise' is true iff Socrates is wise. While this approach fails for the above reasons - truth theories provide too little information to yield knowledge of meaning - knowledge of the relevant *metalanguage* is not the issue. Being competent in, say, English will not help you in utilizing the relevant information to come to understand German, since this information has simply nothing to do with English. Knowledge of the metalanguage adds nothing; it neither helps nor hinders. Second, take Davidson's more elaborate proposal to construe a meaning theory as a theory about the *content* of a truth theory.[8] This approach, too, does not presuppose linguistic knowledge in the relevant sense. According to this proposal, a meaning theory will tell us things like the following: There is a truth theory for German which states that 'Sokrates ist weise' is true iff Socrates is wise. This proposal fails for reasons similar to those that sunk the simple proposal.[9] But, more importantly for my present concern, knowledge of the relevant *metalanguage* is again immaterial. Being competent in, say, English will not help you in utilizing the relevant information to come to understand German since, again, this information has nothing to do with English. Finally, take Higginbotham's proposal, according to which a meaning theory tells us things like the following: It is common knowledge among competent speakers of German that 'Sokrates ist weise' is true iff Socrates is wise.[10] Again, while this proposal faces serious problems,[11] presupposing linguistic competence is not one of them. Just as before, being competent in, say, English will not help you in utilizing the relevant piece of information to come to understand German, since the fact that something is common knowledge among speakers of German has nothing to do with the English language (even if this fact is *stated* in English). It cannot be said, then, that presupposing linguistic knowledge in the relevant sense is a problem approaches from the truth-theoretic camp face in general.

[8] Davidson (1976: 174f.).
[9] See e.g. Soames (2008).
[10] Higginbotham (1992).
[11] Again, see Soames (2008).

c. Lepore & Ludwig's proposal

L&L's proposal is what we might call a *meta-theoretic* approach; they suggest that a meaning theory for a language should be construed not simply as a truth theory, but as a theory *about* a truth theory. To keep things simple, I focus on a meagre fragment of German containing only the sentence 'Sokrates ist weise', and ignore all complicating factors like ambiguity and context sensitivity. The following two axioms form an interpretive truth theory T_G for the relevant fragment of German:

S Ref_G['Sokrates'] = Socrates.
W $\forall t$ ($t^\frown$'ist weise' is true_G iff $\text{Ref}_G[t]$ is wise).

Using the rules of universal instantiation and identity elimination, we can derive the interpretive T-sentence '"Sokrates ist weise" is true_G iff Socrates is wise' from T_G. Furthermore, restricting ourselves to these two rules, we are not able to derive any non-interpretive T-sentences; i.e. with respect to T_G and the two rules $\forall$E and =E, we can define a *canonical proof in* T_G as a finite sequence of formulas the last of which is a T-sentence, such that every element is either an axiom of T_G or can be derived from previous elements by $\forall$E and =E. We now have everything in place to give the axioms that – according to L&L – would constitute a meaning theory for *G* ('*LT*' denotes the language of the truth theory T_G):[12]

A_1 Every instance of the following schema is true_{LT}:

S means_L that *p* iff it is canonically provable on the basis of the axioms of an interpretive truth theory for *L* that *S* is true_L iff *p*.

A_2 T_G is an interpretive truth theory for *G* whose axioms are 'Ref_G["Sokrates"] = Socrates', '$\forall t$ ($t^\frown$"ist weise" is true_G iff $\text{Ref}_G[t]$ is wise)'.

A_3 'Ref_G["Sokrates"] = Socrates' means_{LT} that the referent_G of 'Sokrates' is Socrates.
'$\forall t$ ($t^\frown$"ist weise" is true_G iff $\text{Ref}_G[t]$ is wise)' means_{LT} that (the concatenation of a term with 'ist weise' is true_G iff the referent_G of that term is wise).

A_4 A canonical proof in T_G is a finite sequence of formulas the last of which is a T-sentence, such that every element is either an axiom of T_G or can be derived from previous elements by $\forall$E and =E.

I will call this theory 'M_G'. Before turning to my critique, it might be helpful to sketch why one may hope that theories like M_G could qualify as meaning theories. Roughly, the thought is this: an interpretive truth theory in conjunction with a canonical proof apparatus will yield a T-sentence '*S* is true iff *p*' iff *S* means that *p*. Thus, in some sense, there is a very close connection between the way a sentence is characterized by such a theory (e.g. as *being true iff* Socrates is wise) and the way it should be characterized by a genuine meaning theory (as *meaning that* Socrates is wise). But the theory itself does not *state* that this connection obtains and hence does not supply information sufficient for interpretation. The idea behind theories along the lines of M_G is that, once we

[12] In various places, L&L present versions of their proposal that differ somewhat from each other. I will concentrate on their 2007 monograph, which includes the most comprehensive presentation of their approach; see e.g. Lepore & Ludwig (2007: 90).

supplement a suitable truth theory with explicit information about the connection in question, we have filled the gap between a mere truth theory and a meaning theory.

Suppose that at t_1, a subject α neither knows *G*, the objectlanguage, nor *LT*, the language of the truth theory, and that at some later time t_2, α acquired all the information expressed by the axioms of M_G. Consider the following two claims:

C_1 At t_2, α is in the position to know *G*.
C_2 At t_2, α is in the position to know *LT*.

Here is my argument against L&L:

1. M_G is a meaning theory for *G* iff C_1 holds.
2. C_1 holds only if C_2 holds.
3. C_2 must be rejected.
Therefore: M_G is not a meaning theory for *G*.

The first premise simply follows from the relevant notion of a meaning theory. I will now argue for the second and third premise.

C_1 holds only if C_2 holds. In other words: If C_2 does not hold, neither does C_1. Assume that C_2 does not hold. If C_1 holds, then, at t_2, α should be in the position to know that 'Sokrates ist weise' $means_G$ that Socrates is wise. Is this the case under the current assumption? From her knowledge of M_G, α can extract the following information:

I_1 T_G is an interpretive truth theory for *G*. [From A_2].
I_2 There is a canonical proof in T_G of '"Sokrates ist weise" is $true_G$ iff Socrates is wise'. [From A_2 and A_4].
I_3 There is canonical proof from the axioms of an interpretive truth theory for *G* of '"Sokrates ist weise" is $true_G$ iff Socrates is wise'. [From I_1 and I_2].

Under the current assumptions, knowing $I_1 – I_3$ is not sufficient to know the meaning of the target sentence: Since α does not know *LT*, the information contained in $I_1 – I_3$ is restricted to *(i)* syntactic information about sentences not understood by α, and *(ii)* information to the effect that certain expressions of an alien language (*LT*) are synonymous with certain expressions of another alien language (*G*). Which additional information is available to α? Since α knows that A_3, she knows what the axioms of the truth theory T_G mean. On the current assumption, this information is not sufficient to know what the *theorems* of T_G mean; in particular, it will not be sufficient to know what the T-sentence '"Sokrates ist weise" is $true_G$ iff Socrates is wise' means. It seems clear, then, that the information provided by A_3 is of no help in the current case. This leaves A_1. This axiom tells α that all instances of a certain schema are $true_{LT}$. The pertinent instance is:

IN 'Sokrates ist weise' $means_G$ that Socrates is wise iff it is canonically provable on the basis of the axioms of an interpretive truth theory for *G* that 'Sokrates ist weise' is $true_G$ iff Socrates is wise.

Since α knows that A_1, we can add another entry to the list of things α knows:

I_4 **IN** is true_{LT}.

Recall that we are currently working under the assumption that M_G does not put α into the position to know *LT*. Hence, the information expressed by I_4 is nothing but further information about the truth of a sentence that α does not understand. In particular, while α is in a position to know that I_4, this does not put her into the position to know that IN. This exhausts the information provided by M_G. Premise 2 is thus substantiated: If C_2 does not hold, neither does C_1.

Let us now turn to the third premise. According to this premise, a theory like M_G does not provide information sufficient for knowing the language of the truth theory. Would knowing M_G put someone into the position to know *LT*? L&L do not address this question explicitly, but there is some indication that they believe this. For why else would they have their 'explicit meaning theories' include information about the meaning of the axioms of the truth theory? Knowing what the axioms mean could only be relevant if it resulted in an understanding of the *theorems* – for it is the theorems, the interpretive T-sentences in particular, that may be used to connect objectlanguage sentences with their synonymous counterparts of *LT*. But it is hard to see how the one could yield the other: Knowing of a few sentences what they mean – *en bloc*, as it were – does not suffice to know the meaning of all the infinitely many sentences that can be derived from these sentences by purely syntactic transformations (the canonical proof apparatus), the *semantic* significance of which remains completely opaque. But then knowing M_G does not suffice for being in a position to know *LT*, the language of the truth theory. Thus, the third premise is substantiated. Hence, L&L's proposal does not deliver meaning theories.

Could theories like M_G be supplemented in such a way that they provided information sufficient for knowing the language of the truth theory? If this was possible, then the resulting theories would indeed count as meaning theories for the objectlanguage. But it is important to note that this is just the problem of giving meaning theories all over again: in order to use a theory like M_G to interpret an infinite objectlanguage, we need to fully understand the language of the truth theory, which must itself be infinite. Hence, in order to be a meaning theory for the objectlanguage, the supplemented version of M_G would in effect have to be a meaning theory for the language of the truth theory. But the question with which we started – and which thus far has not been answered – was precisely how meaning theories could be formulated. It falls short of an answer to point out that, given I already have a meaning theory for some infinite language *L*, I can produce a meaning theory for some other language *L**.

d. Conclusion

A meaning theory in Davidson's sense must provide information which, taken by *itself*, would suffice to interpret all the sentences of the objectlanguage. Such a theory must not presuppose *any* further information in the sense that this further information, while not being provided by the theory, is necessary to utilize it to interpret the objectlanguage. But theories along the lines suggested by L&L presuppose linguistic knowledge: they can be used to interpret the objectlanguage only by someone who knows the language of the truth theory, while this knowledge

cannot rest solely on the information provided by the theory in question. L&L's 'explicit meaning theories' are not meaning theories.

It might be helpful to dispel a misplaced worry that could arise in connection with this line of argument. Here is the worry: As semanticists, don't we *always* have to presuppose knowledge of a language in the pertinent sense? After all, *any* semantic theory will have to be formulated in a language that is at least as rich as the objectlanguage; there simply is no way to describe or explicate the function of language without using language. This last observation is clearly correct. But it just as clearly does not substantiate the worry at issue. Yes, meaning theories could only be formulated by *using* language. But the same goes for theories about the function of the human lungs (or anything else, for that matter). Surely we should not take *this* to entail that there could be no theory knowledge of which would suffice to understand the workings of the human lungs without presupposing knowledge of, say, English. Theories typically do not, and thus do not have to, presuppose knowledge of a language in the relevant sense. This goes for theories that happen to be about language as much as it goes for theories about lungs.

The last remarks should also make it clear that I am not putting forward a *general* critique of the project of giving meaning theories. My argument is directed against one particular approach from within one particular tradition and, as I have indicated, it could not be levelled against other proposals in that tradition. Unlike L&L's account, Davidson's simple and revised proposal or Higginbotham's approach do not talk about the language of the truth theory at all, and their failure is thus *not* due to presupposed linguistic knowledge in the relevant sense. They fail because the information they provide about objectlanguage sentences is too weak to yield knowledge of meaning. L&L's hope was that by going meta-linguistic, this problem could be circumvented. I have argued that this move is unsuccessful: it introduces the language of the truth theory as a further object of inquiry without ensuring that this language be understood.[13]

Campbell, J.K. et. al. (ed.) (2002): Meaning and Truth: Investigations in Philosophical Semantics, New York: Seven Bridges Press.

Davidson, D. (1973): „Radical Interpretation", in Davidson (2001), 125–139.

— (1976): „Reply to Foster", in Davidson (2001), 171–179.

— (1977): „Reality Without Reference", in Davidson (2001), 215–225.

— (1990): „The Structure and Content of Truth", *The Journal of Philosophy* **87**, 279–328.

— (2001): *Inquiries into Truth and Interpretation*, 2. Auflage, Oxford: Clarendon Press.

Foster, J. (1976): „Meaning and Truth-theory", in: Evans, G. & McDowell, J. (Hg.) (1976), 1–32.

Heim, I. & Kratzer, A. (1998): *Semantics in Generative Grammar*, Oxford: Blackwell.

Higginbotham, J. (1992): „Truth and Understanding", *Philosophical Studies* **65**, 3–16.

Kölbel, M. (2001): „Two Dogmas of Davidsonian Semantics", *Journal of Philosophy* **98**, 613–635.

— (2002): *Truth without Objectivity*, London: Routledge.

Larson, R. & Segal, G. (1995): *Knowledge of Meaning*, Cambridge MA: MIT Press.

Lepore, E. & Ludwig, K. (2005): *Donald Davidson: Meaning, Truth, Language, and Reality*, Oxford: Oxford University Press.

— (2006): „Ontology in the Theory of Meaning", *International Journal of Philosophical Studies* **14**, 325–335.

— (2007): *Donald Davidson's Truth-Theoretic Semantics*, Oxford: Oxford University Press.

— (2011): „Truth and Meaning Redux", *Philosophical Studies* **154**, 251–277.

[13] For discussion and helpful comments I would like to thank Lars Dänzer, Nick Haverkamp, Kirk Ludwig, Benjamin Schnieder, Moritz Schulz, Alex Steinberg, and an anonymous referee.

Ludwig, K. (2002): „What is the Role of a Truth-theory in a Meaning-theory?“, in: Campbell, J.K. et. al. (Hg.) (2002), 142–163.
Soames, S. (1992): „Truth, Meaning, and Understanding“, *Philosophical Studies* **65**, 17–35.
— (2008): „Truth and Meaning in Perspective“, *Midwest Studies in Philosophy* **32**, 1–19.