Ontology in the Theory of Meaning

Ernest Lepore and Kirk Ludwig

Abstract

This paper advances a general argument, inspired by some remarks of Davidson, to show that appeal to meanings as entities in the theory of meaning is neither necessary nor sufficient for carrying out the tasks of the theory of meaning. The crucial point is that appeal to meaning as entities fails to provide us with an understanding of any expression of a language *except* insofar as we pick it out with an expression we understand which we tacitly recognize to be a translation of the term whose meaning we want to illuminate by the appeal to assigning to it a meaning. The meaning drops out as irrelevant: the work is done, and can only be done, by matching terms already understood with terms they translate.

Keywords: theory of meaning; meanings; compositionality; ontology; Davidson

You say: the point isn't the word, but its meaning, and you think of the meaning as a thing of the same kind as the word, though also different from the word. Here the word, there the meaning. The money, and the cow that you buy with it. (But contrast: money, and its use.)

(Wittgenstein)

1 Introduction

Philosophers since Frege have quantified over meanings to help us to understand how we understand the languages we speak. There are been notable sceptics of this tradition, such as Quine (Quine, 1953, 1960), Davidson (Davidson, 2001 (1967)), and also, in a different tradition, as our epigraph indicates, Wittgenstein (Wittgenstein, 1950). Quine urged complete nihilism not only about meanings as entities, but about even the notions of synonymy and analyticity. Davidson has urged that all the work of the theory of meaning can be done within a framework that makes no essential appeal to meanings as entities. This paper advances a general argument, inspired by some remarks of Davidson, to show that appeal to meanings as entities in the

theory of meaning is neither necessary nor sufficient for carrying out the tasks of the theory of meaning. The crucial point is that appeal to meanings as entities fails to provide us with an understanding of any expression of a language, *except* insofar as we pick it out with an expression we understand which we tacitly recognize to be a translation of the term whose meaning we want to illuminate by the appeal to assigning to it a meaning. The meaning drops out as irrelevant: the work is done, and can only be done, by matching terms already understood with terms they translate. This makes way for seeing a statement of appropriate knowledge about a truth theory doing all the work that needs to be done and that can be done in the theory of meaning, and it shows that there is an interesting sense – though it is not the one he intended – in which Wittgenstein's claim in the *Tractatus Logico-Philosophicus* (Wittgenstein, 1961), that the facts about how our language represents the world cannot be stated but can only be shown, is correct.

2 The Project

Construed broadly, the project of the theory of meaning is to explain how we understand the languages we speak. To conceive of it as a philosophical project, we want to abstract away from facts about how any particular set of speakers understand the languages they speak and focus on facts about what's involved in any conceivable speaker understanding a language. This involves saying both how it is that speakers understand individual words and how speakers understand complex expressions – ultimately and centrally, sentences.

The introduction of meanings as entities to help us understand individual words seems on the face of it fatuous. We might stretch a point and allow as Russell did that the meaning of a proper name is the individual it refers to, so that we are indeed informed of the meaning of 'Sir Walter Scott' by being informed that it refers to Sir Walter Scott – provided that we can do this in a way that does not simply use the words whose meaning we want to be informed about. We might point, for example, to the individual, saying, 'That's him', or identify him as the author of *Waverley*. Let us try to explain what a noun such as 'author' means, however, by saying that it is the sense or meaning of 'author', and to explain our understanding of the word by saying that it consists in 'grasping' its meaning, and it is immediately apparent that we are merely playing with words. No one, given these explanations, would be any the wiser about what 'author' means in English or what understanding it comes to.

Meanings, construed as entities, begin to look more useful when we come to try to explain how we understand complex expressions on the basis of our understanding of the simpler expressions that are combined in them and their arrangement. Individual words are meaningful. The meaningful complexes in a language obviously are understood on the basis of their parts

and mode of combination. Assign the individual words meanings, i.e., things which we call meanings, and we can then assign to the complex a meaning which we think of as composed in some suitably abstract sense out of the meanings of the words. We have then a structured entity at the level of meaning that corresponds to the structured syntax of the complex expression. The illusion of understanding is increased when we realize that this makes available to us the apparatus of quantification theory in giving a systematic account of the meanings of complex expressions on the basis of the meanings of their parts and mode of combination. The sense of understanding is illusory, however, because what is essential to this approach can be preserved while leaving us completely in the dark about the language for which we give such a theory.

To show this, we first lay down a criterion of adequacy on a meaning theory which is to enable us to understand complex expressions on the basis of understanding their parts and mode of combination:

[C]   A meaning theory M for a language L is adequate only if it enables someone who understands it to understand any potential utterance of a sentence in the language given an understanding of its primitive expressions.

In the next section, we give a sample meaning theory, in a neo-Fregean style, that satisfies [C], for a compositional language with an infinity of non-synonymous sentences. In the section following, we show that what is essential to it, the systematic assignment of meanings as entities to expressions, can be retained *without* satisfying [C], and identify the crucial mechanism at work in satisfying [C]. We then draw some general conclusions about the inutility of meanings in the theory of meaning, where illumination in the theory of meaning is to be sought, and what kinds are available.

3   A Neo-Fregean Meaning Theory

Davidson is famous for having claimed that there are insuperable difficulties in the way of formulating a compositional meaning theory which quantifies over meanings (Davidson, 2001 (1967): pp. 19–21). However, it can be done, for a well-understood language, with the resources of classical quantification theory, if the only object is to generate true theorems for each object-language sentence of the form [M]:

[M]   *s* means *p*

The trouble is not that it cannot be done, but that the meanings we quantify over do no real work.

The theory we present in this section treats every meaningful word unit as having assigned to it an entity which is understood to be its meaning. These entities are to be individuated as finely as equivalence classes of synonymous expressions, and thus as finely as Fregean senses. Departing from Frege, we will treat expressions as referring to their meanings. We will suppose also, in contrast to Frege, that the meaning of a proper name is just its referent (though this is inessential). The basic idea is to introduce a rule giving the meaning of a complex expression as a function of the meaning of predicative terms, treated as functional terms, and their argument terms.

Take the simplest case of a subject–predicate sentence. Let us interpret 'means' as 'refers to'. We begin with the following axioms. We presuppose appropriate definitions of 'formula', 'sentence', and the other terms employed below for expressions in various syntactic categories.

A1   Means('Caesar', *Caesar*)
A2   Means('x is ambitious', *x is ambitious*)
A3   For any proper name α, for any predicate $\prod$, the result of placing α in argument position for $\prod$ means the value of the meaning of $\prod$ given the meaning of α as argument.
A4   The value of any sentential function for an argument denoted by a referring term is denoted by the expression that results from placing the referring term in the argument place of the sentential function.

Instantiate A3 to 'Caesar' and 'x is ambitious' to get 1,

1   'Caesar is ambitious' means the value of the meaning of 'x is ambitious' given the meaning of 'Caesar' as argument.

The meaning of 'x is ambitious' is *x is ambitious*, and the meaning of 'Caesar' is *Caesar*, by A2 and A1 respectively. The value of the meaning of 'x is ambitious' given the meaning of 'Caesar' as argument is *Caesar is ambitious*, by A4. So, we can infer 2:

2   'Caesar is ambitious' means *Caesar is ambitious*.

Now, let us add axioms for connectives, which we will treat as having meanings which take us from meanings of sentences or formulae to meanings of sentences or formulae. An axiom for negation and for conjunction will suffice for the purposes of illustration ('P' and 'Q' are to play the role of 'x' above).

A5   Means('P and Q', *P and Q*)
A6   Means('~ S', *~ S*)

A7   For any binary sentential connective Δ, and any formulae ϕ, ψ, the result of placing ϕ and ψ in the first and second argument places of Δ means the value of the meaning of Δ given the meaning of ϕ and of ψ as first and second arguments.

A8   For any unary sentential connective Δ, any formula ϕ, the result of placing ϕ in the argument place of Δ means the value of Δ given the meaning of ϕ as argument.

A9   The value of any sentential connective for a sequence of arguments denoted by a sequence of formulae is denoted by the expression that results from placing the formulae sequentially in the argument places of the connective.

Instantiate A8 to '~' and 'Caesar is ambitious' to get 3:

3   '~ Caesar is ambitious' means the value of the meaning of '~' given the meaning of 'Caesar is ambitious' as argument.

With A9, this gives us 4,

4   '~Caesar is ambitious' means *~Caesar is ambitious*.

Now let's introduce an axiom for a universal quantifier:

A10   Means('For all x: F', *For all x: F*)

A11   For any unary quantifier Q, any formula ϕ, the result of placing ϕ in the argument place of Q means the value of the meaning of Q given the meaning of ϕ as argument.

A12   The value of the meaning of any unary quantifier for an argument denoted by a formula is denoted by the expression that results from placing the formula in the argument place of the quantifier.

Instantiate A11 to 'For all x: F' and '~x is ambitious' to get:

5   'For all x: ~x is ambitious' means the value of 'For all x: F' given the meaning of '~x is ambitious' as argument.

From A12 we get 6:

6   'For all x: ~x is ambitious' means *For all x: ~x is ambitious*.

This generalizes to relational predicates and multiple quantifiers. Every expression is assigned a meaning, and we can produce for any complex expression an assignment of meaning that seems intuitively to give the right result.

The theory can be generalized so that it is given in a language that does not embed the object language. We need to modify A4, A9, and A12 in this case. We indicate the modification to A4 as a guide.

A4′   The value of the meaning of any sentential function for an argument denoted by a referring term is denoted by any expression that results from placing a term that refers to the argument in the argument place of a term that refers to the meaning of the sentential function.

This enables us then to use axioms of the form of A1 and A2 to produce a term in the metalanguage that refers to the meaning of the sentence 'Caesar is ambitious' in the object language so as to produce a theorem of the form of 2 with a metalanguage sentence on the right-hand side even when the metalanguage does not embed the object language. Thus, for example, we might replace A1 and A2 with A1′ and A2′ (understanding 'Means' and other semantic predicates relative to the object language, regimented French in this case):

A1′   Means('César', *Caesar*)
A2′   Means('*x* est ambitieux', *x is ambitious*).

4   The Inutility of Meanings

Thus a meaning theory that works by way, it seems, of assigning meanings to every expression of the language and which exhibits the meanings of complex expressions as a function of the meanings of their parts and mode of combination. It would appear that the appeal to meanings as entities in the theory of meaning has been vindicated.

But this is an illusion. First, the appeal to meanings is not what does the work. Second, what work is done could be done (almost) as well by a translation theory. Third, a theory that does just as well is easily constructed exploiting exactly the features of the original which were important for its fulfilling its purpose, without, however, any quantification over meanings.

To show that the appeal to meanings is not doing the work, we alter the base clauses of the theory. The base clauses still assign meanings to entities systematically, and still allow us to say systemactically what each expression of the language, simple and complex, means, that is to say, what meaning each expression has, when we understand this as what meaning entity is assigned. But it does this in a way that provides no insight into how to understand any expression of the language. The alterations are as follows:

A1*   Means('Caesar', *Caesar*)
A2*   Means('*x* is ambitious', *Brutus*)

A5* Means('P and Q', *Cassius*)
A6* Means('~ S', *Antony*)
A10* Means('For all *x*: F', *Portia*).

Here 'Brutus', 'Cassius', etc. name meanings (not individuals); 'Brutus' is a name for the meaning of '*x* is ambitious', 'Cassius' for the meaning of 'P and Q', and so on. Now, with A3, A8, and A11, we can derive 7–9:

7   'Caesar is ambitious' means the value of Brutus given Caesar as argument.
8   '~ Caesar is ambitious' means the value of Antony given the value of Brutus given Caesar as argument.
9   'For all *x*: ~*x* is ambitious' means the value of Portia given as argument the value of Antony given Brutus as argument.

It is clear that 7–9 do not enable us to understand any of the object-language expressions. Yet, one cannot fault them for failing to tell us, as least as well as the original theory, what each expression of the object language means, in the sense of telling us what meaning entity is assigned to it.

What is the crucial difference? In our original theory, we assigned meanings to primitive expressions in our object language by using expressions in the metalanguage which were in the same grammatical category, and which it was tacitly assumed were translations of the object-language terms. This is what enabled us to understand the object-language primitive expressions. A4, A9, and A12 then told us that certain forms of complex expressions in the metalanguage which we already understand had the same meaning as certain corresponding expressions in the object language, thereby allowing us to match systematically complex object-language expressions with complex metalanguage expressions alike in meaning, i.e., synonymous with them. But the key to our being able to come to know what the complex object-language expressions meant lay in our already understanding the metalanguage expressions in the sense in which we wished to understand the object-language expressions and being given information sufficient to know which metalanguage expression was synonymous with which object-language expression. Being told what meanings, construed as entities, each expression is to be assigned is not sufficient. Being given a way to match an object-language expression with an already-understood metalanguage expression that we know is synonymous with it was. The assignment of entities to expressions, which was to be the key to a theory of meaning, turns out to have been merely a way of matching object-language expressions with metalanguage expressions thought of as used (in referring to their own meanings), so that we are given an object-language expression and a matched metalanguage expression we understand, in a context which ensures that they are synonymous.

That nothing more is involved can be shown by noting that the original theory gives all the information we need to write out a shorter recursive translation theory. To increase interest, let us shift to a regimented fragment of French as the object language. (We use brackets for Quinean corner quotes.)

T1   Translates('César', 'Caesar') [cf. A1]
T2   Translates('x est ambitieux', 'x is ambitious') [cf. A2]
T3   For any P, α, translates([α P], trans(α)⁀trans(P)). [cf. A3–4]
T5   For any P, Q, translates([P et Q], trans(P)⁀'and'⁀trans(Q)])
T6   For any S, translates([~ S], '~'⁀trans(S)) [cf. A7–9]
T7   For any F, translates([Chaque x: F], 'For all x: '⁀trans(F)]) [cf. A10–12]

In the above, 'translates' is a two-place predicate relating an expression to an expression that translates it (we suppress explicit relativization to the two languages here). 'Trans(x)' is a function that yields the translation of the object-language expression x into the metalanguage, where for all y, trans(x) = y iff translates(x, y). Given knowledge of the metalanguage, this provides as much information as our original theory. It might be said that there is still a crucial difference, however, namely, that what the original says is sufficient whereas what the translation theory says is not. But this difference has to do only with the fact that using expressions to refer to their own meanings makes it look as if to understand the meaning statements we must understand the used expressions, so that understanding the theory gives us knowledge of the meanings of the relevant translations into the metalanguage of the object-language sentences. However, in effect, the one grip we have on what entity is supposed to be associated with an object-language expression is by way of a description we can construct using the metalanguage expression, namely, in the case of, e.g., A2, 'the meaning of "is ambitious"'. Reference to the meaning, whatever it is, is beside the point. We think of the meaning in this way, but we understand the expression we use to pick it out, and that understanding and the assumption of synonymy are what enable us to understand the object-language expression.

With a minor modification to our translation theory, we can duplicate this effect of using the metalanguage expression without appeal to meaning entities. For this purpose, we replace 'translates' with 'means'. But we do not interpret 'means' as a predicate relating an expression and a meaning, but as a predicate relating an expression and another expression, which, however, we require someone to understand in order to understand the sentence. We write this, for example, as follows:

'est ambitieux' means *is ambitious*:

This is true just in case 'est ambitieux' in its language is translated by 'is ambitious' in the language of the sentence. The requirement that one understand

'is ambitious' to understand the sentence is simply an additional convention governing its use. Thus, we have the effect of using 'is ambitious' in the sentence, though its usual extensional properties are irrelevant to the truth of the containing sentence.

Now we rewrite the translation theory with some additional axioms to provide a meaning theory as follows:

T1   Means('César', Caesar)
T2   Means('x est ambitieux', x is ambitious)
T3   For any P, α, means([α P], means(α)ˆmeans(P)).
T5   For any P, Q, means([P et Q], means(P)ˆˋ and 'ˆmeans(Q)])
T6   For any S, means([~ S], '~'ˆmeans(S))
T7   For any F, means([Chaque x: F], 'For all x:'ˆmeans(F)]).

Where 'means' has two arguments, it is the relational term; where it has one, it is a function yielding as value the metalanguage expression that translates the object-language expression which is its argument. We add that the convention for substituting for a variable in the context following 'means' is to substitute the expression which is the value of the variable without quotation marks but in italics (note that in T3–T7 we have descriptions in the second argument place for 'means', and thus quantifiers binding the variable there). This signals the convention that the expression is both mentioned and understood. (Consider the standard use of quotation marks for dialogue, where a similar convention seems to hold.) This theory allows us to infer, e.g.,

'Chaque x: ~x est ambitieux' means *for all x: ~x is ambitious*

and so on. The features of our original theory which quantified over meanings which enabled it to serve the purposes of helping us to understand object-language sentences have been preserved, but without any commitment to meanings.

5   Conclusions

The first conclusion to draw from this is that quantification over meanings in the theory of meaning serves no real purpose. The introduction of meanings as entities is not sufficient to enable us to understand object-language sentences. The introduction of meanings as entities is not necessary. The illusion that they are helpful is generated by choosing terms that refer to meanings in the metalanguage in a way that matches object-language terms with metalanguage terms in the same semantic category which are understood to be translations of the object-language terms. The use of them to refer to meanings gets them out of quotation marks so that we have the illusion that

in understanding the theory we must understand the terms that refer to the meanings. Thus the theory seems to state something knowledge of which is sufficient for understanding object-language terms. But it is not much different from a translation theory together with knowledge of one of the languages. And the crucial elements, matching object-language sentences with metalanguage sentences that translate them, but using the sentences in a way that requires understanding them in understanding the theory, can be replicated without the pointless quantification over meanings.

One thing this shows is that to give a compositional meaning theory for a language, it is necessary to have a metalanguage we understand which has the same expressive resources as the object language. A theory that aims to produce from statements about primitives and the complexes we can grammatically form from them what every expression means in a way that ensures that understanding the theory helps us understand the language requires basically that we match object-language sentences with metalanguage sentences 'in use' in a way that enables us to see that the metalanguage sentences are translations of the object-language sentences. There is no way to state what an object-language expression means (as opposed to explaining it, which we do when explaining the meaning of an expression to someone who knows no expression with that meaning) without using terms the same in meaning as it, and so understood. The understanding is achieved not by sudden grasp of an abstract object associated with the expression, but by showing what it means by using an expression the same in meaning as it.

This is how the project of giving a theory of meaning by way of an interpretive truth theory works (see Ludwig, 2002). We use axioms in the truth theory which employ terms in the metalanguage which are translations of them, or provide, relative to contextual parameters, the contribution of the expression in use to what is said in using it. The 'minimal' proof of a T-sentence then provides a sentence in the metalanguage we can know to be synonymous with the object-language sentence. The truth theory uses the metalanguage expressions used to give truth conditions, ensuring that we are, in light of our knowledge about the truth theory, in a position to interpret the object-language sentence (perhaps relative to an occasion of use for a context-sensitive language). But the use of a truth theory has an additional advantage. For the proof of the T-sentence from assignments of reference and satisfaction conditions to primitive expressions, non-recursive and recursive, exhibits how the expressions in the object-language sentence contribute, through their meaning, mirrored in the metalanguage expressions, to the extensional properties of the terms relevant to determining whether the sentence is true or false. We gain genuine insight in this way into the compositional mechanisms of the language, namely, how truth conditions for sentences are built up compositionally out of extensional properties of their component expressions and their mode of combination. Our sample meaning theory above (the one

dispensing with meanings), while it matches object with synonymous meta-language sentences, does not exhibit any information about the systematic contribution of expressions in the object language to conditions under which they are true. It is not clear how else to get this kind of illumination than by way of an interpretive truth theory. The truth theory does not state any of this, of course; it rather shows it. If we are right, there is no other route to the same effect. Thus, as we said at the outset, there is an interesting sense in which facts about how our language represents the world cannot be stated but can only be shown.

Rutgers University, USA and University of Florida, USA

References

Davidson, D. (2001 (1967)) 'Truth and Meaning', in *Inquiries into Truth and Interpretation,* 2nd edn, New York: Clarendon Press, pp. 17–36. Original publication: *Synthese* 17 (1967): 304–23.

Ludwig, K. (2002) 'What is the Role of a Truth Theory in a Meaning Theory?', in D. Shier, J. K. Campbell, and M. O'Rourke (eds) *Meaning and Truth: Investigations in Philosophical Semantics,* New York: Seven Bridges Press, pp. 142–63.

Quine, W. V. O. (1953) 'Two Dogmas of Empiricism,' in *From a Logical Point of View,* 2nd edn, Cambridge, Mass.: Harvard University Press.

—— (1960) *Word and Object,* Cambridge, Mass.: MIT Press.

Wittgenstein, L. (1950) *Philosophical Investigations,* London: Macmillan.

—— (1961) *Tractatus Logico-Philosophicus,* London: Routledge & Kegan Paul.