Metametaphysics and Semantics

(draft of paper to appear in *Metaphilosophy*)

Timothy Williamson

Faculty of Philosophy, University of Oxford

New College, Oxford OX1 3BN, U.K.

timothy.williamson@philosophy.ox.ac.uk

Abstract: Metaphysics faces a threat from apparently metaphysics-friendly non-epistemic forms of semantics, on which sentences express 'worldly' propositions, e.g. functions from worlds to truth-values. The threat goes back to Wittgenstein's *Tractatus Logico-Philosophicus* and is pressed in different forms by various contemporary philosophers. It is that metaphysical claims turn out either trivially true or trivially false, because they express the same proposition as a tautology or contradiction. The problem is shown to generalize to accounts on which sentences express Russellian structured propositions. It applies to logic and mathematics as well as metaphysics. Attempts to solve it by reinterpreting apparently non-contingent claims as contingent metalinguistic claims or by invoking Fregean semantics are shown to fail. The underlying problem concerns necessary equivalence, not necessary truth, and arises in all domains. To solve it, we must recognize that the form of our representations plays an ineliminable cognitive role which cannot be reduced to their content.

Keywords: Guises, hyperintensional, intensional semantics, metaphysics, Russellian propositions, Stalnaker, triviality.

Of all branches of philosophy, metaphysics has probably attracted the most opprobrium. It is the one most easily represented as a lazy, dogmatic, obsolete rival of natural science. It is also the most discursively abstract branch. Predictably, it is the one most often accused of being *nonsense*.[1]

When meaning is understood in epistemic terms, the charge of meaninglessness turns into the charge that metaphysics is epistemically inadequate: it lacks proper methods for achieving knowledge, or even reasonable belief, in its domain. The logical positivists gave a salient version of such a critique, wielding their verification principle as a blunt instrument. Since putative truths of metaphysics are neither analytic nor empirically verifiable, they are meaningless by logical empiricist standards. Of course, such an accusation is largely bluff without an adequate verificationist theory of meaning in the background, and the logical positivists made very little progress towards developing such a theory. Nevertheless, the logical empiricist dichotomy of all cognition into 'empirical' and 'conceptual' aspects continues to have its adherents: for example, the work of Amie Thomasson (2015, 2020) is in the tradition of Rudolf Carnap (1950), and more distantly of David Hume's dichotomy of 'relations of ideas' and 'matters of fact', though she has more interest than Carnap in non-scientific language. Between the conceptual and the empirical, no room seems left for substantive unconfused metaphysical theorizing.

Unfortunately, like 'analytic' and 'synthetic', the terms 'conceptual' and 'empirical' are far more problematic than they first appear. There is a crude stereotype of the conceptual, and a crude stereotype of the empirical, but the assumptions built into those stereotypes are unclear. What is clear is that both stereotypes, separately and together, are utterly inadequate for making sense of logic and mathematics, let alone of metaphysics. I have explained the difficulties elsewhere (most recently, Williamson 2007, 202X) and will not repeat those considerations here.

This article addresses a different challenge to metaphysics. It is more urgent, because its starting-point is less hostile. The new challenge is semantic, like the logical empiricist critique, but unlike the latter it does not depend on an epistemic conception of semantics. Instead, one might even say, it depends on a *metaphysical* conception of semantics. But that does not make the new challenge self-defeating. For if metaphysics is already in tension with a metaphysics-friendly approach to semantics, that is bad news for metaphysics.

Uncompromising metaphysics, both ancient and modern, aspires to discover the necessary nature and structure of reality. Its primary interest is in the world, not in our thought or talk about the world—of course, our thought and talk are part of the world, but (except under extreme forms of idealism) only a very small part. Thus a worldly approach to semantics, on which the semantic value of a linguistic expression in a context is a worldly item, looks like a good fit with metaphysics. For example, such a theory may identify the semantic value of a declarative sentence in a context with a proposition, understood as the set of possible worlds at which the sentence is true, or as a complex of the objects, properties, and relations the sentence is about (in both cases, relative to that context). In brief, such semantics correlates metaphysicians' words with the very metaphysical entities they wish to discuss (if there are such entities). That suggests a fully cooperative attitude of semantics to metaphysics. Any tension between the two is therefore all the more disturbing—as though semantics, with the best will in the world, still leaves no room for metaphysics.

1. *Intensional semantics*

The problem arises in an especially stark form within a standard framework for intensional semantics, the mainstream of contemporary formal semantics as a branch of linguistics. The

approach is referential and truth-conditional. We consider it first in its simplest form. It has been elaborated and modified in various ways, but they turn out to make no essential difference to the problem.

Each expression of the language is assigned a *content* (henceforth, relativization to a context of utterance will be left implicit). The assignment is compositional: the content of a complex expression is determined by the contents of its constituent expressions and how those constituents are put together. This determination is implemented within a framework of *worlds*, treated as parameters of semantic evaluation. The semantics is *intensional* because, for many central types of expression, including predicates and (declarative) sentences, the content of an expression is its intension, a function mapping each world to its extension at that world. In particular, since the extension of a sentence at a world is its truth-value at that world, its intension is a function from worlds to truth-values (truth or falsity). We may call such sentential intensions *propositions*. Thus the sentence 'There is a god' expresses the proposition that there is a god, the function mapping each world in which there is a god to truth and each world in which there is no god to falsity.

Any such intensional framework determines a *distinguished modality*, characterized by the condition that, for any sentence '*A*', 'Necessarily *A*' is true at a world *w* if and only if '*A*' is true at every world, while 'Possibly *A*' is true at *w* if and only if '*A*' is true at some world. In what follows, the words 'possible' and 'necessary' will be used for that modality. Thus the worlds in the framework are all and only the *possible* worlds. On the compositional semantics, the logical connectives behave classically at each world, so any truth of classical propositional logic is true at every world. It is then easily shown that every theorem of the well-known modal system S5 is true at every world on this interpretation. In particular, it validates the distinctive theses that whatever is necessary is necessarily necessary (the S4

axiom) and whatever is possible is necessarily possible (the S5 axiom): matters of necessity or possibility are not themselves contingent.

Since our interest is in metaphysically-oriented semantic theories, we should interpret this distinguished modality as broadly objective rather than merely epistemic in nature. Indeed, we may conceive it as the broadest kind of objective possibility, since it excludes no worlds in the framework. An attractive hypothesis is that such a maximal objective modality is just what is usually called 'metaphysical modality' (Williamson 2016). But that is far from uncontroversial. For instance, the S4 axiom has been denied for metaphysical modality (Salmon 1989), and the S5 axiom has been denied for the broadest modality (Bacon 2018). A further complication is that a deeper understanding of modality may well not treat possible worlds as basic, but instead start from the distinction between the possible and the impossible itself, perhaps in the setting of higher-order logic (Williamson 2013). However, such alternatives are all compatible with some version of intensional semantics; none of them avoids the problems discussed below. For present purposes, we can ignore the differences between them.

An immediate corollary of this approach is that propositions are very coarse-grained. *Necessarily equivalent propositions are identical*: they output the same truth-value for any given world as input, and so are the same function. Thus, in particular, there is only one necessary proposition and only one impossible proposition. If you know one necessary truth, you know them all. For instance, reading 'god' in a strong sense, on which being a god is a necessary property: whatever has it in a world has it in any other world too (it may also entail other standard attributes, such as omniscience, omnipotence, omnipresence, omnibenevolence, and eternity; for present purposes we omit ineffability, since it might cause distinctively semantic problems). Thus it is either necessary or impossible that there is a god. If it is necessary, the proposition that there is a god is just the proposition that all cats are

cats. If it is impossible, the proposition that there is a god is just the proposition that some cats are not cats. So, on this view, when atheists argue with theists, the two propositions in dispute are that all cats are cats and that some cats are not cats, one way round or the other. Surely such a dispute is a waste of time. The moral seems to be: insofar as metaphysics concerns the non-contingent, intensionalism *trivializes* metaphysics.

Unlike empiricist and logical positivist critiques of metaphysics, the argument from intensionalism has no epistemological premises, and its conclusion is not distinctively epistemological; the argument is just semantic. Nevertheless, it reaches a similar conclusion: there is nothing non-trivial for metaphysical claims to mean. Such arguments have had significant influence. They can be traced back to Wittgenstein's *Tractatus Logico-Philosophicus*, where ultimately every declarative sentence is to be analysed as a truth-function of atomic sentences expressing simple, mutually independent states of affairs. If it is true on every assignment of truth-values to those atomic sentences, it is merely tautologous. If it is false on every assignment, it is merely contradictory. If it is true on some assignments and false on others, it is merely contingent. This taxonomy leaves nowhere for metaphysics to hide. A conception of impossibilities as trivially false may explain the claim, widespread even amongst contemporary Wittgensteinians, that it is meaningless to assert an impossibility. In contemporary philosophy, Robert Stalnaker has been a leader in pressing the radical consequences of intensionalism, though with a scaffolding of possible worlds rather than simple, mutually independent states of affairs (1984, 1999).[2] Such intensionalist sympathies can also be found in the works of David Lewis (1970, 1996) and, in less committed form, Saul Kripke (1979), despite their major contributions to metaphysics. More recently, Eli Hirsch (2021) has extended his nuanced semantic critique of (some) metaphysics by connecting it with the kinds of coarse-grained, worldly, semantics, including intensionalism, which look friendly to out-and-out metaphysics.

Of course, intensionalist trivialization threatens more than metaphysics. It concerns any inquiry into the non-contingent. Logic and mathematics are salient examples. They can hardly be dismissed as trivial. If the proof of Fermat's Last Theorem was just a proof that all cats are cats, why did it take centuries to find? However, many philosophers find less difficulty in convincing themselves that logic and mathematics are somehow purely formal, not really concerned with *how the world is*, so not in need of non-trivial content. By contrast, traditional metaphysics stubbornly enquires into the necessary nature of the world; for it, the threat that only triviality that way lies is existential. In what follows, the focus will be on metaphysics, not on logic and mathematics as such, but the conclusions will apply to the latter too, providing a way for them to be as worldly as metaphysics, with which they indeed overlap (Williamson 2013).

2. *Generalizing the problem*

How robust are the trivializing consequences of intensionalism? Do they survive motivated generalizations of the intensional framework?

First indications offer metaphysics little hope. For instance, many versions of intensional semantics add a parameter for *times* to that for *worlds* in semantic evaluation, to handle tense. Then sentences express the same content if and only if they have the same truth-value at every world-time pair. But that makes no significant difference to the problem. On the operative reading of the word 'god', we may assume, the property of being a god is eternal as well as necessary: something is a god at a world and time if and only if it is a god at *every* world and time. Hence either there is a god at every world and time or there is a god at

no world and time. Thus, as before, ‘There is a god’ has the same content as either ‘All cats are cats’ or ‘Some cats are not cats’.

A more far-reaching modification of the framework is to work with *possible situations* instead of *possible worlds*, to handle the locality of much discourse (for instance, Elbourne 2005). Situations are something like parts of worlds. A sentence is neither true nor false in situations which include too little to determine its truth-value. Presumably, then, sentences express the same content if and only if they have the same truth-value (if any) in every situation. But that still makes no crucial difference to the problem. For on the operative reading of the word ‘god’, we may assume, something is a god in a situation if and only if it is a god in *every* situation (a form of necessary omnipresence). Hence either there is a god in every situation or there is a god in no situation. If a situation *s* has a god in it, ‘There is a god’ is true in *s*. If *s* has no god in it, there must be no god, so ‘There is a god’ is false in *s*. Thus if ‘All cats are cats’ is true and ‘Some cats are not cats’ false in every situation, ‘There is a god’ still has the same content as either ‘All cats are cats’ or ‘Some cats are not cats’. There is a slight complication: some versions of situation semantics may determine no truth-value for those ‘cat’ sentences in situations which exclude some cats. By contrast, ‘There is a god’ is true or false in such situations, as just explained. In that case, ‘There is a god’ *differs* in content from both ‘All cats are cats’ and ‘Some cats are not cats’. But this technicality will not solve the problem. For we can introduce a logically constant sentence ⊥ for absurdity, governed by the stipulation that ⊥ is false, and so its negation ¬⊥ true, in each situation. Then ‘There is a god’ has the same content as either the trivially false ⊥ or the trivially true ¬⊥.

A more radical strategy is to allow *impossible worlds*, understood in an ontologically harmless way as arbitrary sets of sentences of the object language. A sentence is evaluated as true at such a world if and only if it is a member of that world. To use this apparatus to individuate content more finely, we can stipulate that sentences express the same content if

and only if they are true at the same worlds, possible and impossible. Then ‘There is a god’ differs in content from any other sentence *S*, for the simple reason that ‘There is a god’ is true at the world {‘There is a god’}, while *S* is not true at that world. But this strategy trivializes sameness of content by reducing it to sameness of sentence. For even if the words ‘god’ and ‘deity’ are *synonyms* by normal standards, the sentences ‘There is a god’ and ‘There is a deity’ still count as differing in content, for the reason just given (let *S* = ‘There is a deity’). No such verbal manoeuvre will rescue the ambitions of traditional metaphysics.

A more hopeful-looking move is to abandon the identification of sentential contents with (perhaps partial) functions from circumstances of evaluation to truth-values, and adopt a more structured conception instead. In particular, one might identify the content of a declarative sentence with a *Russellian proposition*, a complex built out of the objects, properties, and relations the sentence is about, and structured according to the structure of the sentence. For example, the proposition that there is a god might be something like <∃, divinity>, the ordered pair of the second-order property ∃ of being instantiated and the first-order property divinity, of being a god. Then a proposition *p* is the proposition that there is a god only if *p* has divinity as a constituent.

An immediate concern is that the individuation of Russellian propositions is itself hostage to the individuation of properties and relations. In particular, suppose that properties are identical if and only if they are necessarily coextensive. Then if it is in fact impossible to be a god, the property of being a god is necessarily coextensive, and so identical, with the property of being a round square; thus the proposition <∃, divinity> is just the proposition <∃, round-squarehood>, and the threat of trivialization returns. So far, this is just an isolated case; there is no such elementary argument on the alternative hypothesis that it is possible to be a god.

However, we can develop a much more general threat of trivialization for Russellian propositions. We keep ‘There is a god’ as our sample sentence of metaphysics, but without exploiting its specific details. We introduce a new singular term ‘D’, governed by this stipulation:

If there is a god, ‘D’ names 1.
If there is no god, ‘D’ names 0.

The stipulation is to be understood as belonging to the metasemantics of ‘D’, not to its semantics, in Kripkean terms, to *fix the reference* of ‘D’, not to *give its meaning* (Kripke 1980). Thus ‘D’ is not to be understood as abbreviating a definite description like ‘the number *n* such that either there is a god and *n* = 1 or there is no god and *n* = 0’. Rather, ‘D’ is simply a name of a natural number; the stipulation specifies which number. Consequently, the Russellian proposition semantically expressed by the equation ‘D = 1’ has none of the complex structure of the definite description, but is simply something like <identity, <D, 1>>. Obviously, the sentences ‘There is a god’ and ‘There is no god’ have the same truth-values as the equations ‘D = 1’ and ‘D = 0’ respectively. If we want to argue about whether there is a god, we can argue about whether D = 1; it makes no dialectical difference.

Of course, whichever side is wrong about the metaphysics also has a false belief about the reference of ‘D’, given that they have been introduced to the name by the stipulation above. If there is a god, atheists falsely believe that ‘D’ names 0; if there is no god, theists falsely believe that ‘D’ names 1. But that does not mean that one side or the other *misunderstands* the name ‘D’. It is like the name ‘Jack the Ripper’, introduced by the description ‘whoever committed the grisly Whitechapel murders’. Some people may still

falsely believe the wild theory that Edward VII committed the grisly Whitechapel murders, so that Jack the Ripper was Edward VII; familiar with the name 'Jack the Ripper' in the usual way, they falsely believe that it names Edward VII, but they do not thereby *misunderstand* the name 'Jack the Ripper'.[3]

On the Russellian view, if there is a god, the sentence 'D = 1' expresses the same proposition as the trivially true sentence '1 = 1'; if there is no god, the sentence 'D = 0' expresses the same proposition as the trivially true sentence '0 = 0' (on the intensional view, the corresponding necessary propositions are identical too).

The threat of trivialization has returned in completely general form. One could substitute any other sentence for 'There is a god' in the preceding argument. Nothing here depends even on the non-contingency of 'There is a god'. The argument works in the same way if one substitutes 'There is intelligent life in other galaxies' for 'There is a god':

If there is intelligent life in other galaxies, 'G' names 1.
If there is no intelligent life in other galaxies, 'G' names 0.

Everything proceeds as with 'D'. In particular, since 'G' is a proper name, it is a rigid designator, even though it is contingent whether there is intelligent life in other galaxies. Thus, if there is intelligent life in other galaxies, we use 'G' to designate 1 even with respect to counterfactual possibilities in which there is no intelligent life in other galaxies. Equally, if there is no intelligent life in other galaxies, we use 'G' to designate 0 even with respect to counterfactual possibilities in which there is intelligent life in other galaxies. Thus, we are in an epistemic position to assert both 'There is intelligent life in other galaxies if and only if G = 1' and 'Either there could have been no intelligent life on other planets while G was 1 or there could have been intelligent life on other planets while G was 0'. The biconditional is

similar to proposed examples of contingent *a priori* truths (Kripke 1980). If we want to argue about whether there is intelligent life in other galaxies, we can argue about whether G = 1; it makes no dialectical difference. If there is intelligent life in other galaxies, the sentence 'G = 1' expresses the same Russellian proposition as the trivially true sentence '1 = 1'. If there is no intelligent life in other galaxies, the sentence 'G = 0' expresses the same Russellian proposition as the trivially true sentence '0 = 0'. Again, in both cases, the corresponding necessary propositions in the intensional framework are identical too.

Such examples cast doubt on any attempt to interpret the semantic considerations as revealing some pathology of metaphysics, for the question whether there is intelligent life in other galaxies is uncontentiously non-pathological.

Related examples occur quite naturally, with no need of artificial stipulations. For instance, the terms 'furze' and 'gorse' are simply two natural kind terms for the very same genus of thorny shrub (the example is in Kripke 1979). There is no semantic difference between them in English. Yet by normal standards someone could understand both terms without recognizing that they co-refer. Perhaps at one time in one place you were shown a green bush with yellow flowers and told 'That's furze', while at another time in another place you were shown a brown bush with no flowers and told 'That's gorse'. Those are both adequate ostensive definitions by normal standards. You might well not realize that the differences were largely seasonal, and that you had been introduced to the same genus twice. Of course, *you* think of 'furze' as green with yellow flowers and 'gorse' as brown with no flowers, but that is purely idiosyncratic. Someone else may think of 'furze' as brown with no flowers and 'gorse' as green with yellow flowers. As words of English, they are synonyms. On the Russellian approach (if not that of the historical Russell), the English sentence 'Furze is gorse' expresses the Russellian proposition <identity, <furze, gorse>>, which just is the obviously true Russellian proposition <identity, <furze, furze>>. Yet you could sensibly ask

yourself ‘Is furze gorse?’ out of simple non-pathological botanical interest. A similar issue arises on a natural implementation of the intensional approach, since ‘furze’ and ‘gorse’ are rigid designators of the same genus, so ‘Furze is gorse’ is true at all possible worlds.[4]

The evidence so far supports at least two conclusions. First, for the problem of trivialization, it makes little difference whether we adopt intensionalism or some sort of Russellian hyperintensionalism. Second, the problem is generic; it shows nothing distinctive about specific forms of enquiry. In particular, it shows nothing pathological about metaphysics. Nor does it show anything special about logic or mathematics.

3. The metalinguistic strategy[5]

Some philosophers are still tempted by the idea that the ignorance or at least non-triviality displayed in the cases described is fundamentally semantic, that there are serious obstacles to knowing the semantic values of some of the words or sentences in play. For instance, it is hard to know *which* numbers the names ‘D’ and ‘G’ designate. This ignorance would be of a familiar, unpuzzling kind, and pose no threat to the favoured semantic framework.

Consider ‘furze’ and ‘gorse’. The obvious line for proponents of the metalinguistic strategy is to insist that anyone—such as an expert botanist—with full, non-deferential understanding of both ‘furze’ and ‘gorse’ *is* in a position to know that they co-refer. Everyone else has at most partial understanding of at least one of the two terms. Thus the problem is fundamentally one of semantic ignorance.

Such an account may apply to this particular case, though what the ‘full understanding’ might be with which ‘partial understanding’ is implicitly contrasted is far from clear. In any case, we can vary the example. In one variant, set many centuries ago, the

shrub in question is rare and grows only in remote places. It has been seen only occasionally, but never studied scientifically, and no specimens have been observed over extended periods. The term ‘furze’ was introduced by travellers who saw green bushes with yellow flowers, and in practice only bushes in that condition are recognized as ‘furze’. Similarly, the term ‘gorse’ was introduced by travellers who saw brown bushes with no flowers, and in practice only bushes in that condition are recognized as ‘gorse’. Not even the best botanists in our community realize that ‘furze’ and ‘gorse’ co-refer; they may regard it as an open question. Nevertheless, despite the community-wide difference between ‘furze’ and ‘gorse’ in associated recognitional capacities, there is no strictly *semantic* difference between the two words. They are both simply natural kind terms for what is in fact the very same natural kind. In that sense, they are synonyms. In these circumstances, a fully non-deferential understanding of both terms does not put one in a position to recognize their co-reference. To resolve our ignorance, our primary need is to know more botany, not more semantics.

In another variant, everything is like the original case, but without deference, since the community does not recognize the status of scientific expertise. Instead, natural kind terms are treated more as words like ‘if’ and ‘know’ are actually treated. Although some people devote themselves to studying conditionals or knowledge, they play no privileged role in the social practice of using the corresponding words, because the community has no tendency to defer to them in applying the words. Similarly, in the imagined case, even if some people devote themselves to studying shrubs, they play no privileged role in the social practice of using the words ‘furze’ and ‘gorse’, because the community has no tendency to defer to them in applying the words. At least for those who have been introduced to them ostensively, the ethos in applying them is that everyone is entitled to their own opinion. For that large, at least minimally competent group, there is no deferential partial understanding, because there is no deference to a higher level of competence. ‘Furze’ and ‘gorse’ are still treated as natural kind

terms, but in an unscientific spirit. Many speakers fully competent by communal standards with both terms cannot recognize that they co-refer. Unlike the previous case, there is no community-wide difference between ‘furze’ and ‘gorse’ in associated recognitional capacities; such differences obtain only at the level of individual speakers. There is also no strictly semantic difference between the two words. They are both simply natural kind terms for what is in fact the very same natural kind. They are synonyms. In these circumstances too, a fully non-deferential understanding of both terms does not always put one in a position to recognize their co-reference. To resolve one’s ignorance, one’s primary need is to know more botany, not more semantics.

At first sight, the artificially introduced names ‘D’ and ‘G’ look more promising as candidates for semantic ignorance. The associated stipulation might plausibly be denied by itself to enable one to know *which* number ‘D’ or ‘G’ designates. Currently, someone familiar with the stipulation may be uncertain whether ‘D’ co-refers with ‘1’ or with ‘0’. But that is because they are uncertain whether there is a god: the semantic ignorance seems to depend on prior metaphysical ignorance, contrary to the metalinguistic strategy. However, proponents of the strategy may respond that the relevant semantic ignorance is at the level of the *sentence*, not the singular terms: the underlying uncertainty is as to *which proposition* the sentence ‘There is a god’ expresses. This may seem more promising. Semantic ignorance of individual words in the sentence would be mere linguistic incompetence, which is an implausible diagnosis of the problems of metaphysics. Of course, the word ‘god’ is hardly straightforward, but for present purposes we may assume that it has been stipulatively defined in terms of a list of attributes. The picture is that we know the semantic values of the atomic constituents of the sentence, but cannot work out which proposition results from composing them in the relevant way.

Such an account makes more sense for intensional than for Russellian propositions. For the latter, if we know that a sentence is composed of a constituent expressing the second-order property ∃ predicated of a constituent expressing the first-order property of divinity, we can easily work out that the sentence as a whole expresses the Russellian proposition <∃, divinity>: there is no mystery as to which proposition that is, because the notation is already so perspicuous.[6] By contrast, if the proposition is a function from worlds to truth-values, but one is uncertain whether it maps all worlds to truth or all to falsity, one might well be counted uncertain as to which function the sentence expresses. If sentences are individuated syntactically, not semantically, it is contingent which proposition a sentence expresses, so the apparent metaphysical uncertainty has finally been traced to uncertainty about something contingent.

However, the proposal does not withstand further scrutiny. Let '*S*' abbreviate whichever is false of the quotations 'There is a god' and 'There is no god'. Thus *S* is the false one of those two sentences. Whatever proposition *S* semantically expresses is impossible. Consider a metaphysician uncertain whether *S* expresses a true proposition. Of course, that uncertainty is uninteresting if it results from lack of native speaker knowledge of English. We must assume our metaphysician to know what the words and modes of composition in *S* mean. Thus we assume that she knows that *S* has semantic features *F*, fully characterizing the semantics of *S*'s atomic constituents and modes of composition. Consequently, since our metaphysician is rational, she is also uncertain over the conjunction that *S* both has *F* and expresses a true proposition. But the conjunction is itself impossible, for since the semantics is compositional, a necessary consequence of the first conjunct (that *S* has *F*) is that the proposition *S* expresses is impossible, which is incompatible with the second conjunct. Each conjunct is possible, but they are not compossible. Thus our metaphysician's uncertainty

extends to something impossible, contrary to the metalinguistic strategy of confining the relevant uncertainty, ignorance, or error to contingent linguistic matters.

A back-up tactic for the metalinguistic strategy is to divide an agent's beliefs into separate subsystems, individually possible but jointly incompossible (Stalnaker 1984). However, in the present case, separating the metaphysician's belief that *S* expresses a truth from her belief that *S* has *F* misses the depth of the problem. She does not have to *ignore* her understanding of *S* in order to believe that it expresses a truth; she believes that *S* is true *in the light of* her understanding of *S*. Positing a cognitive wall between her belief that *S* expresses a truth and her belief that it has the semantics *F* makes no sense of the example.

Although one could develop other ways of implementing the metalinguistic strategy, they are all vulnerable to the sort of problem just explained (to my knowledge, first pointed out by Kripke in an unpublished lecture). Thus the metalinguistic strategy fails.

4. Reconceiving the problem[7]

In order to make progress, what one must take to heart is that the underlying problem is not about necessary or impossible propositions. It is about necessarily equivalent propositions, whether they are contingent or not. For instance, the sentences 'There is furze in Edinburgh' and 'There is gorse in Edinburgh' express the same contingent proposition, on the worldly approach to semantics under discussion, even though a speaker who understands both sentences need not be in a position to know that they have the same truth-value.

Of course, this is just a variant on the problem of cognitive significance, which Frege introduced his distinction between sense and reference to solve. One might therefore hope

that switching to a Fregean framework would help, by building modes of presentation into the semantics. But the present context makes two worries for Fregeanism salient.

The first worry concerns the big picture for metaphysics. Fregean thoughts—the senses of declarative sentences—are *perspectival* in a sense in which worldly intensional or Russellian propositions are not. A Fregean thought is a *mode of presentation* of a truth-value, presumably to a notional subject. By contrast, functions from worlds to truth-values and structured complexes of objects, properties, and relations are normally presentation- and subject-independent.[8] Thus Fregean thoughts seem less apt than such worldly propositions for being *what is objectively at stake* in an out-and-out metaphysical dispute, as traditionally conceived (in a way Kantians might describe as pre-Kantian). For thoughts can differ while the relevant non-presentational objects, properties and relations stay the same. In such cases, one might think, what is objectively at stake stays the same, while Fregean thoughts vary, so what is objectively at stake is no Fregean thought. Of course, Frege himself did not treat mathematics as lacking in objectivity; his approach was explicitly, indeed prototypically, anti-psychologistic. The commitments inherent in a Fregean semantic framework are unclear. Nevertheless, those engaged in a dispute over what they understand as a purely objective metaphysical question may be suspicious of treating what is at stake as a Fregean thought.

The second worry concerns the detailed implementation of Fregean semantics. Recall the potential cognitive differences between ‘There is furze in Edinburgh’ and ‘There is gorse in Edinburgh’ for an individual speaker. As explained in section 3, they do not depend on any community-wide cognitive difference between the two sentences. They can arise for normal speakers through accidental features of the process by which they acquire the words ‘furze’ and ‘gorse’. In such cases, the cognitive difference for the individual speaker is not explained by any difference between the senses attached to the two words at the level of the community. One might therefore be tempted to apply Fregean semantics at the level of the

individual speaker, to a family of more or less similar idiolects. But that brings back problems of its own. It ignores the lessons of social externalism and the division of linguistic labour (Putnam 1975). In particular, for purposes of metaphysics as a shared enterprise, we want to work in a common language. Frege himself insisted that thoughts (such as mathematical theorems) must be capable of forming part of the common heritage of mankind. Moreover, an individualistic account of senses has difficulty explaining their role in the inter-personal ascription of thoughts. When I say 'Mary thinks there is gorse in Edinburgh', do I attribute a thought involving *my* individual sense of 'gorse' to Mary? But I know that her individual sense of 'gorse' probably differs from mine. Or do I attribute a thought involving *her* individual sense of 'gorse' to Mary? But how can I do that when I do not know what her individual sense of 'gorse' is? And how do you understand me when your individual sense of 'gorse' may be different from both mine and Mary's? It looks as though a Fregean account of the inter-personal ascription of thoughts in natural language may be forced back into working with minimal community-wide senses after all, and so back to its failure to discriminate in sense between 'furze' and 'gorse'. Historically, it is no accident that Fregean semantics has largely dropped out of discussions of propositional attitude ascriptions in natural languages, despite the apparent head-start it gained from the sense-reference description: that is just not how natural languages work, or even *could* work. Fregean semantics does not solve our problem.

The moral to draw from 'furze' and 'gorse' and similar cases is not that semantic properties are Fregean. It is that cognitive significance does not supervene on semantic properties. At both the individual and the community levels, two sentences may have all the same semantic properties, yet differ in cognitive significance. Tracking cognitive significance is not just a semantic exercise. We must track the vehicles of semantic content too, the very sentences which have semantic content and their contexts. For example, we can distinguish

between believing the proposition that there is furze in Edinburgh under the guise of the sentence ‘There is furze in Edinburgh’ and believing the same proposition under the guise of the sentence ‘There is gorse in Edinburgh’ (in the terminology of Salmon 1986). To believe a proposition *simpliciter* is to believe it under some guise or other, where the believing-under relation has an extra argument-place for a guise. Similarly, to *know* a truth *simpliciter* is to know it under some guise or other. By treating the sentential guise as an extra parameter, we liberate semantics itself from pressure to make cognitive distinctions it is ill-suited to making; we thereby avoid distorting the semantic framework. Even the quasi-syntactic structure of Russellian propositions may reflect such inappropriate pressure on the semantics, by contrast with a purely intensional approach (Salmon 1986 works within a broadly Russellian framework). Just as we should not project the difference between ‘furze’ and ‘gorse’ onto their worldly semantic values, so we should not project differences in syntactic structure between sentences onto *their* worldly semantic values.

In thought, guises are not *what* we think, and not normally what we think *of*; they are what we think *with*. Similarly, in speech, when you make an assertion, a guise is not *what* you assert, and not normally what you assert it *of*; it is more like what you assert it *with* (though the hearer may receive it under a different guise). We must keep track of linguistic or, more generally, representational differences, without confusing them with differences at the level of reference.

Often, more than the linguistic expression type must be put into the guise to capture cognitive significance. This is clear for demonstratives: in the same context, someone may wonder ‘Is that gull that gull?’, where the first occurrence of ‘that gull’ refers to a seagull as she sees it in the distance, while the second refers to the same bird as she hears its cry. The case of someone who does not realize that the politician Paderewski and the pianist Paderewski are the same man also calls for such further differentiation of guises (Kripke

1979). Since full guises are not normally what need to be communicated, individuating them very finely carries little cost. Nor need guises always include something linguistic: the guise of a spatial thought might be more like a picture, seen or imagined.[9]

This separation of content from guise is not transparent to normal language-users in producing and comprehending ascriptions of propositional attitudes. As Kripke (1979) has emphasized, our ordinary practice can easily run into trouble with tricky cases. That does not mean that ordinary practice is somehow 'conceptually incoherent'. As with many cognitive challenges, we may be relying on *heuristics*, tests which are quick and easy to apply, and work well enough in most normal cases, but are not perfectly reliable. Indeed, the very disquotational principles that Kripke identifies as getting us into trouble may be just such helpful heuristics: for instance, "*A normal English speaker who is not reticent will be disposed to sincere reflective assent to 'p' if and only if he believes that p*". In using such rules of thumb, one may have no privileged access to their status as mere fallible heuristics, just as we have no privileged access to the heuristics on which we rely in making ordinary perceptual judgments (Williamson 2020). We rely naïvely on our heuristics, getting things mostly right, sometimes wrong, until philosophers force us to consider the inconsistencies into which we have been led, and even then the nature of the problem remains opaque to us, though some mix of philosophical, linguistic, and psychological investigation may eventually get us to the solution.

In any case, we can provisionally use the approach of ascribing acceptance or rejection of coarse-grained intensional propositions under guises to track what is going on in enquiries into non-contingent matters, such as logic, mathematics, and metaphysics. In those enquiries, propositions usually come under sentential guises, but not always: in geometry, for example, a proposition may come under the guise of a diagram. The trap not to fall into is that of thinking that the need for tracking sentential guises shows anything distinctive about

those fields—for instance, that they are somehow partly linguistic enquiries in some sense in which more 'empirical' enquiries are not.

Admittedly, fields may differ in how far we can use differences in proposition expressed as convenient proxies for cognitive differences between sentences—doing so works much better in history than it does in mathematics—but in principle the two levels are *never* equivalent, and in practice the inequivalence will sometimes obtrude in every field, though more frequently in some than in others. For example, in ancient history, doubt is not uncommon as to whether the same name in different documents refers to one person or two.

Of course, this separation of semantic value from cognitive significance forms a coherent picture only if there are systematic connections between the two levels. Compositional semantics provides such connections. Although the semantic structure of a sentence is not even roughly similar to any structure intrinsic to the proposition it expresses, the former determines the latter in more or less principled ways, described by a compositional semantic theory for the language. Even in discourse where the only propositions are the necessary truth and its contradictory, a multitude of properties and relations are normally in play as the semantic values of predicates. Thus a standard first-order language for arithmetic can express infinitely many distinct monadic properties (intensions) of natural numbers. The case of metaphysics is analogous. When things go well, we learn *how* the properties and relations of interest are necessarily interconnected.

One may still feel puzzled. For when we learn how those properties and relations are necessarily interconnected, *what* we learn are necessary truths, which by intensionalist lights are all one. Indeed, if metaphysical truths are all necessary, how do we learn anything in metaphysics, since presumably we already knew the trivial necessary truth before we started doing metaphysics? In response, a first point is that calling the necessarily true proposition 'trivial' already confuses the issue, because the distinction between 'trivial' and 'non-trivial',

like that between ‘obvious’ and ‘non-obvious’, arises primarily at the cognitive level: the trivial is the very easily known. The necessarily true proposition is trivial under the guise of the equation ‘2 + 2 = 4’ but highly non-trivial under the guise of a statement of Fermat’s Last Theorem. Similarly, in such cases, learning and discovery must themselves be understood with respect to guises: mathematicians who already knew the necessary truth under one guise came to know it under another. The novelty was in the guise, not in the proposition known. Again, the same points apply to logic and metaphysics.

But if you know a truth under one guise, why bother to learn it under another? That would be a good question if knowledge were valued as a miser’s hoard of true propositions. But not even true propositions have intrinsic value: the value is in how we are cognitively related to them. We can bear dramatically different cognitive relations to the same proposition under different guises. In learning an old truth under a new guise, we acquire a potentially valuable new cognitive relation to the old truth.

None of this involves a return to the discredited metalinguistic view. The latter makes the mistake of trying to get the content to do all the cognitive work, forgetting that even a metalinguistic content can be presented to the subject under different guises. The point is rather that *any* content is *present* to a subject at a time only in some form or other; that form is its guise. Even physical aspects of linguistic form are cognitively significant, because they facilitate or impede cognitive manipulation. Mathematicians know this well; metaphysicians would be well-advised to know it too. As Bertrand Russell observed, ‘a good notation has a subtlety and suggestiveness which at times make it seem almost like a live teacher’; ‘Notational irregularities are often the first sign of philosophical errors’ (1922: xix). That is why definitions matter in metaphysics, even though they merely abbreviate longer forms of words: a good definition makes salient and handy a distinction which cuts at the joints. In that respect, even metaphysics is a kind of embodied cognition.[10]

References

Bacon, A. 2018: 'The Broadest Necessity', *Journal of Philosophical Logic*, 47: 733-783.

Cappelen, H., and Dever, J. 2013: *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.

Carnap, R. 1950: 'Empiricism, Semantics, and Ontology', *Revue Internationale de Philosophie*, 4: 20-40.

Dorr, C. 2016: 'To be F is to be G', *Philosophical Perspectives*, 30: 39-134.

Elbourne, P. 2005: *Situations and Individuals*. Cambridge MA: MIT Press.

Hirsch, E. 2021: 'Ontology by Stipulation', in J.T.M. Miller (ed.), *The Language of Ontology*, 7-22. Oxford: Oxford University Press.

Kripke, S. 1979: 'A Puzzle about Belief', in Avishai Margalit (ed.), *Meaning and Use*: 239 -283. Dordrecht: Reidel.

Kripke, S. 1980: *Naming and Necessity*. Oxford: Blackwell.

Lewis, D. 1970: 'General Semantics', *Synthese*, 22: 18-67.

Lewis, D. 1979: 'Attitudes *De Dicto* and *De Se*', *Philosophical Review*, 88: 513-543.

Lewis, D. 1996: 'Elusive Knowledge', *Australasian Journal of Philosophy*, 74: 549-567.

Magidor, O. 2015: 'The Myth of the De Se', *Philosophical Perspectives*, 29: 249-283.

Marconi, D. 2011: 'Wittgenstein and Williamson on Conceptual Analysis', in Richard Davies (ed.), *Analisi: Annuario e Bollettino della Società Italiana Analitica (SIFA) 2011*, 91-102. Milan: Mimesis.

Putnam, H. 1975: *Mind, Language and Reality: Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press.

Rayo, A. 2013: *The Construction of Logical Space*. Oxford: Oxford University Press.

Russell, B. 1922: ‘Introduction’, in L. Wittgenstein, *Tractatus Logico-Philosophicus* (trans. C.K. Ogden), ix-xxv. London: Routledge and Kegan Paul.

Salmon, N. 1986: *Frege’s Puzzle*. Cambridge, Mass.: MIT Press.

Salmon, N. 1989: ‘The Logic of What Might Have Been’, *Philosophical Review*, 98: 3-34.

Stalnaker, R. 1984: *Inquiry*. Cambridge, Mass.: MIT Press.

Stalnaker, R. 1999: *Context and Content*. Oxford: Oxford University Press.

Stalnaker, R. 2011: ‘The Metaphysical Conception of Analyticity’, *Philosophy and Phenomenological Research*, 82: 507-514.

Thomasson, A. 2015: *Ontology Made Easy*. New York: Oxford University Press.

Thomasson, A. 2020: *Norms and Necessity*. Oxford: Oxford University Press.

Williamson, T. 2007: *The Philosophy of Philosophy*. Oxford: Wiley-Blackwell. 2nd edition 2021.

Williamson, T. 2011a: ‘Reply to Stalnaker’, *Philosophy and Phenomenological Research*, 82: 515-523. Reprinted in Williamson 2021: 471-480.

Williamson, T. 2011b: ‘Three Wittgensteinians and a Naturalist on *The Philosophy of Philosophy*’, in R. Davies (ed.), *Analisi: Annuario e Bollettino della Società Italiana Analitica (SIFA) 2011*, 127-137. Milan: Mimesis. Reprinted in Williamson 2021: 481-483, 553-562.

Williamson, T. 2013: *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

Williamson, T. 2016: ‘Modal Science’, *Canadian Journal of Philosophy*, 46: 453-492.

Williamson, T. 2020: *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford: Oxford University Press.

Williamson, T. 202X: ‘Disagreement in Metaphysics’, in M. Baghramian, J. A. Carter and R. Rowland (eds.), *Routledge Handbook of the Philosophy of Disagreement*, London: Routledge, forthcoming.

Williamson, T. 202Y: ‘Moral Anti-Exceptionalism’, in P. Bloomfield and D. Copp (eds.), *The Oxford Handbook of Moral Realism*. Oxford: Oxford University Press, forthcoming.

Notes

1. I use the term ‘metaphysics’ as it is standardly used in contemporary philosophy, with a standard view of what counts as metaphysics. The arguments of this paper are robust to minor variations in that respect.

2. For an exchange on the Wittgensteinian claim about impossibility, see Marconi 2011 and Williamson 2011b. For an exchange on Stalnaker’s view see Stalnaker 2011 and Williamson 2011a. For a view which combines ideas from Wittgenstein and Stalnaker see Rayo 2013.

3. For relevant discussion of what counts as understanding, see Williamson 2007: 97-8 and the exchange between Stalnaker 2011 and Williamson 2011a. More generally, the discussion of analyticity in Williamson 2007 supports the arguments of this section.

4. The semantics works most smoothly with the stipulation that a rigid designator for *x* designates *x* even with respect to worlds at which *x* is not concretely present. After all, the semantics characterizes how *we* use words, speaking in our world *about* actual and

counterfactual worlds, not how those words *would have been used* in those counterfactual worlds.

5. The leading defender of intensionalism about content is Robert Stalnaker (1984, 1999). Since I have engaged in detail with his application of intensionalism to content in philosophy elsewhere (Stalnaker 2011, Williamson 2011a), I will not do so here. In effect, his approach is a version of the metalinguistic strategy; my concern in this section is with the general strategy.

6. The notation is perspicuous because <X, Y> = <X*, Y*> just when X = X* and Y = Y*, so one can individuate the whole by individuating its constituents. This justifies the ordered pair notation, though Russellian propositions need not literally *be* ordered *n*-tuples. But this fineness of grain also generates the Russell-Myhill paradox, which makes a pure Russellian account inconsistent. See Dorr 2016 for discussion.

7. The approach in this section builds on the proposal in Williamson 2007: 66ff.

8. David Lewis (1979) turned the intensional framework perspectival by reworking it in terms of *centred* worlds, with a distinguished agent and time, although he did not build in other aspects of modes of presentation. For criticism of such hybrid approaches see Cappelen and Dever 2013 and Magidor 2015.

9. What sort of Russellian proposition would correspond to a picture?

10. Thanks to participants in the 2021 ‘New Directions in Metaphilosophy’ conference at the University of Kent and a Lugano Philosophy Colloquium (both virtual), to Daniel Kodsi and Luis Rosa (in correspondence), and to two anonymous referees for helpful comments on earlier versions of this material.