2. Middle Views

2.0 Introductory comments

Ramsey's prescience in suggesting a redundancy theory, and separating the question of the semantics of belief from that of the nature of truth, is all the more impressive for coming at a time when the theory of judgement was almost universally allowed to dictate the terms on which a theory of truth could be constructed. Despite a change of vocabulary, this underlying approach remained the norm even through the change in methodology witnessed in the middle years of the century and later labelled 'the linguistic turn'. In the theory of truth this change meant that the focus of debate shifted to questions surrounding the use and meaning of the truth-predicate. However, despite a common linguistic focus, there was little consensus during this period about what the purpose of a theory of truth is. In particular, two important dimensions of disagreement concerning how to theorize about truth opened up. On the one hand, there was disagreement about whether theorizing should take the form of an investigation into the formal, logical properties of the truth-predicate, or that of an investigation into the linguistic practices in which the use of such a predicate has a home (i.e. 'fact-stating discourse'). On the other hand, philosophers disagreed as to whether they should be investigating our actual truth-predicate or devising some preferable replacement predicate. Nevertheless, throughout this period, it was almost always a theorist's conception of semantics that dictated the framework in which they investigated truth.

2.1 Logical empiricism (I)

Like Ramsey, the logical empiricists were deeply influenced by Wittgenstein's *Tractatus* and at least in the first years of the Vienna Circle, and despite their various differences, they adopted a number of that book's fundamental theses, including that of extensionality: that the only propositions with sense are either atomic propositions or those constructed truth-conditionally from them. But they interpreted Wittgenstein's sparse, formal conception of language in the light of pragmatist considerations. For example, they usually read him, tendentiously, as claiming that atomic propositions were in some sense about immediate experience. This conception of the relation of language to experience, together with the thesis of extensionality, led the empiricists to their famous slogan that the only meaningful statements are those that are verifiable. They also inherited one of the major tensions of the *Tractatus*. Both Wittgenstein's picture theory and the verificationist criterion of meaningfulness seem to entail that logical or analytic truths (or falsehoods) are meaningless and thus neither true nor false. On the other hand, these propositions are as truth-functional as any proposition and thus seem apt to be considered true or false.[1] This tension made it difficult for them to formulate a unified conception of truth.

The logical empiricists' conception of meaningfulness as verifiability led them to reject a number of traditional philosophical questions as, at least in some sense, meaningless. In particular, because they took questions about the nature of extra-linguistic reality to lead to

[1] For a pellucid description of the tension as it appears in the *Tractatus* (and Wittgenstein's attempt to resolve the tension) see Fogelin 1987: 45ff.

pointless, irresolvable disputes, semantic questions were impugned too, on the grounds that they concerned mysterious relations between language and this extra-linguistic reality. One of their early projects was thus to create a language suitable for science that would prevent such questions being formulated. To this end, the fundamental expressions of such languages were to be words for the description of our immediate experience. The only other allowable expressions were logical and syntactical expressions and those that could be defined out of the observational expressions using the logical resources of the language. To the extent that semantic questions can be asked in these languages, they must be put in syntactic terms. For example, Carnap suggested that instead of the question 'Is this book about Africa?' we should ask 'Does this book contain the word "Africa?"' [*sic*].[2] In short, for the early logical empiricists the logic of scientific languages was to be the object of study for philosophers and the method was to be purely syntactic.

In the light of this brief characterization, it is not surprising that the popular view of the logical empiricists' theory of truth is that they either rejected truth as a metaphysical pseudo-concept or endorsed a version of the coherence theory of truth by identifying truth with verification or confirmation. The popular view captures some of the positivistic motivations for their various conceptions of truth. Truth was obviously an uncomfortably semantic notion for many of them and such a concept would have to prove both its empiricist credentials and its usefulness in describing relations amongst sentences of a scientific language. The most obvious truth-related concept that met these criteria is verification. But in fact the popular view overlooks the variety of accounts endorsed by the logical positivists. For example, Moritz Schlick shared neither of these attitudes to truth.

Schlick saw himself as holding firm to the Tractarian correspondence theory despite its involving the sort of metaphysical talk that the logical empiricists were trying to purge from serious theorizing – in this case an appeal to both facts and a semantic relation between language and the world. Schlick argued that there was nothing mysterious about the claim that we compare propositions and facts to see whether a proposition is true. Doing so is just what it is to verify a proposition:

> I have often compared propositions to facts; so I had no reason to say that it couldn't be done. I found, for instance, in my Baedeker the statement: "This cathedral has two spires," I was able to compare it with 'reality' by looking at the cathedral, and this comparison convinced me that Baedeker's assertion was true … I meant nothing but a process of this kind when I spoke of testing propositions by comparing them with facts. (Schlick 1935: 400)

Unlike Wittgenstein's theory of truth, Schlick's did not involve the claim that propositions picture reality. Instead, like James and Dewey, Schlick explained the correspondence relation in terms of verification. However, for Schlick verification was understood according to a foundationalist epistemology whose basic propositions were verified by direct confrontation with sense experience. These are captured in protocol sentences 'which express the result of a pure immediate experience without any theoretical addition',[3] so that the experiencer could not make a significant contemporaneous error in judgement. This entails that, once justified, a

[2] Carnap 1935: 65.

[3] Hempel 1935: 11. See, in particular, Schlick 1934 and 1935.

sentence cannot later become unjustified. According to Schlick, this view also entailed that verification involved directly confronting propositions with facts. This treatment of verification was what the logical empiricists saw as justifying the claim that Schlick's position was a correspondence theory, on the grounds that, on this view, being verified is a *stable* property that propositions have in virtue of comparison with facts.[4] However, given its appeal to extra-linguistic facts and the difficulty it faces in assigning its claims to either the class of empirical or that of syntactic sentences, Schlick's theory of truth was not popular amongst the other members of the Vienna Circle and, with his death in 1936, was entirely abandoned.

At first, Carnap and Neurath shared Schlick's foundationalism. However, under the influence of Duhem and Poincaré, they came to accept confirmation holism and thus abandoned the idea that any proposition could be tested or verified on its own. Instead, they insisted that statements could only meaningfully be compared with other statements. On this view, even basic propositions may well be given up in the face of an inconsistency if this is the simplest way to re-establish a consistent system of belief. They thus replaced Schlick's foundationalist epistemology with a coherence theory of justification. As Hempel emphasized, this move to a holistic view of justification also led them to abandon the thesis of extensionality, for no proposition could be regarded as intrinsically atomic. It also meant that they could not accept Schlick's correspondence theory of truth, for it implies that sentences can change their epistemic status over time and thus contrasts with what they often called the "absolutist" nature of both correspondence truth and our "ordinary" conception of truth. However, *pace* Hempel, their coherence theory of justification does not imply that they were thereby committed to a coherence theory of truth.[5]

That said, this divergence between the "ordinary" notion of truth and their conception of confirmation left Carnap and Neurath with a choice: either abandon the notion of truth altogether, or replace our "ordinary" notion with a notion defined in terms of confirmation. Neurath chose the latter course and so made the move from a coherence theory of justification to a coherence theory of truth:

> If a statement is made, it is to be confronted with the totality of existing statements. If it agrees with them, it is joined to them; if it does not agree, it is called 'untrue' and rejected; or the existing complex of statements of science is modified so that the new statement can be incorporated; the latter decision is mostly taken with hesitation. *There can be no other concept of 'truth' for science*.
> (Neurath 1931: 53; his italics)

In this passage, Neurath not only endorses a coherence theory of truth but suggests that there may be other notions of truth available that have no place in science. In fact, it is clear that Neurath rejected the ordinary notion of truth because it was different from a coherence notion of justification. So while Neurath in a sense offers a coherence theory of truth, he openly

---

[4] This identification of what is at issue was anticipated by Bradley (1909: 201-7), in criticism of Russell 1907 and Stout 1907.

[5] Hempel (1935: 14) seems to follow Neurath in seeing this move as virtually automatic. Walker (1989: ch. IX) is admirably clear that the coherence theory of truth is not implied by the coherence theory of justification.

acknowledges that this is not the "ordinary" conception of truth. Importantly, unlike Joachim's and Blanshard's, his coherence theory was not based on claims about the essential nature of thought. Nor would Neurath have thought kindly of the metaphysical doctrine that reality is ideally coherent. Rather his holistic, coherence theory of justification together with a distrust of metaphysics led him to a position much like that of the pragmatists.

The case of Carnap is rather more difficult and any interpretation of his pre-Tarski views on truth is bound to be controversial. Many commentators attribute to him a coherence theory of truth based on his coherence theory of justification.[6] But while he may have once agreed with Neurath, most of his published writings on truth in the early thirties present a rather different attitude. In *The Logical Syntax of Language*, despite the remarkably rich notion of syntax to which he appealed, Carnap nevertheless assumed that 'truth' could not be defined syntactically.[7] This was not because he thought that any such definition would lead to contradiction via the liar sentences: he explicitly states a way to avoid the liar paradox by distinguishing clearly between object-languages and metalanguages. Rather, he is clear that once this problem is resolved we have to admit that 'truth and falsehood are not proper syntactical properties; whether a sentence is true or false cannot generally be seen by its design, that is to say, by the kinds and serial order of its symbols.'[8] Of course, he thought that the truth of some sentences, the analytic ones, could be discovered merely by observing their syntactical properties (and one of his ongoing projects was to provide an adequate syntactic definition for the analytic truths). Nevertheless, Carnap is obviously right that no syntactic definition of truth will be possible such that the definition will serve to demarcate the truths from the falsehoods.

Moreover, like Popper,[9] Carnap thought that there was no need for a truth-predicate in logic. He claimed that we could translate the majority of sentences containing 'true' into sentences that do not. In particular, '"p" is true' can be translated by 'p' and 'If "p" is true, then so is "q"' by '"q" is a consequence of "p"'. So, for Carnap at this time the truth-predicate was to be avoided primarily because it could not be given an adequate syntactic definition and seemed to lack any legitimate, non-redundant uses. The idea that the truth-predicate is redundant bears an obvious resemblance to Ramsey's redundancy theory of truth. And although neither Carnap nor Popper refers to Ramsey in this context, it is clear that Carnap had read Ramsey and that Popper had read Carnap.[10] However, Carnap did not make the claim that our ordinary truth-predicate was redundant, only that he saw no use for it in logic.

A. J. Ayer, on the other hand, quite explicitly followed Ramsey in claiming that, owing to the redundancy of the truth-predicate, there was no problem of truth.[11] Ayer, unlike Ramsey, did

---

[6] See, for example, Ayer 1959: 20 and Hempel 1935: 14.

[7] For helpful discussions of Carnap's syntactical resources and his claim that truth could not be defined syntactically, see Ricketts 1996 and Coffa 1991: 285-306.

[8] Carnap 1934: 216.

[9] Popper 1934: 274f.

[10] Carnap refers to Ramsey 1931 a number of times in his 1934 and Popper refers to Carnap's 1934 in his discussion of truth (1934: 275).

[11] Ayer 1936: ch. 5.

not turn his mind to the problematic cases for this theory and so offered no further hint as to how they might be handled. However, he did go further than Ramsey in concluding that the correct analysis of truth reveals that there is no property of truth to wonder about the nature of:

> There are sentences … in which the word "truth" seems to stand for something real; and this leads the speculative philosopher to enquire what this "something" is. Naturally he fails to obtain a satisfactory answer, since his question is illegitimate. For our analysis has shown that the word "truth" does not stand for anything, in the way which such a question requires.
> (Ayer 1936: 89)

It is surprising that a theory of truth so amenable to the logical empiricists was not more generally adopted. In fact, it took the ground breaking logical work of Alfred Tarski to open the eyes of many of the logical positivists to the availability of a theory of truth that was as deflationary as Ramsey's.[12]

2.2 Tarski's semantic conception of truth

Tarski was well aware that the concept of truth was viewed with suspicion by many of his contemporaries. He himself was suspicious of the everyday use of truth-predicates in natural languages. His main concern was that such use led to inconsistency because of the possibility of formulating liar sentences like 'This sentence is false'.[13] One of Tarski's main goals in supplying a definition for a truth-predicate was to ensure that any such term would avoid these paradoxical consequences. Thus, based on his diagnosis of the conditions that allow the liar paradox to be generated for the 'ordinary' truth-predicate, Tarski proposed a number of formal constraints on any adequate definition of a truth-predicate. In particular, he claimed that all semantically closed languages are inconsistent, where a semantically closed language is one that has the means of referring to its own expressions and itself contains semantic expressions like 'true' and 'false'. As well as being semantically closed, Tarski thought that natural languages do not allow us to set up precise formal conditions for the adequate definition of a truth-predicate because we cannot exactly specify the formal structures of these languages (including which of the expressions are meaningful). These two features of natural languages led him to abandon any attempt to define either our ordinary truth-predicate or any other truth-predicate for a natural language. Instead, his project was to define a truth-predicate (or provide a recipe for defining truth-predicates) for formal languages. Moreover, Tarski was only interested in defining the term 'true sentence' for these formal languages and not in any notion of truth as it applies to propositions or beliefs.

To make sure that the predicate he defined was worthy of being called a *truth*-predicate, Tarski suggested his famous material adequacy condition, Convention T. Convention T asserts that any adequate definition of a truth-predicate must entail all instances of the T-schema:

---

[12] We do not mean to suggest that Tarski's definition obviously amounts to a deflationary theory.

[13] Amongst the theorists we are considering, Tarski was largely alone at the time in allowing his concern with the semantic paradoxes to shape his views about truth. In these days of paraconsistent logic, it may no longer be as obvious as he assumed that such inconsistencies are an overwhelming cause for concern.

(T) X is true iff p

where 'X' is replaced by a name or structural description of some sentence and 'p' is replaced by that sentence or a translation of it in the meta-language. (An instance of the schema would thus be "'*Der Schnee ist weiß*' is true iff snow is white", with German as the object- and English as the meta-language.) T clearly captures a central feature of the notion of truth and as trivial as it may seem, Tarski's clear articulation of this adequacy condition was a major contribution to our understanding of truth. It is important to note, though, that in first offering this condition Tarski rather cautiously claimed that this was only a condition of adequacy on a particular conception of truth, namely the semantical conception, according to which 'a true sentence is one which says that the state of affairs is so and so, and the state of affairs indeed is so and so.'[14] Tarski claims that this type of statement is essentially an elaboration on what he calls the '*classical* conception of truth ('true – corresponding with reality').'[15] However, despite his suggestions it is not clear that a semantic definition of truth (and his definition in particular) must amount to a correspondence theory of truth (though it does seem to capture the 'classical', Aristotelian conception).

Finally, it is crucial to Tarski's enterprise that he not use any undefined semantic terms in his definition. Otherwise, the role of primitive concepts would be 'played by concepts which have led to various misunderstandings in the past.' Moreover, 'it would then be difficult to bring this method into harmony with the postulates of the unity of science and of physicalism (since the concepts of semantics would be neither logical nor physical concepts).'[16] The influence and the vocabulary of the logical positivists is quite plain here: Tarski's goal was not just to avoid the antinomies, but also to provide a definition of truth that would be acceptable to those with positivistic or empiricist suspicions about semantics. Thus he hoped to define semantical concepts 'in terms of the usual concepts of the metalanguage' so that they are 'reduced to purely logical concepts, the concepts of the language being investigated and the specific concepts of the morphology of language.'[17]

Tarski's aim, and his self-imposed constraints, were accordingly these: to produce a definition of 'true sentence' for a formal object-language free from semantic terms and with a precisely specified structure, such that this definition entailed all instances of the T-schema and was formulated in a metalanguage that was about this object-language and used no semantical terms as primitives.

The most obvious way to construct a definition that meets these constraints would be to list all the instances of the T-schema. After all, Tarski himself claims that these instances are partial definitions of truth. For any language with an infinite number of sentences, however, this approach will not provide us with an explicit definition of truth. For such languages the obvious step would be to generalize on the T-schema itself to get something like TG:

---

[14] Tarski 1933: 155.

[15] ibid. 153. Tarski claims that the semantical definition of truth is an elaboration of this classical conception (and quotes Aristotle's famous slogan about truth as an example of the semantical conception) at 155f.

[16] Tarski 1936: 406.

[17] ibid.

(TG) ∀p (‘p’ is a true sentence iff p).

As should be clear from our discussion of Ramsey, however, whether we can do this depends on how we read quotation marks and the universal quantifier. If we read the quantifier objectually, then, even if the variables were to range over propositions, the sentence would be ungrammatical. On the other hand, we could read it substitutionally. Substitutional variables do not range over a set of objects so that when we instantiate formulas containing them they are replaced with names of these objects. Instead, they are associated with a class of expressions (in this case sentences) so that when we instantiate we replace the variable with an expression and not a name for the expression. However, whether this approach works will depend on how we read the quotation marks. As Tarski points out, the approach will not work if we treat “‘p’” like a syntactically simple name so that the expression inside the quotation marks is no more in quantifiable position than ‘lip’ is in ‘Philip was sleeping’. Rather, we must treat quotation marks as functions that take an expression as input and create a name for it. This allows us to read TG so that it not only comes out as grammatical but also seems to provide a definition of truth that meets Tarski’s constraints.[18] The only significant reason Tarski gave for not adopting this definition is that he thought that it generates contradiction.[19] Yet if he had been as careful in distinguishing between object-language and metalanguage in this case as he was for his own definition this inconsistency would not arise.[20] There are other reasons for worrying about using substitutional quantification in this context, but Tarski did not share them and we shall see that it is anyway far from clear that they are conclusive.[21]

Yet another option would be to provide a recursive definition of truth by first defining truth for atomic sentences using the appropriate instances of the T-schema and then defining the circumstances under which complex sentences built out of atomic sentences and logical operators are true. For example, one part of such a definition would be <(A & B)> is true iff <A> is true and <B> is true.[22] This is obviously much like the logical atomist approach of Russell and Wittgenstein except that the T-schema is used instead of providing a general, correspondence account of truth for atomic sentences. However, Tarski rejected the idea that quantified sentences should be treated like infinite disjunctions or conjunctions and so abandoned the idea that one could define truth for all sentences on the basis of a definition of truth for atomic sentences.

To overcome these problems Tarski began with a recursive definition of a near substitute of

---

[18] Field offers a similar (but importantly different) approach using schematic variables. Field 2001: 115 and 141ff.

[19] Tarski 1933: 161. Both Mackie and Soames, for example, agree that this is the only substantial reason Tarski offers. Mackie 1973: 32 and Soames 1999: 89.

[20] Marcus 1972: 246f. In Marcus’s own words “considerations of definitional adequacy require that for the substitution class of sentences the condition on quantification be

(Q’) If A = (x)B(x) then v(A) = T iff v(B(t)) = T for each sentence t such that (B(t)) contains fewer quantifiers than A.”

[21] These objections include complaints of circularity (see for example Platts 1980: 14f and Horwich 1998a: 25f, but see also Camp 1975 and Soames 1999: 90ff for responses) and also complaints that substitutional quantifiers are inconsistent with convention T (see, for example, Wallace 1971, and Kripke 1976 and Camp 1975 for responses).

[22] We have used angle brackets as a convenient substitute for corner quotes in this sentence.

sentence truth, namely satisfaction of sentential functions, and showed how we could define sentence truth on the basis of this notion. A sentential function is an open sentence, or a formula that contains variables that are not bound by any quantifiers (e.g. 'x is married to Queen Elizabeth'). To give an idea of what satisfaction is, we can speak of a sentential function's being satisfied by certain objects when replacing the variables with names for these objects results in a true sentence (e.g. Prince Philip satisfies 'x is married to Queen Elizabeth'). However, given that we want to use the notion of satisfaction to define truth, we need some other way to *define* satisfaction. Moreover, Tarski's initial truth definition concerned languages that do not have names, so we will follow him in this and then point out the obvious way to extend his definition. For a finite number of predicates and relations, a simple way to define satisfaction would be to simply list each predicate and relation and the conditions under which it is satisfied (e.g. for all x, x satisfies the predicate 'is married to Queen Elizabeth' iff x is married to Queen Elizabeth).

The task of providing a general definition of satisfaction, however, is harder. Because open sentences can have more than one variable and because they can be closed by applying quantifiers to bind the variables as well, Tarski adopted the technique of talking about sentential functions' being satisfied by *sequences* of objects. A sequence of objects is basically a potentially infinite list of objects where the way the objects are ordered in the list is crucial. When considering whether a sentential function is satisfied by a sequence Tarski pairs the variables in the sentential function with objects in the sequence so that if we have variables x_1, x_2, x_3, they are paired with the first, second and third objects of a sequence. So, if our sentential function was 'x_1 is the husband of x_2' then it would be satisfied by any sequence of objects that had Prince Philip as its first member and Queen Elizabeth as its second (as well as many other sequences, of course). Dealing with the quantifiers is less straightforward, but the general idea is the same. We consider sequences of objects such that the variables of the sentential function are paired with objects in the sequence based on their position in the list.[23]

The general pattern (although not the logical machinery) of Tarski's truth definition is quite easy to grasp. First, he provides a recursive definition of satisfaction of sentential functions by sequences of objects. This crucially involves assigning free variables to objects in the sequences. For sentences there are no free variables and so, whatever sequence we are talking about, this sequence will either satisfy the sentence or not. Once we notice this, the move from a definition of satisfaction to a definition of truth is obvious. True sentences are those that are satisfied by all sequences. False sentences are satisfied by none.

With this ingenious approach, Tarski managed to provide a definition of true sentence that satisfied all the constraints he had imposed. Of course, the definition he supplied is a definition of truth relative to a particular formal language, so his approach could never count as a definition of truth *simpliciter*. That is, as Max Black emphasized,[24] although his work can

[23] For the few who are interested and yet have not met Tarski's treatment of the quantifiers, his treatment of the existential quantifier is such that a sequence s_1 satisfies some existentially quantified sentential function $\exists x_k A$ iff there is some other sequence, s_2, that satisfies A and differs at most from s_1 at the k^{th} place.

[24] Black 1948.

be applied to different formal languages, the definition of 'true sentence' is unique to the particular language in question. If, as it seems, our ordinary notion of truth is not relative to a language in this way, Tarski has failed to capture the meaning of 'true'. In particular, if we change the language for which we have provided a Tarskian truth-definition by as much as adding one predicate, the definition becomes obsolete. The definition does not tell us how to incorporate this new vocabulary. Moreover, his definitions only apply to a certain sort of *formal* language. Whether or not a Tarski-style definition can be extended to natural languages is of great philosophical significance. One way in which it can be easily extended is to incorporate languages with names. For a language with a finite list of names, a definition of denotation can be given merely by listing the names and the objects they denote. We can then define truth for atomic sentences quite simply along the following lines:

'Fa' is true iff the object denoted by 'a' satisfies the predicate 'F'.

Of course, as was the case with predicates, any language with an infinite number of names will resist being treated in this way. However, there are also many other types of natural language expressions that resist being treated in a Tarskian fashion; such as indexical expressions, reported speech, adverbial constructions and sentences in non-indicative moods.[25]

So much about Tarski's goals and the extent to which he achieved them is relatively uncontroversial. What is far more difficult to achieve consensus on is what, if anything, Tarski's definition contributes to the philosophical debate about truth. In particular, does it serve to rehabilitate the correspondence theory, as he himself sometimes seems to suggest? Karl Popper certainly thought so, arguing that part of Tarski's great contribution was to point out the adequacy conditions on any correspondence theory of truth.[26] Popper quite rightly sees that a correspondence theory needs the capacity to describe truth-bearers, a way of referring to all the facts described in the object language and an account of the correspondence relation that holds between truth-bearers and facts. But Popper also reads these constraints into Tarski's definition of truth. As we saw, Tarski insisted that the meta-language have the resources to describe or name all sentences of the object language and also to provide translations for all these sentences. For Popper's characterization of Tarski to be right, this latter translation constraint must amount to a way of referring to or describing facts. In other words, as Popper says, the right hand side of the T-schema must state the fact to which the sentence referred to on the left hand side corresponds.[27] However, if we treat facts as whatever is stated by statements like 'Snow is white,' surely we reduce the notion of a fact to something that even coherence or pragmatic theories could admit. There is nothing preventing coherence theorists from accepting that 'Snow is white' is true iff snow is white, provided that they read the sentence as an expression of the belief that snow is white and they commit themselves to the claim that snow is white iff the belief that snow is white is part of the ideally coherent set of beliefs. This is not to agree with Black's claim that Tarski's theory is philosophically insignificant because it is consistent with all philosophical theories. It is simply to point out that Tarski's T-schema (which is not his definition of truth) should be

[25] Although much work has been done to show how these constructions can be treated in a Tarskian way by those committed to Davidsonian semantics. For some examples of this work see Davidson 1980 and 1984.

[26] Popper 1974: 401.

[27] Popper 1974: 402.

accepted by all theories of truth.

Of course, this is also not yet to deny that Tarski's theory might be a correspondence theory. At one time Donald Davidson argued that it was a virtue of Tarski's account that it offered a correspondence theory without needing to appeal to fact-like entities.[28] Tarski's *definition* of 'true sentence' is that a true sentence is satisfied by all sequences. Davidson thus understood sequences as Tarski's replacement for facts. On this reading of Tarski, it is the world that makes sentences true, but it is not the case that each sentence corresponds to some part of the world (a particular fact). It is all the possible sequences of objects taken together that determines that a sentence is true (together with the semantic facts about the sentence). In a sense, then, it is the entire universe that either does or does not make a sentence true. For Davidson, this is just as well. According to Davidson, there is a simple proof, first offered by Frege and now generally known as the 'Slingshot', that shows that any attempt to appeal to facts as the referents of sentences collapses into the claim that there is only one Big Fact that makes all true sentences true.[29] A Tarski-style approach to correspondence theories would also have the rather significant virtue of putting an end to annoying questions about negative facts, general facts, mathematical facts and so on.

However, an appeal to sequences does not on its own constitute a correspondence theory any more than does Tarski's appeal to the availability of metalanguage translations. At a minimum, Tarski's theory would also need to explain the relation that holds between sentences and sequences such that the holding of this relation will constitute the correspondence of the sentence with the world. Davidson and others have claimed that Tarski's definition of satisfaction is precisely what fills this role and so completes the correspondence account of truth. On this reading, Tarski has discovered the relation of satisfaction to be the correspondence relation that philosophers have been searching for. Correspondence thus turns out to be a semantic relation between sentences and the world. In calling such a position a correspondence theory, though, Davidson et al. commit themselves to treating all theories of the following form as correspondence theories:

A sentence is true iff the sentence means that p and p

This would mean that even Ramsey's redundancy theory is automatically a correspondence theory.[30] In fact, it looks as if, in general, treating Tarski's definition as a version of a correspondence account means that Ramsey's account will need to be treated as a correspondence theory too. Like Tarski, Ramsey noticed that the problem of truth would be easy to solve if all occurrences of 'true' were of the form 'It is true that bitumen is black.' Also like Tarski, Ramsey saw that there were grammatical difficulties in extending this story to cover all cases and that the solution lay in uncovering all the different logical forms that judgements could have and then explaining truth recursively. It is difficult to see what parts

[28] Davidson 1969.

[29] There is a large literature concerning the cogency of this argument and the consequences of its conclusion. For an excellent overview and an extensive bibliography see Neale (1995).

[30] Field 1986 argues that Ramsey's theory is a version of the correspondence theory, based on reading Ramsey as offering a pragmatic account of what makes it the case that judgements have the truth-conditions that they do. According to the way of characterizing truth theories currently under discussion in the text, even if Ramsey had said that there is nothing to say about what makes judgements have truth-conditions he would still have been offering a correspondence theory.

of Tarski's superior logical apparatus could turn such a redundancy theory into a correspondence theory.[31]

2.3 Logical empiricism (II): the impact of Tarski

It is worth returning to the logical empiricists to see how the definition of truth given by one of their close associates affected their own conceptions of truth. The difference between Neurath's and Carnap's reactions to Tarski's work is telling. Neurath agreed that Tarski had provided a definition that made sense of our familiar notion of truth rather than the verificationist notion Neurath himself had been offering. Nevertheless, Neurath continued to worry that this notion of truth carried with it absolutist implications, i.e., that there was both a reality independent of the language in which we describe it and that we could have some sort of absolute certainty in the truth of propositions about this reality [32] Thus Neurath retained his scepticism towards 'true' until he could be shown that the Tarskian notion of truth was both a useful notion and did not carry such implications.

Carnap's reaction could hardly have been more different. When Tarski described his method for defining truth to Carnap he felt that the scales had fallen from his eyes. Afterwards he embraced semantics and went on to write some of the most influential books on semantics of the twentieth century. Surprisingly, it was not the rather ingenious logical machinery Tarski had used in his definition that Carnap felt was so revolutionary. Rather it was use of the seemingly trivial T-schema as a criterion of material adequacy. For in his (1934) Carnap had most of the logical machinery required for Tarski's definition already in place. Thus in many ways, the step from syntax to semantics was not a major one for Carnap. Yet, for a time, Carnap was blind to the possibility of giving a definition of truth in terms of the conditions that each sentence had to meet in order for it to be true.[33] It took Tarski to show him the way truth could be so defined and to make him enforce the distinction between truth and confirmation.[34]

The difference between the reactions of Carnap and Neurath shows both that Tarski's T-sentences are helpful in showing which issues are about truth and which issues are not, and also how difficult it can be to heed the lesson. In this case, the T-sentences show quite clearly that truth is not the same as confirmation, that we can abandon the idea of absolute knowledge and yet happily endorse the use of the truth-predicate. For in claiming that 'Snow is white' is true iff snow is white we maintain only that whatever uncertainty there is in the claim that snow is white applies also to the attribution of truth to the sentence 'Snow is white'. Similar things can be said about the independence of reality from the language we use

[31] There is much more to be said about the philosophical significance of Tarski's definition. Some of this will emerge as we continue.

[32] See, for example, the last sentence of Neurath 1937.

[33] The exact reason for the blindspot is hotly debated amongst commentators. Coffa 1991 claims that Carnap's verificationism led him to look for a syntactic criterion for truth. Ricketts 1996 claims that it was Carnap's single-mindedness in giving a complete definition of truth that led him to overlook a materially adequate definition.

[34] See Carnap 1949 for the classic attack on the conflation. Carnap does not mention his former self as one of those guilty of making the conflation.

to ask it questions.

2.4 Correspondence vs redundancy: the Austin/Strawson debate

One of the central issues in the extended debate between J. L. Austin and P. F. Strawson turned on the question of whether yet another issue, namely the theory of meaning, is in fact properly considered part of the theory of truth.[35] As we have seen, Tarski's definition of truth is capable of being turned into a correspondence theory of truth. The fact that it can be so developed stems from its basis in what Tarski calls the semantic conception of truth – a sentence S is true iff S says that some state of affairs obtains and that state of affairs obtains. As we have seen, an attentive correspondence theorist is likely to claim that this meaning relation that is assumed to hold between the sentence and a state of affairs is the correspondence relation they have been attempting to articulate. Thus the crucial question between a redundancy theorist like Ramsey and a correspondence theorist of this variety is whether an analysis of the meaning relation belongs to the theory of truth.

Austin's correspondence theory seems to be a disguised version of the latter, two-stage, approach to the theory of truth. However, it is important to note that Austin shared the logical positivists' taste for attempting to tackle metaphysical questions through an analysis of language:

> We approach ['truth'] cap and categories in hand: we ask ourselves whether Truth is a substance … or a quality … or a relation … . But philosophers should take something more nearly their own size to strain at. What needs discussing rather is the use, or certain uses, of the word 'true'.
>
> (Austin 1950: 149)

For Austin, however, the theory of truth should proceed by giving an analysis of our *ordinary* uses of 'true'. The motivation for such an approach is the idea that concepts like *truth* are falsely taken as enigmatic by philosophers who have paid insufficient attention to the way the word 'true' is really used. That said, we will largely ignore Austin's linguistic analyses and merely describe the theory of truth that he ends up endorsing.

Austin complained that previous correspondence theories made two basic mistakes. First, they supposed that the correspondence relation was some sort of real relation between truth-bearers and facts such as an isomorphism or congruence relation.[36] Second, and relatedly, previous correspondence theories incorrectly populated the world with "linguistic *Doppelgänger*" which could stand in such real relations to truth-bearers.[37] In contrast, he urged, the correspondence relation is "*absolutely and purely* conventional."[38] This is because what our words mean is a matter of convention. Once we appreciate this, Austin claims, there is no need to posit facts with the same structure as sentences. Instead, we can say that for a truth-bearer to be true is for there to be some convention that determines that the truth-bearer

[35] The debate between Austin and Strawson began in 1950. It marks the emergence of the continuing feud between deflationary and correspondence theorists.

[36] Austin 1950: 154ff.

[37] ibid. 154.

[38] loc. cit. Austin's italics.

means that p and for it to be a fact that p. To understand Austin's view in more detail we need to follow him in spelling out the nature of his truth-bearers and conventional relations.

Austin is quite clear that the bearers of truth are neither sentences nor propositions nor beliefs but statements. What is less clear is what statements are meant to be. Austin describes a statement as something that is made (and so is an historic event) in words, but also that one sentence can be used to make different statements (you say 'It is mine' and I say 'It is mine').[39] While it is difficult to be sure, it seems as if Austin means that a statement is an utterance-as-understood, or a pair consisting of an utterance and its interpretation.

Central to Austin's account of the conventional nature of correspondence was the claim that there are two importantly different types of linguistic convention. Descriptive conventions relate sentence types (not statements) to a certain type of state of affairs. Demonstrative conventions relate statements to particular, historic states of affairs. To see the difference between these two conventions, consider a stop sign. There are descriptive conventions that tell us that this sign is a *stop* sign; that the type of thing this sign is telling us to do is stop. But there is also a demonstrative convention that tells us where we should do this – namely in the area that the sign has been planted.[40] Austin's claim is that statements have a similar duality. On the one hand they refer to a particular state of affairs and on the other they assert that this particular state of affairs is of a certain type. This gives us an obvious way of defining truth.

> A statement is said to be true when the historic state of affairs to which it is correlated by the demonstrative conventions (the one to which it 'refers') is of a type with which the sentence used in making it is correlated by the descriptive conventions.
> (Austin 1950: 152f)

Another way of putting the same point would be to say that a statement is correlated with a set of states of affairs and a particular state of affairs and the statement is true if the latter is a member of the former. This analysis of correspondence truth also makes a simple account of falsity available that avoids the sort of worries about meaningful falsehoods that Russell struggled with. On Austin's account, statements always refer to actual, real states of affairs. False statements simply assert that they are of a type that they are not.

On the surface, Strawson could hardly have had more to disagree with in Austin's theory. He rejected Austin's account of truth-bearers, truthmakers, the correspondence relation and the types of linguistic convention. As so often, though, their differences were sometimes only apparent.

The key to much of Strawson's dislike for Austin's theory lies in Austin's attempt to elucidate truth using the words 'fact' (or 'states of affairs') and 'statement'. According to Strawson, along with 'true' these words are all part of a linguistic practice, or a particular way we have of communicating with each other, which we can call fact-stating discourse. Strawson's worry is that any attempt to use some of these words to elucidate the others will result in nothing but vacuous truisms. The idea seems to be that one grasps the way to use these inter-definable words all at once as one is introduced into the practice of fact-stating.

---

[39] ibid. 151.

[40] Austin suggests the example of traffic signs in 1950: 153, note 10.

Strawson's claim is thus that trying to elucidate truth by appealing to these cognate concepts is really an attempt to elucidate truth through elucidating this whole practice. But this elucidation cannot get anywhere if we use notions like 'fact' and 'statement' because if we did so we would find that the 'words occurring in the solution incorporate the problem.'[41] Thus Strawson found it no surprise that Ramsey discovered that if we have first got clear about what a statement (or judgement) is there is no further problem about truth. Presumably Strawson thought that getting clear about the nature of making a statement involves understanding that, within the practice of stating facts, statements aim at the truth.

Strawson's famous attack on the correspondence account of facts exemplifies these rather abstract claims. He began by urging that, *pace* Austin, we do not use statements to refer to facts. If anything, statements *state* facts. However, even if Austin were to give up the claim that statements refer to facts, Strawson argued, there are no such constituents of reality. Facts are just the shadows of true statements. For one, we have no way of individuating facts except for using the statements which the correspondence theory claims are made true by the fact: we can refer to the fact that John is driving only by using the statement that John is driving. It is thus no more of a surprise that statements correspond to facts than it is that there is no further problem of truth once we have understood what it is to make a statement. This is the point of Strawson's oft-quoted remark 'Of course statements and facts fit. They were made for each other.'[42] But it is important to see this claim in the right light. On the one hand, it is meant to suggest the truistic nature of any claims about true statements' corresponding to the facts. But it is also meant to help in throwing doubt on the existence of facts. Because we have no other way of individuating facts except by using the statements that they make true, and because they serve no other role than to be 'the tautological accusatives' of true statements',[43] it seems gratuitous to suppose that such things as facts are part of language-independent reality.

It is easy to miss the seriousness of this attack. Austin, for one, complained that Strawson had been unfair to facts because we can equally well point out that targets and well-aimed shots are made for each other and yet we do not suppose that targets are bogus entities.[44] In making this parallel Austin rightly argues that it is not sufficient to point out that facts are the internal accusatives of true propositions in order to conclude that the former are the mere shadows of the latter. Yet Strawson's criticism is not meant to rely solely on this point. In the case of targets and well-aimed shots, targets can be set up as the things at which to aim because they can be individuated without having to fire a shot. Further, they can be such that any number of shots are considered well-aimed in virtue of hitting that target. Strawson's worry is that neither of these features is shared by facts. Austin has not shown us how to individuate a fact as the thing to which we should try to get a statement to correspond; nor has he given us reason to think that it is not the case that each true proposition has its own fact that makes it true. However, on this last point, notice that Austin's *statements* (unlike Strawson's) are historical occurrences and so, as for well-aimed shots and targets, any number of them could

[41] Strawson 1950: 171.
[42] ibid. 168.
[43] Armstrong 1997: 19.
[44] Austin 1961: 188.

be made true by the same fact. Here the difference in Austin's and Strawson's choice of truth-bearers seems to have resulted in their talking past each other. For when Strawson claims that facts are the shadows of true statements, he does not mean that they are the shadows of historical occurrences.[45]

This difference between Austin and Strawson leads to Strawson's most serious criticism of the Austinian project. For Strawson claims that Austin's approach rests on a basic confusion between, on the one hand, the conditions that must obtain for our attributions of truth to be true and, on the other, what we actually assert when we make an attribution of truth. By focusing on the first issue, Strawson claims that Austin is led to an attempt at elucidating fact-stating discourse by disclosing the conventional relations between our statements and the world. But, according to Strawson, the real issue is what we assert in making an attribution of truth and we do not assert that such conventional relations hold.[46] Instead, Strawson says, we should focus on the role of 'true' within fact-stating discourse and not attempt to step outside this type of discourse to explain truth. In saying this, Strawson repeats an objection to the correspondence theory that we have found being made by Joachim, Blanshard, the pragmatists and the logical positivists – namely, that the correspondence theory attempts to step outside language or the mind to compare language to the world when the right (and only possible) approach is to attempt to understand truth from within this discourse. We will see Quine make the same sort of remark.

Strawson's conclusion, then, is that Austin has illicitly included claims about the nature of representation into the discussion of truth. As such, Strawson's attack clearly generalizes to all theories that attempt to explain correspondence by focusing on representation. Notice, though, that the difference between Austin and Strawson on the role of the theory of meaning may well reduce to the differences in their choice of truth-bearers. For if Strawson is right and we should treat propositions as the bearers of truth, then it seems to be entirely superfluous to add a discussion of how it is that our sentences get to express the propositions that they do. However, if Austin was focusing on sentences as the bearers of truth then perhaps a case could be made for the inclusion of such a discussion, as Strawson appeared to recognize:

> If someone wishes to contend that we do not *really*, or do not *fully*, know the meaning of 'is true' unless we know what types of conventional relation obtain between words and things when something true is stated or otherwise expressed in words, then the contention seems to me by no means extravagant. …
>
> Better, perhaps, let the theory of truth become, as it has shown so pronounced a historical tendency to become, part of some other theory: that of knowledge; or of mind; or of meaning.
> (Strawson 1964: 232 and 233)

Here Strawson articulates another common theme, this time that of the deflationist. Those who follow Ramsey's approach to truth see the theory of truth as fairly straightforward. It is only by mixing up truth with other topics that we come to see it as problematic.

---

[45] Strawson spells out his conception of statement at 1950: 162-5.

[46] ibid. 172ff.

In fact, Strawson himself championed a Ramsey-style approach to the nature of truth. Focusing on how we use 'true', Strawson agreed with Ramsey that to state that p is true is to make no further statement than would be made by stating that p, but unlike his predecessor Strawson made much of the idea that we use 'p' and '"p" is true' differently.[47] He claimed that sentences containing 'true' are much like performatives such as 'I promise to clean the house.' So that although we assert nothing more than that p with the statement that p is true, we do more than just assert that p. We also perform the speech act of endorsing or confirming someone's statement. According to Strawson, attributions of truth require someone to have first made the statement to which we are attributing truth so that we can confirm or endorse it. In his earlier presentations of this view (1949, 1950) Strawson argued that this fact tempts people to the view that we are attributing a property to a statement. However, we are not attributing truth to anything, we merely require an appropriate context, namely someone's making an assertion, before we can signal our agreement. This view can look as if it may be required by any Ramsey-style, or deflationary, theory of truth. For if "'p' is true" asserts no more than 'p', then, given that 'p' is not about a statement, our "attributions" of truth should not be either.

Austin, for one, rejected the idea that we do not use 'true' to say something about statements. The rejection is *prima facie* reasonable, as the move from (1) and (2) to (3) seems natural:

(1) John said that Brian is sleeping.
(2) It is true that Brian is sleeping.
(3) Something John said is true.

But how could we make sense of this style of reasoning if (2) was not about anything at all? In any event, Strawson backed away from this view in later writings. In his 1964, for example, he went to some length to argue that deflationary theories are in fact compatible with the claim that attributions of truth are in fact statements about statements. He suggested that we make the following paraphrases of statements that attribute truth:

(4a) John's statement that p is true.
(4b) As John stated, p.
(5a) It is true that p.
(5b) As may be urged or objected or … , p.

Strawson's hope in offering these paraphrases was that he could provide us with statements that (i) did not use 'true', (ii) were equivalent to statements that did use 'true' and (iii) were about (in some sense of 'about') statements. (5b) in particular, however, seems a strange paraphrase of (5a) precisely because (5a) does not seem to be making a claim about what may be urged or objected or … . Yet for Strawson's performative approach to work, (5a) must be read in such a way because Strawson insists that such attributions, if they are about something, are about actual or possible acts of stating.

[47] Ramsey (1923: 142) said only that we use 'true' for different emotive or stylistic reasons or to indicate the position of the proposition in an argument. Carnap (1942: 26) also claims that while 'p' and "'p' is true" mean the same in a logical or semantic sense, the two expressions have 'different features and different conditions of application; from this point of view we may e.g. point to the difference between these two statements in emphasis and emotional function.'

Strawson's comments about the role of 'true' can be seen as a suggestion about the utility of the truth-predicate. If attributions of truth say no more than the statement to which they attribute truth, then why do we have a truth-predicate at all? Strawson's suggestion is that the predicate is useful as a performative device. One question that should be asked of any such approach is whether it can do better than Ramsey's in dealing with the use of 'true' in statements like 'Everything the Pope says is true'. If 'true' is regarded purely as a performative device, it is difficult to see how to read such statements. Strawson's original suggestion was something like (6):

(6) The Pope has made some statements. I confirm them all.

But Strawson's view is that we do not actually *assert* that we confirm them; rather we just confirm them. This entails that the only thing asserted by (6) is that the Pope has made some statements, which of course is true. But, as Soames points out, this means that even if some things the Pope says are false, according to Strawson's analysis the statement 'Everything the Pope says is true' will come out true.[48] In the light of such difficulties, Strawson later altered his theory by de-emphasizing the confirming use of 'true' as only one use among others. The analysis he later gives of these problematic statements seems to follow Quine.

2.5 Quine and disquotation

Like those of the logical empiricists, Quine's views on truth were shaped by a distrust of semantic concepts. In particular, Quine is famous for his strong claim that there is no such thing as objective, interlinguistic synonymy, and that it only makes sense to compare the meanings of expressions that belong to the same language. He was led to this conclusion by his commitment to meaning holism, which, like Joachim and Blanshard, Quine used to reject the notion of individual propositions associated with individual sentences. However, if there is no isolable meaning for each sentence, then the correspondence theory is threatened: we cannot say that a sentence is true iff it represents some particular state of affairs as obtaining and that state of affairs obtains.

Continuing his agreement with the early coherence theorists, Quine insisted that only whole theories can be compared with reality.[49] This confirmation holism, which in part drove his meaning holism, was a crucial feature of his attack on the distinction between analytic and synthetic truths.[50] No truth, he claimed, was immune to revision in the event of a theory's being confronted with a recalcitrant experience. The difference between the sort of statements that had been labelled analytic and those that had been labeled synthetic was thus merely one of degree. Although this attack on the analytic/synthetic distinction was largely directed at the logical empiricists, removing this distinction from their theories of truth would have removed a worrying tension in their views that we remarked on earlier (§2.1, opening paragraph), namely, that between their truth-functional account of molecular propositions and their claim that analytic propositions are neither true nor false.

[48] Soames 1999: 236f.

[49] Quine 1970: 1-8, spells this out clearly.

[50] See Quine 1953.

Yet Quine's meaning holism also led him to reject epistemic theories of truth like coherence and pragmatic theories. As we have seen, such theories are constantly in danger of succumbing to relativism. To avoid this, their proponents often resort to the idea of an ideal system of beliefs (perhaps the set of beliefs we would adopt at the end of inquiry or perhaps as that ideally coherent set that is identical with reality). Any particular belief is then said to be true iff it is a member of this ideal set. As Quine points out, a serious problem arises if we combine this approach with meaning holism (as Joachim and Blanshard did). For if we cannot ask about synonymy interlinguistically (intertheoretically), then we cannot make sense of the idea that a belief that belongs to our current theory also belongs to some ideal theory.[51] For no two beliefs (sentences) that belong to different sets of beliefs (theories) can have the same content. Thus meaning holism undermines epistemic theories of truth as well.[52]

In contrast to both coherence and correspondence theorists, Quine urged that truth was 'immanent'. In other words, the truth-predicate can be meaningfully applied only to sentences within our own language. He argued that to the extent that foreign sentences can be called true, this is only relative to some translation scheme that we employ to translate these sentences into sentences of our language. Yet, once truth is treated as internal to a language and in consequence (because Quine thought that people with different theories thereby speak different languages), as internal to a theory, problems about its nature dissolve in the way a deflationary reading of Tarski would suggest. With an empiricist's distrust of propositions, Quine takes sentences as truth-bearers, though, to avoid problems concerning indexicals and other context-dependent expressions, he is careful to define a class of sentences he calls 'eternal', which function as the primary bearers of truth. Eternal sentences are sentence types whose tokens always have the same truth-value. For example, while 'I am hungry' is not an eternal sentence, 'James Brown is hungry as at 10.15 p.m. on 13th August 1968' supposedly is. For these sentences, Quine claimed that truth is disquotational. An attribution of truth to a sentence merely undoes the effects of quotation marks that we have used to form a name for the sentence. As Ramsey noted, this means that in simple contexts the truth-predicate is redundant. This is seen in Tarski's T-schema, and more clearly in what has come to be called the disquotation schema:

(DS) 'P' is true iff P.

Whereas the original T-schema allowed an arbitrary name for the sentence, the disquotation schema requires that the name be formed by the addition of quotation marks.

Where Ramsey seems to have gone wrong is in not noticing that the difficult cases for the redundancy view actually show the *raison d'être* of the truth-predicate. For in cases of generalization (like 'Everything the Pope says is true' or "All instances of 'If p, then p' are true") we cannot dispense with the truth-predicate precisely because these are the cases in which it does its work. In these cases we want to generalize in a way analogous to the generalization from a sentence like 'Socrates is mortal' to get 'For all x, if x is a man, then x is mortal'. However, we cannot do this because the move from 'If time flies then time flies'

[51] Quine 1960: 23f.

[52] Of course, a pragmatist or coherentist could accept the charge of relativism and thus accept both meaning holism and an epistemic theory of truth. Rorty sometimes seems to be happy to adopt this position, for example.

to ‘For all x, if x then x’ ends with a string of expressions that is incoherent if the quantifier is read objectually. The beauty of the truth-predicate is that it allows us to make these generalizations:

> We could not generalize as in ‘All men are mortal’, because ‘time flies’ is not, like ‘Socrates’, a name of one of a range of objects (men) over which to generalize. We cleared this obstacle by *semantic ascent*: by ascending to a level where there were indeed objects over which to generalize, namely linguistic objects, sentences.
>
> (Quine 1990: 81; his italics)

Because ‘x’ is equivalent to ‘‘x’ is true’, ‘If time flies then time flies’ is equivalent to ‘If ‘time flies’ is true, then ‘time flies’ is true’. But the final sentence can be used to create a coherent generalization: in fact, just the generalization which seemed at first to pose a problem for a disquotational view of truth. According to Quine it is thus the transparent or disquotational function of ‘true’ that allows us to make a technical, semantic ascent to talk about sentences, while still talking about the world. With this suggestion about the utility of the truth-predicate, Quine’s disquotationalism purports to solve two of the more difficult problems facing deflationary accounts of truth. On the one hand, it explains why it is we have a truth-predicate if its uses are often redundant. On the other, Quine’s explanation of the utility of the predicate explains how ‘true’ functions in precisely those cases of generalization that proved a stumbling block for both Ramsey and Strawson.

Thus disquotationalists treat ‘true’ as a meta-linguistic predicate applying to sentences. Doing this allows us to explain generalization contexts without appealing to any non-standard understanding of the quantifiers. However, as Quine essentially admits by focusing on eternal sentences, disquotationalism has difficulty with a vast range of sentences of natural language. For example, treating “‘I am hungry’ is true” disquotationally implies that it is equivalent to ‘I am hungry’. Of course, this implication means that we cannot use the disquotational account to describe the conditions under which some other person’s utterance of ‘I am hungry’ is true. In attempting to do so, we would be forced to say that their utterance is true iff *I* am hungry. Another difficulty arises if the disquotation schema is treated as expressing a necessary equivalence, as seems required if disquotationalism is to be in conflict with inflationary theories such as correspondence or coherence. This reading of the schema implies that the truth-conditions of sentences belong to them necessarily. In other words, in a possible world where we used the sentence ‘Rabbits are furry’ radically differently, and so the sentence’s meaning differed from that in this world, the sentence would still be true iff rabbits are furry in that world. Intuitively, though, it seems that sentences, as opposed to propositions, have their truth-conditions contingently. These problems, and the closely related problems associated with the disquotationalist’s claim that ‘true’ is confined to a particular language, suggest that the simple disquotationalist account will need to be modified if it is to capture the ordinary meaning of ‘true’.[53]

2.6 Dummett

Michael Dummett, however, raised a number of further difficulties for the deflationary approach in the course of mounting a defence of a verificationist theory of truth inspired by

[53] See David 1994, ch. 5, for a discussion of the more pressing problems facing the disquotationalist.

intuitionist and constructivist accounts of mathematics.[54]

Dummett expressed his reservations about deflationary theories of truth as concerns about the attempt to define truth in terms of the equivalence thesis:

(ET) <It is true that A iff A>.[55]

Each of these concerns has been influential in the ongoing debate about deflationary theories of truth. But each rests on some controversial, though not necessarily implausible, assumptions. First, Dummett argued that the equivalence thesis runs into contradiction if we deny the law of bivalence.[56] The simple argument is that if a statement, say A, is neither true nor false, then it is false that A is true. However, this means that while <A> is neither true nor false, <it is true that A> is false, and so <A> and <it is true that A> are not equivalent. This argument draws on the obvious inconsistency between the following:

(7) Some statements are neither true nor false (Denial of Bivalence)
(8) For all statements A, A is true iff A (Equivalence Thesis)
(9) For all statements A, A is false iff not A.

Of course, whether or not someone like Quine should worry about this inconsistency depends on whether one thinks that (9) reflects an adequate conception of falsity and whether one thinks that we should abandon bivalence. We shall explore Dummett's verificationist reasons for abandoning bivalence below.[57]

Dummett's second influential criticism is the now familiar one that the redundancy theory does not account for truth's being a normative goal of assertion. [58] And he insists that the redundancy theory cannot be fixed by simply adding the claim that in making an assertion we aim at the truth. What he requires by way of supplementation 'is a description of the linguistic activity of making assertions; and this is a task of enormous complexity.'[59] In other words, he is demanding what Strawson denied was the task of a theory of truth, namely the analysis of fact-stating discourse. However, it is worth recalling from §1.5 that Ramsey had already offered an explanation of the normative role of truth that is consistent with a redundancy theory of truth. If one provides a version of a 'success semantics' then it is quite clear why we aim at the truth: true beliefs are those that lead to successful action. Dummett's demand had been met before it had been made.

Dummett's concern that the activity of asserting has been neglected leads fairly directly to his third influential claim about the redundancy theory. He argued (or rather claimed it was too obvious to be able to argue) that the redundancy theory was incompatible with truth-

---

[54] Dummett 1959. Although he ends this paper by rather strangely suggesting that he adhered to a version of the redundancy theory, the commitment fitted badly with much of the rest of that paper. He later admitted that this way of characterizing his position was misleading. See his 1978: xxii.

[55] Dummett 1973: 445. Dummett uses corner quotes to express the thesis. For convenience, we have substituted angle brackets.

[56] Dummett 1959: 4-7. See also Dummett 1973: 445-446.

[57] Notice that Dummett only thinks we should abandon bivalence and not deny it. He distinguishes between (7) and the rejection of the claim that every statement is either true or false.

[58] Dummett 1959: 2f.

[59] Dummett 1978: 20.

conditional theories of meaning; that is, theories of meaning that took truth as their fundamental concept. Traditionally, he claimed, theories of truth have tried to do more than spell out the conditions under which assertions are true. They have tried to give an account of truth that would serve as the basis for a theory of meaning. However, if one tries to explain 'true' by appealing to instances of the equivalence thesis, one cannot then use these instances of that thesis to explain the meaning of the sentences. For one can understand the explanation of 'true' only if one already understands the sentence that is used on the right hand side of the equivalence thesis:

> But in order that someone should gain from the explanation that P is true in such-and-such circumstances an understanding of the sense of P, he must already know what it means to say of P that it is true. If when he enquires into this he is told that the only explanation is that to say that P is true is the same as to assert P, it will follow that in order to understand what is meant by saying that P is true, he must already know the sense of asserting P, which was precisely what was supposed to be being explained to him.
> (Dummett 1959: 7)

Dummett has argued that this problem is especially apparent if we consider a Tarskian truth-definition in which the metalanguage is not an extension of the object-language.[60] If we do not understand the object-language, then this sort of truth-definition will not help us understand its sentences because we will have no idea what is accomplished by pairing them with truth-conditions: the appeal is to his second objection, that deflationary theories cannot tell us the point of calling sentences true. Furthermore, these theories seemingly assign truth-conditions to sentences without basing these assignments on the way each sentence is used. This purported feature of deflationary theories of truth-conditions would make them implausible theories of meaning.

Dummett's conception of theories of meaning is crucial to his famous argument(s) for anti-realism. However, the connections between this conception and his conception of realism are tortuous; our account must be brutally brief. He characterizes realism as the belief that statements 'possess an objective truth-value, independently of our manner of knowing it: they are true or false in virtue of a reality existing independently of us.'[61] Dummett thus treats realism as a doctrine about truth, namely, that it is recognition-transcendent. He also urges that this recognition-transcendence is necessary to guarantee the principle of bivalence. That is, in cases where we have no means of determining whether a sentence is true or not (where no condition of which we are capable of becoming aware determines whether or not the sentence is true), it is only if there is some recognition-transcendent condition which, if met, determines that the sentence is true (and if not that it is false) that the sentence must be either true or false. In other words, without recognition-transcendence, bivalence, and hence the objectivity of truth, is at risk.[62]

Without a guarantee for the principle of bivalence, however, so Dummett argues, we cannot adopt a truth-conditional semantics. He argues at length that the nature of assertion is such

---

[60] Dummett 1978: xxi.
[61] ibid. 146.
[62] Dummett 1959: 14ff.

that we only speak of an assertion's being correct or incorrect and have no need and no room for the idea that there is a third possibility that lies between these poles.[63] The reason he concentrates on this supposed feature of assertion is this: if the principle of bivalence cannot be guaranteed, and so when it comes to truth there can be more than two possibilities to consider, then truth-conditional semantics cannot be the right approach to the theory of meaning. (At least, not if truth-conditions are understood recognition-transcendentally.) Thus, Dummett suggests, realist/anti-realist debates about a particular subject matter ought to be reformulated as debates about what notion of meaning, and, because meaning should be understood truth-conditionally, what notion of truth is appropriate for a particular area of discourse.[64] Accordingly, the question as to the nature of truth-in-general is better formulated as 'What is the notion of truth appropriate for this particular area of discourse?' We shall see that the pluralism about the nature of truth implied by this line of thought has been developed more recently by Crispin Wright and others.

The general structure of Dummett's argument for anti-realism is that a recognition-transcendent notion of truth-conditions is untenable.[65] We have already seen that he conceives a theory of meaning as a theory that specifies for each sentence the conditions under which the understanding of the sentence counts as being manifested. If the obtaining of truth-conditions of a sentence is beyond our recognitional capacity, he claims, we could never manifest our understanding of these sentences (the manifestation argument). Moreover, we could never be taught the meaning of these sentences to begin with (the acquisition argument).[66] He concludes that we should adopt instead a notion of truth that is relative to our epistemic powers: in short, a verificationist conception of truth. The right notion of truth-conditions to use for a theory of meaning is therefore a verificationist notion of truth-conditions.[67] This allows us to abandon bivalence, because some statements will be neither verified nor have their negations verified. However, an assertion will still only be either correct or incorrect, because every assertion will be either verified or not and we can thus say an assertion is correct if and only if it has been verified. The compatibility of these two claims is ensured by the fact that, unlike truth, verification is such that the fact that a statement is not verified does not entail that its negation is verified. For example, although we have not verified the claim that there is life on the planet Venus, this certainly does not entail that we have verified that there is not life on Venus.

---

[63] Dummett 1959. This is not to suppose that Dummett rejects outright the possibility of three truth-values. He claims that we may need to appeal to more than two truth-values to explain the way the meaning of complex sentences depends on the meaning of atomic sentences. However, he later (1973: 446f) distinguishes between the "ingredient sense" (the sense relevant to its contribution to more complex sentences) of a sentence and its content. As far as a sentence's content is concerned only two possibilities can be of concern.

[64] The situation is a bit delicate here. Dummett argues that if we give up recognition-transcendent truth-conditions we must give up the idea that truth is the *central notion* in a theory of meaning. However, identifying truth with verification still allows us to treat meaning as truth-conditional in the sense that meaning is identical with verification conditions. More on this below.

[65] See Dummett 1976 for an extended discussion of his conception of theories of meaning.

[66] A brief version of this line of thought is presented at 1959: 17. See also Dummett 1973: 467f.

[67] Although Dummett sometimes characterized his claim as being that we should abandon truth-conditional theories of meaning he later realized that what he meant was that we need to shift conceptions of what a truth-condition is because we need to treat truth as verification.

Dummett's arguments for verificationism about truth have been repeatedly criticized. It is worth pointing out just a few of the more salient of their controversial assumptions. For one, his acquisition argument relies on the assumption that to learn the meaning of a sentence we must be acquainted with the conditions under which it is *conclusively* verified.[68] Dummett here adopts an extreme version of verificationism about meaning that even the logical positivists were forced to abandon. He himself relinquished this position too,[69] but it is unclear how his argument survives without it. He also rejects Quine's arguments for meaning holism. For it is crucial to Dummett's manifestation argument that each statement can be treated as having its own verification conditions and that a theory of meaning can spell out the conditions under which understanding of a particular sentence can be manifested. Whatever the merits of Quine's arguments in general, it is difficult to see how one can deflect their force once the verificationist background is accepted.

Finally, it is worth noting that verificationists about meaning are not *ipso facto* verificationists about truth. In particular, even if Dummett is right that the meaning of a statement is the conditions under which it is verified, this entails a claim about truth only if we take the meaning of a statement to be the conditions under which it is true. For example, there is nothing preventing a verificationist about meaning from following Ayer in rejecting truth-conditional theories of meaning and embracing a redundancy theory of truth. Nor is it clear that Dummett's arguments for anti-realism should be construed, as he does construe them, as arguments about the nature of truth. Although we cannot argue for it here, it seems to us, and to many others, that realism (or anti-realism) is a metaphysical addition to semantics and not built into it.[70] This seems to have been Quine's point when he claimed that from within the language there is nothing more to say about truth than to point out its disquotational properties. Perhaps, then, the outstanding question is whether the right perspective to take on the nature of truth is from inside or outside our language. Indeed, this question proved important to a number of theorists of truth in the latter stages of the century.

[68] This is made clear at Dummett 1976: 132.

[69] Dummett 1978: xxxviii.

[70] For a brief and clear argument to this effect see Soames 1999: 32-9.