Truth and meaning

Rumfitt, Ian

DOI:
10.1111/j.1467-8349.2014.00231.x

TRUTH AND MEANING

Ian Rumfitt

In *Proceedings of the Aristotelian Society Supplementary Volumes* **88** (2014): 21-55

Abstract: Should we explicate truth in terms of meaning, or meaning in terms of truth? Ramsey, Prior, and Strawson all favoured the former approach: a statement is true if and only if things are as the speaker, in making the statement, states them to be; similarly, a belief is true if and only if things are as a thinker with that belief thereby believes them to be. I defend this explication of truth against a range of objections.

Ramsey formalized this account of truth (as it applies to beliefs) as follows: B is true $=_{\mathrm{df}}$ $(\exists P)$(B is a belief that $P \wedge P$); in §I, I defend this formula against the late Peter Geach's objection that its right-hand side is ill-formed. Davidson held that Ramsey & co. had the whole matter back to front: on his view, we should explicate meaning in terms of truth, not *vice versa*. In §II, I argue that Ramsey's approach opens the way to a more promising form of semantic theorizing than Davidson's. Ramsey presents his formula as a definition of truth, apparently contradicting Tarski's theorem that truth is indefinable. In §III, I show that the contradiction is only apparent: Tarski assumes that the Liar-like inscription he uses to prove his theorem has a content, but Ramsey can and should reject that assumption. As I explain in IV, versions of the Liar Paradox may be generated without making any assumptions about truth: paradox arises when the impredicativity that is found when a statement's content depends on the contents of a collection of statements to which it belongs turns pathological. Since they do not succeed in saying anything, such pathological utterances or inscriptions pose no threat to the laws of logic, when these are understood as universal principles about the way things may be said or thought to be. There is, though, a call for rules by following which we can be sure that any conclusion deduced from true premisses is true, and hence says something. Such rules cannot be purely formal, but in §V I propose a system of them: this opens the way to construction of deductive theories even in circumstances where producing a well-formed formula is no guarantee of saying anything.

I want to elaborate and defend a thesis once put forward by a local hero. In *On Truth*, a book he drafted in 1927-9 and which appeared some sixty years after his death (Ramsey 1991), F.P. Ramsey takes the primary bearers of truth and falsity to be *beliefs* or *judgements*—terms he uses synonymously to apply to particular mental states which have both propositional content and 'some degree of the affirmative character' that is present in judging that such-and-such is the case but absent from merely wondering whether it is the case (*op. cit.*, p.8). He then proposes a definition of truth as it applies to beliefs in this sense:

> Any belief whatever we may symbolize as a belief that *p*, where '*p*' is a variable sentence just as '*A*' and '*B*' are variable words or phrases (or terms as they are called in logic). We can then say that a belief is true if it is a belief that *p*, and *p*. In Mr Russell's symbolism
>
> B is true :=: $(\exists p)$. B is a belief that p & p. Df' (*op cit.*, p.9, incorporating p.15, *n*.7).

Similarly, in an updated symbolism, B is false if and only if $(\exists P)$(B is a belief that $P \wedge \neg P$) (*op. cit.*, p.11). These purported definitions of truth and falsity, and what they presuppose about the contents of beliefs and other bearers of truth, are my topic today.

The present account of truth differs radically from the redundancy theory that Ramsey had advanced in 'Facts and Propositions', his celebrated contribution to the Joint Session of 1927.\[1]/ In that paper, Ramsey had taken the primary bearers of truth and falsity to be propositions—i.e., the contents of beliefs or judgements. Thus my particular belief that Caesar was murdered is true by virtue of the truth of the proposition that Caesar was murdered. Ramsey also held, though, that apparent reference to propositions dissolves under logical analysis. In particular, 'The proposition that Caesar was murdered is true', which appears to ascribe truth to a proposition, 'means no more than that Caesar was murdered' (Ramsey 1927, p.38). This is the basis for his famous claim that 'there really is no separate problem of truth but merely a linguistic muddle' (*ibid.*).

The theory advanced in *On Truth* is quite different, and the differences stem from Ramsey's changed view of what are the primary bearers of truth.\[2]/ Particular states of belief are not mere logical constructs, apparent reference to which dissolves under analysis. To the contrary, they are elements in

---

[1] Ramsey's respondent was G.E. Moore, who found problems with, and obscurities in, pretty well all the central contentions of his paper. Moore's objections may have prompted Ramsey to develop the very different account of truth that we find in his draft book.

[2] For more on the differences between Ramsey's two theories of truth, see Rumfitt 2011.

the causal order. In tandem with my desire to tell you about Roman history, it is my belief—call it *B*—that Caesar was murdered that causes me to say 'Caesar was murdered'. Moreover, on Ramsey's later account, an ascription of truth to a belief does not share its content with the belief. The content of *B* is expressed by 'Caesar was murdered'. Ramsey's definition, by contrast, unpacks '*B* is true' as '$(\exists P)$(*B* is a belief that $P \wedge P$)', which clearly means something different from 'Caesar was murdered'. '$(\exists P)$(*B* is a belief that $P \wedge P$)' makes no reference to Caesar, or to murder; instead, it involves an unusual form of quantifier, one that binds a variable that replaces a complete sentence. No such quantifier is to be found in 'Caesar was murdered'.

Ramsey's formula will not succeed in defining truth if that notion must be invoked in explicating '*B* is a belief that *P*', so his definition does not cohere with any account on which a belief's content is constituted by its truth-conditions. Ramsey was aware of this and proposed a pragmatist account of the contents of beliefs: an assertion 'that a man has a belief that such-and-such is the case' is to be cashed out in terms of 'how he would behave' (Ramsey 1991, p.45). As his follower R.B. Braithwaite put it, to believe that *P* (where *P* is contingent) is to be disposed to act as if *P*, in a circumstance where it matters whether *P* (Braithwaite 1933, pp.132-3).\[3]/ We can see why Ramsey took beliefs to be the primary bearers of truth: among those mental states that may be assessed as true or as false, beliefs have the most direct connection with action; from a pragmatist perspective, then, they will be the primary bearers of propositional content.\[4]/

Much of the interest of Ramsey's formula derives from its role in this ambitious reductive programme in the theory of content, but I shall say little about the programme here. Instead, I want to address certain logical problems that confront the formula and thereby threaten to derail the programme before it can get properly under way. Some of these problems pertain to Ramsey's claim that his formula can serve as a *definition* of truth: a famous theorem of Alfred Tarski (1935) appears to show that no such thing is possible. Others cast doubt on the correctness of the equivalence between '*B* is true' and '$(\exists P)$(*B* is a belief that $P \wedge P$)', regardless of whether that equivalence is taken to define truth. I start by considering a challenge to the well-formedness of Ramsey's *definiens* that was thrown down at the 1982 Joint Session.

---

[3] Braithwaite acknowledges a debt to Alexander Bain.

[4] See, though, Mellor 2012 §9 for an attempt to show how dispositions to act also fix the contents of declarative utterances.

I. Geach's objection

However exactly its component quantifier is explained, the *definiens* contains two occurrences of the sentential variable '*P*'. In his provocative paper 'Truth and God', Peter Geach observed that 'such identification of variables as we have here is doubtfully legitimate if there would have to be a shift of meaning in any substitution instance of the formula in which the [sentential] variable is repeated. Frege', he continued, 'powerfully argued that there is indeed a shift of meaning (*Bedeutung*) in such cases. If Frege is right, <Ramsey's definition> must be rejected, or at least is hard to defend' (Geach 1982, p.85).\[5]/

While Geach rejects some details of Frege's theory of indirect discourse, he accepts the key 'negative thesis' that in indirect contexts 'names do not simply name the objects that they ordinarily name' (*op. cit.*, p.86). He illustrates this negative thesis with the following example:

> The Derby winner Running Rein turned out to be a horse Maccabeus disqualified by reason of age from running. In a context like 'Lord George Bentinck discovered that — was four years old', the truth-value might alter according as we inserted the name 'Running Rein' or the name 'Maccabeus': the age of Maccabeus was already well known to the racing fraternity, the age of Running Rein was not…We must say that 'Maccabeus' and 'Running Rein' have not here their straightforward meaning: the meaning they would have, as each naming a certain horse and both of them the same horse, in a context like '— was four years old when he won the Derby'. I think these considerations rob <Ramsey's definition of truth> of its intuitive simplicity and persuasiveness (*ibid.*).

Geach's verdict is uncharacteristically mealy mouthed—not all true philosophical theories are intuitively simple and persuasive—and in fact his objection to Ramsey is flawed. To see this, it helps to compare indirect and modal contexts. Consider 'Caesar was murdered, but it was not necessary that Caesar was murdered'. Not only is this statement readily intelligible, it is true: Caesar could have heeded the soothsayer's warning and stayed away from Pompey's Theatre on the Ides of March. But

[5] At the place indicated, and for the reason given earlier, I have substituted 'Ramsey's definition' for 'the redundancy theory'. Angled brackets in subsequent quotations from Geach mark the same substitution.

we need to assign a richer semantic value to the predicate 'was murdered' to account for the truth of the second conjunct than is needed to account for the truth of the first. To account for the first conjunct's truth, it suffices to assign to 'was murdered' the set of people who have in fact been murdered: the first conjunct is true because Caesar belongs to that set. Such an assignment, though, does not account for the truth of the second conjunct. To account for that, we need to assign a more complex entity to be the semantic value of 'was murdered'—as it might be, the function *f* that maps each possible circumstance to the set of people who have been murdered in that circumstance. It does not follow from this difference, however, that there is any equivocation—any shift in meaning—between the two occurrences of 'was murdered'. Nor does it follow that it is illicit to symbolize our statement in the form '$P \wedge \neg\Box P$'. The semantic value of a substituent for 'P' must be the richer value needed to account for its behaviour inside a modal context, but that richer value yields the simpler value that accounts for the truth or falsity of the unmodalized statement. Thus, in the case where 'P' is replaced by 'Caesar was murdered', the set of people who have in fact been murdered is the result of applying the function *f* to the actual circumstance; once we are given that set, we may account for the truth of the first conjunct precisely as before.

Despite some differences (notably, the modal rigidity of proper names) the case of indirect discourse is parallel in the crucial respect. There is of course less agreement about the shape of a good semantic theory for attributions of belief, knowledge, discoveries, etc., than there is for modal discourse. Following Frege, Geach supposes that 'Bentinck discovered that Running Rein was four years old' is true while 'Bentinck discovered that Maccabeus was four years old' is false; but many philosophers follow Russell in denying that these statements diverge in truth-value.\[6]/ All the same, even if we accept Geach's evaluation of these statements, and even if we accept that a name inserted into the context 'Bentinck discovered that — was four years old' does not *simply* name its customary bearer, it does not follow that there is any equivocation—any shift in meaning—between the two

[6] Care is needed in identifying the proper topic of philosophical debate in this area. It is pretty clear that the maxims we ordinarily go by in reporting the speech and beliefs of others do not place people in relation either to Fregean *Gedanken* or to Russellian propositions; for that reason, there is little mileage in discussing whether *Gedanken* or Russellian propositions best match our pre-theoretic notions of saying or believing the same thing. Those entities are better conceived as constructs, postulated for various theoretical purposes in philosophy, linguistics, and psychology. The proper topic of debate, then, is whether a given construct optimally serves a specified theoretical purpose. It is entirely possible that Fregean *Gedanken* might best serve one such purpose, Russellian propositions another, and indeed Stalnakerian propositions (i.e. sets of possible worlds) a third. For a development of this irenic view, see Moore 1999.

occurrences of ‘Running Rein’ in ‘Running Rein was four years old and Bentinck discovered that Running Rein was four years old’. Nor does it follow that it is illicit to symbolize this conjunction in the form ‘$P \wedge \Phi P$’. As in the modal case, the semantic value of a substituent for ‘P’ must be rich enough to account for its behaviour in the intensional context ‘Φ’, but we can answer the charge of equivocation so long as that richer value also yields what is needed to account for the truth or falsity of the first conjunct. Frege’s own theory of indirect discourse meets this condition. According to that theory, the semantic value of ‘Running Rein’, as it appears in the second conjunct, is the name’s ‘customary sense’.\[7]/ But the theory also says that sense determines reference, so the customary sense of ‘Running Rein’ determines that its customary referent is indeed the horse Running Rein, alias Maccabeus. Because that horse was four years old when Bentinck conducted his inquiry, we can account for the truth of the first conjunct. Whatever special features the name ‘Running Rein’ may possess that distinguish its sense from that of ‘Maccabeus’, it retains those features even when they do not bear directly on the truth-values of statements involving it. There is, then, no reason to doubt the legitimacy of the repeated variables in Ramsey’s formula ‘B is a belief that $P \wedge P$’.

It would, indeed, be astonishing if there were always an equivocation between a free-standing use of a declarative sentence and a use of the same sentence to complement ‘B is a belief that’. For what is it for a belief to be true? One correct answer says: it is for things to be as the relevant thinker, in having the belief, believes them to be. This answer presupposes that there is a way things might be said or thought to be which is at once a way the thinker thinks they are and a way they actually are. Now one way things might be thought to be is the way they are thought to be when someone thinks that Running Rein was four years old. That way was at once a way Bentinck thought things were and a way things actually were. But then a sentence saying that things are that way must be able to serve unequivocally as a complement of ‘Bentinck believed that’ and as a statement of how things are. An example of such a sentence is ‘Running Rein was four years old’. So there need be no equivocation between a free-standing use of that sentence and its use immediately after ‘Bentinck believed that’.

---

[7] Frege’s actual claim was that, in an indirect context, a name’s *Bedeutung* is its customary sense (*gewöhnlicher Sinn*) (Frege 1892, p.28). I follow Dummett in taking ‘semantic value’ to be a good gloss on Frege’s use of the term ‘*Bedeutung*’. Geach’s translation of that term as ‘meaning’ is tendentious in the present context. While Frege (*ibid.*) does speak of words having indirect senses as well as indirect references, he offers no explanation of how an expression’s *meaning* might shift between direct and indirect contexts.

As well as vindicating the denial of equivocation, this reflection provides a first gloss on the quantifier that figures in Ramsey's definitions of truth and falsity. We may read '∃*P*' as 'There is a way things might be said or thought to be' and the attendant variables as 'things are that way' or 'things are thus'. This is essentially the gloss on quantification into sentence position that A.N. Prior proposed in his posthumously published masterpiece, *Objects of Thought* (1971). On this view, it would be mistake to think of '∃*P*' as quantifying over a domain of objects. Ways things might be said or thought to be are not objects, even in the broad logical sense of Frege's *Gegenstände*. They are not objects because 'the way things are thought to be when someone thinks that Running Rein was four years old' is not a genuine singular term or *Eigenname*. Rather, 'There is a way things might be said or thought to be' is tantamount to the adverbial quantifier 'Things might be said or thought to be *somehow*' (Prior *op. cit.*, p.38). Similarly, the corresponding universal quantifier '∀*P*' is best read 'However things may be said or thought to be'.\[8]/ These quantifiers, then, are not 'objectual': they do not range over a domain of objects. Equally, though, they are not substitutional: things might be said or thought to be somehow, even though no sentence in a given substitution class says that things are thus. From a formal point of view, these 'non-nominal' quantifiers are a species of higher-order quantifier. One task of this paper is to investigate the logical properties of this species, but our ordinary understanding of the English quantifiers 'however' and 'somehow' tells us what we are supposed to be investigating.\[9]/

---

[8] '"However he says things are, thus they are" is a very natural rendering of "For all *p*, if he says that *p*, then *p*"' (Prior 1971, p.38).

[9] Scott Soames (1999, pp.39-49) discusses Ramsey's definition under the heading of 'nihilism about truth', and rejects it on the ground that no coherent account can be given of the component existential quantifier. Soames argues first (pp.41-46) that in the context of the proposed definition this quantifier cannot be read substitutionally; I fully agree. He then considers an alternative reading on which it expresses 'higher-order quantification over propositions without a truth predicate' (p.46). Against this, he objects that 'the proponents of the redundancy theory [among whom he names Ramsey and Prior] adopted it not only as a means of eliminating the property <of> truth but also as part of an overall program of avoiding commitment to the existence of propositions' (*ibid.*).

While this may be a fair criticism of 'Facts and Propositions', it does not touch Ramsey's position in *On Truth*. Ramsey certainly wishes to avoid commitment to propositions, but in his draft book he takes truth to be a genuine property of particular states of belief. *Pace* Soames, then, his view is not that 'there is no such property as truth, and the predicate *is true* is not used to describe anything' (Soames *op. cit*,, p.39). Rather, Ramsey aims to define the condition for that predicate to be correctly applicable to beliefs. Moreover, on Prior's reading, the quantifier used in that definition does not 'range over propositions' (p.48), or over any other domain of objects. Instead, it is to be understood as formalizing the English adverbial quantifier 'somehow'. That understanding will be refined as we proceed, but Prior's gloss on '∃*P*' precisely calls into question Soames's unsupported underlying assumption that the 'two standard types of quantification: objectual and substitutional' (p.41) exhaust the possible interpretations of sentential quantifiers.

Ways things might be said or thought to be must conform to the laws of logic. Indeed, Prior's sentential quantifiers may be used in formulating those laws. Thus $\forall P(P \leftrightarrow \neg\neg P)$ expresses the classical law of double negation: however things may be said or thought to be, they are thus if and only if only it is not the case that it is not the case that they are thus. Logical laws in this language-transcendent sense—laws of thought, as we might call them—must be distinguished from the rules we apply in constructing derivations in particular natural or formal languages (more on this in §V).

In Ramsey's definition, the sentential variable '*P*' falls within the scope of the intensional operator '*B* is a belief that'. This, *au fond*, is what worried Geach: 'the ostensible dissolution of the problem about "is true" that <Ramsey's definition> offers is paid for at a dear rate if nasty problems of intensionality are left on our hands' (Geach 1982, p.86). If I am right, then one of those problems—namely, the alleged shift in the meaning of 'Running Rein' between 'Running Rein was four years old' and 'Bentinck discovered that Running Rein was four years old'—is spurious, but there may well be others. We need to assess how nasty the problems of intensionality implicit in Ramsey's definition of truth really are.

II. Truth-theories and meaning

Before turning to this task in earnest, it is worth considering a rather different sort of objection that may be pressed against Ramsey. Under Prior's interpretation of the sentential quantifiers, Ramsey's definition emerges as a semi-formal rendering of a formula that P.F. Strawson recommended on a number of occasions as conveying 'something uncontroversial and fairly general about truth' (Strawson 1971, p.180). A belief, Strawson said, 'is true if and only if things are as one who holds that belief thereby believes them to be'.\[10]/ *Pari passu*, 'one who makes a statement or assertion makes a true statement if and only if things are as, in making that statement, he states them to be' (1971, p.180).

I do not share Ramsey's and Prior's total scepticism about propositions as a species of abstract object. In rather special circumstances, it is possible to specify a coherent criterion for propositional identity. However, those favourable circumstances do not always obtain (see Rumfitt forthcoming), so I do share Ramsey's and Prior's project of explicating the notions of truth, and the contents of truth-bearing utterances and mental states, without presupposing propositions.

[10] I take this formula *verbatim* from a 1973 Open University Third Levels Arts Course programme, entitled 'Problems of Philosophy: Truth', which consists largely of a discussion between Strawson and Gareth Evans. Still visible on You Tube, it is an evocative record of both men in their prime.

Like Ramsey, Strawson takes truth to apply centrally to beliefs (1971, p.189), but a merit of his and Ramsey's formula is that it may be adapted, with only minimal modification, to explicate truth as it applies to other sorts of bearer. Let us call an utterance or inscription that expresses a complete thought a *statement*, whether or not it has the force of an assertion. Then we can formalize the remark last quoted from Strawson as follows: a statement S is true if and only if $(\exists P)$(S is a statement that $P \wedge P$).

In defining truth as it applies to beliefs, Ramsey takes as understood the notion of a belief's being a belief that P. Similarly, in explicating truth as it applies to statements, Strawson takes as understood the notion of a statement's saying that things are thus-and-so.\[11]/ That is, both these philosophers take the *content* of a belief or statement to be prior to its truth in the order of philosophical explanation. It may seem obvious that they are right to do this: only mental states or utterances with particular kinds of content are so much as candidates to be true or false; whether a candidate mental state or utterance is true then depends on its precise content. I think they are right but, by the time Strawson delivered the famous Inaugural Lecture from which I have quoted, this feature of the approach that he and Ramsey shared was far from being uncontroversial. Indeed, one of Strawson's goals in that Lecture was to reassert the order of priority implicit in its title, 'Meaning and Truth', against those who held that Ramsey had the whole matter back to front. For in his article 'Truth and Meaning', Donald Davidson had complained that intensional notions such as 'believes that', 'says that' or 'means that' are 'obscure' (1967, p.23). According to Davidson, the notion of truth was comparatively clear, so the project of defining the latter in terms of the former had to be wrong-headed. Instead, he embarked on an ambitious programme of explicating the notion of meaning, and thereby other sorts of content, in terms of truth. A Tarskian truth-theory, formulated in an 'extensional first-order language' (Davidson 1976, p.171), and delivering T-theorems in the form 'S is true if and only if P', can 'serve' as a theory of meaning (and eventually cast indirect light on belief). By this, Davidson meant that a Tarskian truth-theory 'explicitly states something knowledge of which would suffice for interpreting utterances of speakers of the language to which it applies' (*ibid.*). That is, it explicitly states something knowledge of which would suffice for knowing what those utterances say.\[12]/

---

[11] I shall use the slightly barbarous locution 'Statement S says that P' as an abbreviation of 'In producing S, what its speaker thereby says is that P'.

[12] I take 'interpret' to have precisely this meaning in Davidson's writings—one free from any connotations of hermeneutic deliberation: 'Kurt utters the words "Es regnet" and under the right conditions we know that he has said that it is raining. Having identified his utterance as intentional and

As is well known, this claim of Davidson's faces serious difficulties, first clearly identified by John Foster. Let us call the T-theorem '*S* is true if and only if *P*' *interpretative* when *S* says that *P*. Even a truth-theory that yields interpretative T-theorems for all the statements in the relevant language on the basis of a correct analysis of their compositional structure will not meet the condition quoted from Davidson: one may know a truth-theory that is in fact interpretative for a language without knowing *that* it is interpretative; yet it is the latter knowledge that one needs in order to know what a statement says (Foster 1976, pp.19-20). I may know that the German noun 'Schnee' designates snow, and that the German adjective 'weiß' is true of an object if and only if it is white, and I may on that basis come to know that the German statement 'Schnee ist weiß' is true if and only if snow is white. But if, with Davidson, we read 'if and only if' as a material bi-conditional, that knowledge is entirely consistent with the statement's saying that snow is white and the earth moves. Even if we switch to a modalized metalanguage and take 'if and only if' to be a strict bi-conditional, the knowledge that 'Schnee ist weiß' is true if and only if snow is white is entirely consistent with its saying that snow is white and two is a prime number. Knowledge of a statement's extensional or modal truth-conditions is always consistent with ignorance of its meaning. So, even in the most favourable case, knowledge of a statement's truth-conditions does not suffice for knowing what it says.\[13]/

What I find most striking here is not the inadequacy of Davidson's answer but the strangeness of his question. It is natural to ask what we know that enables us to understand what other people say; because we can understand sentences we have never heard before, this question forces us to address the issues of linguistic creativity that so exercised Davidson. He, though, does not attempt to answer our natural question. Instead, he asks what knowledge *would suffice* for interpreting a speaker or group of

---

linguistic, we are able to go on to interpret his words: we can say what his words, on that occasion, meant' (Davidson 1973, p.125).

[13] James Higginbotham has 'responded' to Foster by suggesting that 'one will understand Gianni [a typical Italian speaker] when one knows what he, Gianni, knows and is expected to know [by other Italian speakers] about reference and truth. The general principles and certain theorems of a theory of truth for Gianni will figure in one's knowledge about him' (1992, p.9). It is not altogether clear what precise general proposal Higginbotham means to advance about the relationship between understanding and knowledge of truth-conditions, but his response appears to face Foster-type problems. Other Italian speakers will expect Gianni to know that speakers make noises, and Gianni knows that they expect this of him. Supposing that those speakers also expect Gianni to have a basic logical competence, they will expect him to know that 'Firenze è una bella città' is true if and only if Florence is a beautiful city and speakers make noises. All the same, one would misunderstand Gianni if one thought that 'Firenze è una bella città' (as it comes from his mouth) says that Florence is a beautiful city and speakers make noises. Hence, one may know something that Gianni knows and is expected to know about the truth-conditions of 'Firenze è una bella città' without understanding his use of that sentence.

speakers, and his answer to *that* question would be an implausible response to the more natural one: ordinary speakers do not know Tarskian truth-theories. As far as I am aware, Davidson never gave any substantial justification for focusing on a question that he acknowledges to be 'hypothetical'; all he says is that 'it is not altogether obvious that there is anything we actually know which plays an essential role in interpretation' (1973, p.125). One is left with the sense that he was convinced that a Tarskian truth-theory was the answer to something, but struggled to say what.

Be that as it may, the following facts seem altogether obvious to me. (1) Assuming that Kurt is a regular German speaker, I know what he says when he utters 'Entweder es regnet, oder es schneit'; he says that it is either raining or snowing. (2) I have this knowledge—which is actual, not hypothetical—by virtue of knowing what the German expressions (a) 'es regnet', (b) 'es schneit', and (c) 'entweder…oder' mean. Accordingly, (3) there is an interesting problem of explaining how these pieces of knowledge combine to account for my knowing what Kurt's statement says.

In addressing *this* problem, we shall have to work in a language with intensional operators like 'says that' and 'means that'. In answering his question, Davidson was eventually forced to invoke an intensional notion too. No truth-theory, he conceded to Foster, is such that knowledge of it would alone suffice for interpretation (Davidson 1976, p.179). Rather, what would suffice for interpreting the statements of a language *L* is knowledge that 'some <interpretative> T-theory for *L* states that…(and here the dots are to be replaced by a T-theory)' (*op. cit.*, p.174; cf. Foster 1976, p.20). 'States that' is intensional, as is the notion of 'entailing that' which Davidson uses to gloss it (Davidson *op. cit.*, p.178). This puts Davidson in an awkward position. If he is allowed to use the notion of *stating that* in answering his question, then it would seem he can solve the problem of specifying knowledge that would suffice for interpretation very simply, and without any reference to truth-conditions or truth-theories. For what would clearly suffice for interpreting the statement 'Schnee ist weiß' is the knowledge that it states that snow is white. Quite generally, one will be able to interpret a statement if one knows what it states.

In any event, if we allow ourselves intensional notions in characterizing the knowledge that underpins understanding, we can use Prior's sentential quantifiers to formulate semantic axioms that yield theorems which directly specify what statements say. For something any competent speaker of German knows is this: however things may be said or thought to be, if one German statement, *A*, says

that things are one way, and another, *B*, says that things are another way, then their disjunction ⌜Entweder *A* oder *B*⌝ says that things are either one way or the other. In symbols:

(*E*) $\forall P \forall Q$ (if the German statement *A* says that *P* and the German statement *B* says that *Q*, then the German disjunction ⌜Entweder *A* oder *B*⌝ says that either *P* or *Q*).\[14]/

Since I also know that 'es regnet' says that it is raining, and that 'es schneit' says that it is snowing, we may appeal to (*E*) in accounting for my knowledge that 'Entweder es regnet, oder es schneit' says that it is either raining or snowing.

Semantic theories in this unabashedly intensional style may be extended to cover other parts of speech. To characterize the semantic contribution of a noun-phrase like 'Schnee', we can say:

(*S*) The German noun 'Schnee' is used to refer to snow.

Here, 'is used to refer to' means 'is used so as to realize the intention of referring to': an adherent of Frege's theory of *Sinn* and *Bedeutung* will hold that 'Running Rein' may be used to refer to Running Rein without being used to refer to Maccabeus. With this notion in play, we can state the semantic contribution made by 'weiß' as follows:

(*W*) If *NN* is a German noun that is used to refer to α, then the German statement ⌜*NN* ist weiß⌝ says that α is white.

Admittedly, (*W*) is a schema rather than a single axiom; we obtain its instances by replacing the letter α by various names and descriptions. But Davidson's early (1965) insistence on finitely axiomatized semantic theories was surely too restrictive. First-order arithmetic gives a schematic characterization

[14] Compare axiom (M2) of Davies 1981, p.42. Davies, though, reads '$\forall P$' and '$\forall Q$' as substitutional quantifiers (p.44) and this leads to a problem. If the result of substituting sentences for '*P*' and '*Q*' in 'says that either *P* or *Q*' is not to be macaronic gibberish, the substituted sentences must be in English, but there is no reason to assume that the content of any German statement has to be expressible in English. Prior's interpretation of the quantifiers avoids this problem. Although I have formulated (*E*) as an English sentence, the truth it expresses does not (on that interpretation) relate particularly to English sentences. Rather, it concerns the ways things might be said to be in any language.

of our mastery of mathematical induction, and there is no evident reason to baulk at a schematic characterization of our mastery of 'weiß'. At any rate, an instance of (*W*) combines with (*S*) to yield the directly meaning-specifying theorem: the German statement 'Schnee ist weiß' says that snow is white. There is no strain in attributing implicit knowledge of axioms and schemata like (*E*), (*S*), and (*W*) to ordinary German speakers. Indeed, one might take these axioms or schemata to spell out the meaning of such locutions as that 'entweder…oder' *means* either…or, 'Schnee' *means* snow, 'weiß' *means* white—locutions that we use all the time, even though philosophers of language have paid scant attention to them. Of course we have here only the beginnings of an intensional semantic theory for German; it would be a major project to see if such a theory could be constructed for that and other natural languages. But we have made a start on the problem of explaining how knowledge of word meanings combines to yield knowledge of what complete statements say—a start that does not proceed via the construction of a truth-theory in Davidson's preferred 'extensional first-order language'.

According to Wittgenstein's *Tractatus*, 'to understand a statement means to know what is the case if it is true…It is understood by anyone who understands its constituents' (Wittgenstein 1922, 4.024). Our analysis is entirely consonant with Wittgenstein's thesis. Everyone ought to agree that understanding a statement is a matter of knowing what it says on the strength of knowing what its constituents mean. Given Ramsey's definition of truth, though, where a statement *S* says that *P*, and only that *P*, *S* is true if and only if *P*, so the condition that has to be satisfied for *S* to be true is precisely that *P*. On this analysis, Davidson's mistake was not to take truth-conditions to be central in an account of meaning and understanding. It was, rather, to try to give a purely extensional account of truth-conditions. The basic condition that must be satisfied for a statement to be true is that what it says should be the case. If 'Running Rein is four years old' and 'Maccabeus is four years old' say different things, they have different truth-conditions, even though it is metaphysically necessary that one truth-condition obtains just when the other does.

Does this mean there is nothing of value in Davidson's philosophy of language? Our analysis exposes severe limitations in his account of the epistemology of understanding: there, his predilection for extensional notions led him to duck the central question in the field. But some of his reflections bear more fruitfully on what one might call the metaphysics of meaning. The fact that 'Schnee ist weiß' says that snow is white is surely not *brute*: a philosopher may reasonably seek to identify the more basic truths in virtue of which it and similar facts obtain. Davidson holds that these more basic

truths concern the way an optimally performing 'radical interpreter' would interpret the statement. More precisely, his answer to the metaphysical question—'In virtue of what does *S* say that *P*?'—begins: *S* says that *P* in virtue of the fact that a semantic theory whose canonical deliverances make optimal overall sense, given the principles of radical interpretation, of the speakers of the language of *S* will canonically yield a theorem saying that *S* is true if and only if *P*. This answer acquires substance as the principles of radical interpretation are spelled out, but since those principles crucially concern the relation between meaning and belief, they may be re-formulated to apply to a semantic theory whose theorems explicitly tell us what statements say. In the last of his Dewey Lectures, Davidson acknowledged this:

> The thesis that a theory of truth-conditions [i.e., a truth-theory cast in a purely extensional language] gives an adequate account of what is needed for understanding the literal meanings of utterances is, of course, much disputed…If the [thesis] is mistaken, much of the detail in what I shall go on to say about the application of the concept of truth will be threatened, but the general approach will, I think, remain valid' (Davidson 1990, p.54).

By 'the general approach' he means his theory of interpretation. I shall not try to assess his 'interpretivist' answer to the metaphysical question today, nor compare it with the more reductive answer that pragmatist writers have developed from Ramsey's own suggestions.\[15]/ All the same, it is

[15] See, most recently, Mellor 2012.

Mellor's paper provides the most detailed and persuasive elaboration to date of Ramsey's rather sketchy account of what determines the contents of beliefs and statements. All the same, the 'exclusionary' semantic theory developed in my 2011 may usefully contribute something that Mellor does not provide—namely, an account of how the contents of complex statements and beliefs relate to those of their components. Mellor follows Braithwaite in taking a belief that *P* to be a disposition to act as if *P*. But how, on this account, does a belief that either *P* or *Q*, say, relate to a belief that *P* and a belief that *Q*? Anyone disposed to act as if *P* will be disposed to act as if either *P* or *Q*, and similarly for *Q*. But the converse does not hold. If I believe that either *P* or *Q* but have no view as to which disjunct true, I shall be disposed to act as if either *P* or *Q* even though I am neither disposed to act as if *P* nor disposed to act as if *Q*.

Exclusionary semantics can help here. For the exclusionist, to have a certain disposition to act is to set aside, or discount, certain possibilities in acting or in making plans for action. Let f_P be the set of possibilities that are discounted in being disposed to act as if *P*. Then we can characterize the relation between believing that either *P* or *Q*, believing that *P*, and believing that *Q* as follows: $f_{P \text{ or } Q} = f_P \cap f_Q$. That is: in being disposed to act as if either *P* or *Q*, one sets aside precisely those possibilities that one sets aside both in being disposed to act as if *P*, and in being disposed to act as if *Q*. (See Rumfitt 2011, pp.235-41, for further development of this point, and for corresponding semantic axioms for conjunction, negation and the conditional.)

noteworthy that Davidson's answer to the metaphysical question may be accepted by those who reject the thesis that knowledge of a truth-theory would suffice for understanding a language.

III. The definability of truth

Ramsey offers a definition of truth, but Tarski is said to have proved that truth is indefinable. Did Tarski refute Ramsey?

Tarski's indefinability theorem applies to a formalized language, L, containing a negation operator $\neg$ and a device $\ulcorner\ \urcorner$ which, when applied to any well-formed expression of L, yields a singular term designating that expression. It is assumed that the syntax of the language is strong enough to prove the diagonal lemma: for any formula $A(x)$ in L, with x free, there is a formula B in L such that B is equivalent to $A(\ulcorner B \urcorner)$. According to Tarski, a *truth-predicate* for L is a one-place predicate $Tr(\xi)$ such that, for any closed formula A of L, $Tr(\ulcorner A \urcorner)$ is equivalent to A. What Tarski proved is that no truth-predicate *for* L can be a predicate *in* L. For suppose $Tr(\xi)$ were a truth-predicate for L in L. Then $\neg Tr(x)$ would be a well-formed formula of L with x the only free variable. By the diagonal lemma, there would exist a closed formula D in L such that D is equivalent to $\neg Tr(\ulcorner D \urcorner)$. Since $Tr(\xi)$ is a truth-predicate for L, we would also have that $Tr(\ulcorner D \urcorner)$ is equivalent to D, so that $Tr(\ulcorner D \urcorner)$ and $\neg Tr(\ulcorner D \urcorner)$ would be equivalent. This contradiction reduces to absurdity the supposition that $Tr(\xi)$ is a truth-predicate in L. So no truth-predicate for L can be in L (Tarski 1935, pp.249-51).

Tarski's theorem appears to contradict Ramsey's definition of truth. We have not specified the precise properties of the language in which that definition is cast, but Ramsey's definition of falsity uses a negation sign. Moreover, whereas Ramsey differs from Tarski in taking the primary bearers of truth to be beliefs, one could adapt Ramsey's proposal so as to define a one-place truth-predicate that applies to the type sentences of a formalized language which lacks any context-dependent expressions: $Tr(S)$ if and only if $(\exists P)(S$ expresses the thought that $P \wedge P)$.

Given Tarski's assumptions, his proof of the indefinability theorem is unassailable, but Ramsey's account of truth provides the resources to challenge one of those assumptions. Tarski assumes that if $Tr(\xi)$ is a truth-predicate for L, then any closed formula A of L is equivalent to $Tr(\ulcorner A \urcorner)$. He thereby presupposes that any closed formula of L has truth-conditions. On Ramsey's view, a

formula has truth-conditions only if it expresses a thought, so Tarski presupposes that any closed formula of the relevant language expresses a thought. More colloquially, he presupposes that any closed formula says something. The very formula Tarski constructs in proving his theorem, however, casts doubt on that presupposition.

To see this, it helps to work with a truth-predicate that applies to the type sentences of English. Since English contains context-sensitive expressions, such a predicate has to be relativized to a context of potential utterance or inscription: it makes no sense to say that the type sentence 'I am tired' is true or false *simpliciter*, but it does make sense to say that the sentence is true, as potentially uttered by me at midnight. Let us, then, introduce a two-place predicate '$Tr\ (A, c)$', understood to mean 'The declarative type sentence A is true as potentially uttered or inscribed in the context c', and a corresponding predicate 'Say (A, c, P)' meaning 'Type sentence A, as potentially uttered or inscribed in the context c, says that P'. We may then define '$Tr\ (A, c)$' in Ramsey's style as follows:

$$(T) \qquad \forall A \forall c(\text{True}\ (A, c) \leftrightarrow \exists P\ (\text{Say}\ (A, c, P) \wedge P)).$$

We may similarly define the corresponding notion of falsity:

$$(F) \qquad \forall A \forall c(\text{False}\ (A, c) \leftrightarrow \exists P\ (\text{Say}\ (A, c, P) \wedge \neg P)).$$

Here, '$\forall A$' and '$\forall c$' are ordinary objectual quantifiers, ranging over type sentences and contexts respectively, while '$\exists P$' is our existential sentential quantifier. (T) and (F) are single formulae, not schemata, and they regulate the application of truth and falsehood to type sentences in all languages; the language to which a sentence belongs may be taken to be determined by the context of utterance.

Tarski uses the diagonal lemma to construct a sentence which 'says of itself' that it is not true. In a similar spirit, let us consider the English sentence

The sentence printed on the last line of page 16 of this essay is not true.

Let '*L*' be a name of the sentence just displayed, and let *c** be its actual context of inscription. We can be sure that if *L* says anything in *c**, it says that the sentence printed on the last line of page 16 is not true in *c**. Given what its component words mean in the actual context of inscription, *L* must say *that*, if it says anything at all. On the assumption that *L* says something in *c**, then, (*T*) yields

(1) True (*L*, *c**) ↔ the sentence printed on the last line of page 16 is not true in *c**.

Now a simple empirical investigation establishes

(2) *L* = the sentence printed on the last line of page 16

and, by Leibniz's Law, (1) and (2) together yield

(3) True (*L*, *c**) ↔ ¬True (*L*, *c**).

As in Tarski's proof of his indefinability theorem, (3) is a contradiction, but we need not take the contradiction as reducing to absurdity the assumption that the language contains the predicate 'True (*A*, *c*)'. Instead, we may take the contradiction to refute line (1). Tarski writes that we 'wish to use the term "true" in such a way that all [T-equivalences in the form "*S* is true if and only if *P*"] can be asserted, and we shall call a definition of truth "adequate" if [and only if] all these equivalences follow from it' (Tarski 1944, p.344). *Mutatis mutandis*, he assumes that any satisfactory account of our two-place truth-predicate 'True (*A*, *c*)' will entail (1). What Ramsey's analysis reveals, though, is that our commitment to assert such T-equivalences is provisional. We shall assert '*S* is true if and only if *P*' only when we believe that *S* expresses a thought, and we shall assert (1) only if we accept the additional assumption that *L* says something when inscribed in *c**. It is, then, open to us to take (3), not to reveal any problem with our Ramseyan definition (*T*), but instead to refute this additional assumption. In other words, the deduction shows that *L* does not say anything when inscribed in *c**. It proves ¬∃*P* Say (*L*, *c**, *P*).

The theory of truth and falsity that comprises (*T*), (*F*) and their logical consequences is, indeed, formally consistent. As Timothy Williamson has observed, we can show this 'by constructing

an unintended model…in which formulas are treated as referring to truth-values, the propositional quantifiers range over truth-values, and all formulas of the forms "Say (A, c, P)", "True (A, c)", and "False (A, c)" are treated as false' (Williamson 1998, p.14). In the context of such a theory, 'the semantic paradoxes are transformed into sound arguments for constraints on what can say what in what contexts' (p.19). Transformations of this kind are familiar. In the context of Zermelo-Fraenkel set theory, Russell's paradox is transformed into a sound argument for the conclusion that there is no set of non-self-membered sets. We shall have to consider whether the hypothesis that the inscription of L says nothing can be sustained: after all, it *seems* to say that the sentence printed on the last line of page 16 of this essay is not true. But the hypothesis is clearly a formal possibility. Tarski has not refuted Ramsey's definition of truth.

This reply to Tarski requires that we allow that a well-formed declarative sentence might fail to say anything in a particular context, or perhaps in any context. Philosophers take similar possibilities seriously in other areas. In circumstances where you have dyed an elephant pink, thereby enraging it, your utterance of 'That pink elephant is about to charge' says something. Yet the departmental dipsomaniac says nothing when he utters the very same words, with the same meaning, while pointing into thin air (see Evans 1982 and McDowell 1982). The explanation of why no thought is expressed will be very different in the two cases. All the same, there is nothing outlandish in the contention that, in uttering a well-formed sentence one might, in unfavourable circumstances, fail to say anything whatever.

The argument that L says nothing in c^* tacitly presupposes that any thoughts a single unambiguous sentence might express in a single context are materially equivalent. Some philosophers take Liar sentences to be counterexamples to this presupposition. Stephen Read (2009) has revived Thomas Bradwardine's idea that a sentence signifies its own truth, as well as what it explicitly says. According to Bradwardine, a sentence (as uttered in a given context) is true only if things are wholly as it signifies in that context: it is not enough that things should be as it explicitly says them to be. On this view, (T) will have to be replaced by

$$(T^*) \qquad \forall A \forall c(\text{True}\ (A, c) \leftrightarrow (\exists P\ \text{Signifies}\ (A, c, P) \wedge \forall Q(\text{Signifies}\ (A, c, Q) \rightarrow Q))).$$

and (F) by

$$(F^*) \qquad \forall A \forall c(\text{False}\ (A, c) \leftrightarrow (\exists P\ \text{Signifies}\ (A, c, P) \wedge \forall Q(\text{Signifies}\ (A, c, Q) \rightarrow \neg Q))).$$

When applied to our paradoxical inscription, the usual reasoning shows that *L* is not true in *c**, but on Bradwardine's account we cannot conclude from this that *L* is true in c* after all. For while *L* explicitly says that *L* is not true—which is the case—it also signifies that *L* is true, which is not the case, so things are not *wholly* as *L* signifies. Within Bradwardine's theory, then, the Liar reasoning is transformed into a sound argument for the conclusion that *L* expresses no *unique* thought in *c**, a conclusion that may be easier to swallow than the claim that it says nothing. Although they deviate from his precise proposal, (*T**) and (*F**) remain entirely in the spirit of Ramsey's account of truth.

IV. Intensional paradox

Ramsey's definitions of truth and falsity, then, have much to recommend them. They are formally consistent, and Geach's objection poses no threat to them. Moreover, they connect with a promising approach to constructing semantic theories for natural languages. Why, one wonders, are they not more widely accepted?

The main reason is a lurking threat of paradox. Williamson demonstrated the formal consistency of (*T*) and (*F*) by considering an unintended interpretation in which the sentential quantifiers range over truth-values, and in which no sentence says anything in any context. But a theory of truth and falsity which works only when nothing is said is clearly of rather limited interest, so we have to check that the Ramseyan theory of truth and falsity remains consistent when we add postulates stating that this utterance, or that inscription, says that such-and-such is the case.

We can *begin* to do this without running into difficulties (although see note 16 below). The theory remains consistent, for example, if we add (*S*), (*W*) and (*E*)—i.e., the intensional semantic postulates for German proposed in §II. Indeed, we can consistently add similar disquotational principles for all sentences not involving propositional attitudes or semantic notions, thereby yielding such theorems as

(*SW*) $\forall c$(*c* is a context in which English is spoken $\rightarrow$ Say ('Snow is white', *c*, snow is white)).

Eventually, though, we shall encounter trouble. In fact, the basic problem does not pertain to (*T*) and (*F*) or to (*T**) and (*F**) *per se*, but to the apparatus of sentential quantification combined with intensional operators that all these definitions use.

The underlying problem manifests itself in a number of ways. The manifestation I shall discuss brings out the central difficulty, but I have no proof that my remedy cures all forms of the disease. If Ramsey is to be fully vindicated against the problem that I now address, more work will be needed.

Prior (1958, 1961) half saw the problem. He considers the case of Epimenides the Cretan, who utters the sentence 'Nothing said by a Cretan is the case'. That utterance cannot be true, of course, from which it appears to follow that something said by a Cretan is the case. Suppose, though, that nothing *else* said by a Cretan is the case. In those circumstances, the supposition that Epimenides succeeds in saying that nothing said by a Cretan is the case yields a contradiction; so it would appear to be 'impossible for Epimenides the Cretan to say that nothing said by a Cretan is the case; whatever noises he makes, he will not under those circumstances be able to say *that* by them' (Prior 1961, p.19). As Prior observes, this result is a simple corollary 'of the obvious truth that *if it is a fact that no fact is asserted by a Cretan, then THIS fact (that no fact is asserted by a Cretan) is not asserted by a Cretan either*' (*ibid.*). There are corresponding truths for other intensional verbs. Thus, if it is a fact that no Cretan believes/grasps/fears-the-obtaining-of/uses-words-which-standardly-express any fact, then no Cretan believes/grasps/fears-the-obtaining-of/uses-words-which-standardly-express THIS fact (namely, that no Cretan believes/grasps/fears-the-obtaining-of/uses-words-which-standardly-express any fact) either.\[16]/

[16] Prior held that similar cases scupper disquotational principles specifying what statements say, even when the sentences used in making those statements contain no semantic vocabulary or propositional attitude verbs. One is apt to believe, with (*SW*), that any English speaker who utters the sentence 'Snow is white' thereby says that snow is white. But Prior contends that this belief is false (1961, p.29). For suppose Tarski utters 'Snow is white' immediately after Prior has said 'If, immediately after this, Tarski says something that is the case, then what I am now saying is not the case'. Since snow is white, we reach a contradiction if we suppose (1) that Tarski says that snow is white, and (2) that Prior says that if Tarski immediately says something that is the case, then what Prior says is not the case. According to Prior, the only possible basis for holding that one of these utterances, rather than the other, succeeds in expressing a content is 'who got his say in first' (*op.cit.*, p.21). Since Prior spoke first, he made it impossible for Tarski to say anything that is the case, so when Tarski uttered 'Snow is white' he could not thereby say that snow is white. The most we can say about that sentence is that an English speaker who utters it says that snow is white, *provided* he says something by it. If it works at

Prior already had misgivings about the conclusion that Epimenides cannot assert, or believe, or otherwise relate to the thought that his utterance would normally express. He was, he confessed, 'driven very unwillingly' to conclude that 'thinking, fearing, etc., because they are attitudes in which we put ourselves in relation to the real world, must from time to time be oddly blocked by factors in that world, and we must just let Logic teach us where these blockages will be encountered' (*op. cit.*, p.32). But even if we can bring ourselves to accept that conclusion, 'revenge' paradoxes lie just around the corner, as J.L. Mackie (1973) pointed out. On Prior's view, even when nothing else said by a Cretan is the case, 'the thing itself—that nothing said by a Cretan is the case—will under those circumstances be true…That is, there [is] nothing wrong with *what we suppose the Cretan to say*, but only with *the supposition that* he says it' (*op. cit.*, p.19). In other words, Epimenides's words express a proposition that he is blocked from saying. Where 'say' means 'assert', this may well be correct: perhaps Epimenides is blocked from asserting a truth that others can assert. But now suppose that Epimenides utters the sentence 'No sentence uttered by a Cretan, standardly construed, makes a true statement', and assume as before that no other sentence uttered by a Cretan, standardly construed, makes a true statement. As Mackie remarks, 'we cannot without contradicting ourselves allow that Epimenides's remark makes a true statement. And yet if it fails *for whatever reason* to make a true statement, we must ourselves say exactly what Epimenides has said; how then can we deny that this is a sentence uttered by a Cretan which, standardly construed, makes a true statement? How can we avoid contradicting ourselves?'(Mackie 1973, p.294).

Mackie's answer is as follows:

> Suppose that we expand 'true' here, replacing 'would make a true statement' with 'would state that things are as they in fact are'. And remember that the things in question include the

---

all, this argument could be applied to refute *any* principle to the effect that an utterance of a sentence *S* always says that *P*.

I do not think that the argument does work, though: it accords quite unreasonable importance to temporal priority. A more sensible view takes account of whether the relevant sentences involve semantic predicates or propositional attitude verbs. As per (*SW*), when Tarski utters the simple sentence 'Snow is white' he succeeds in saying that snow is white, and the fact that he expresses a truth shows it was Prior who failed to say anything by his utterance. On this view, whether Prior's utterance has a content depends on whether Tarski subsequently utters a truth or a falsehood, so no judgement on the contentfulness of Prior's utterance is possible before Tarski speaks. However, I do not think that this consequence is implausible. Prior's words advert to Tarski's forthcoming utterance and, given what they mean, we should accept that whether they express a thought depends on the truth-value of that later utterance.

> success or failure of this sentence itself in this respect. I think we can and must say that because of the very tricky kind of self-reference and consequent self-dependence in this case, there just is no *how things are* in the key respect. Consequently, we cannot either endorse or deny a sentence-token of the same type and with the same reference as [Epimenides's remark]. We cannot now drive a wedge between what we say about that sentence and what we allow it to say about itself: standardly construed, it would say just what our own use of the same type sentence would. We must just admit that the issue it appears to raise is indeterminate, and hope our study of self-reference has explained why this is so. This sentence's indeterminacy with respect to truth is of a kind which prevents our saying even that it is not true, and therefore from arguing, by a further step, that it is true (*op. cit.*, p.295).

The main problem with this answer, though, is that we appear to have a watertight logical deduction whose conclusion is the sentence that Mackie deems to be indeterminate. So far from teaching us where blockages occur, Logic leads us from true premisses to a conclusion that fails even to express a thought.\[17]/

It is remarkable how weak a logic is needed to generate trouble. To see this, it helps to set out the problem formally, and in doing so I shall switch to a slightly different example that makes utterances by other Cretans irrelevant. Let us suppose, then, that Epimenides utters 'Epimenides utters no sentence or closed formula which, when standardly construed, expresses something that is the case'. More exactly, let us suppose that he utters the formula $\ulcorner \neg\exists P(\delta P \wedge P)\urcorner$, which I shall label '$\lambda$', in an interpreted formal language in which '$\exists P$' is Prior's existential sentential quantifier and 'δP' means 'Epimenides utters a sentence or formula which, when standardly construed, expresses the thought that P'. Let us also suppose that λ is the only sentence or formula that Epimenides utters in the relevant period. We cannot assume that Epimenides's words succeed in expressing a thought, but we can be sure that any thought his words express will, when standardly construed, be equivalent to the thought

[17] Even if what I have called 'the main problem' can be solved, Mackie's explanation of why paradoxical utterances and inscriptions fail to say anything is at best incomplete. He hopes that his 'study of self-reference' will provide the needed explanation, but there are paradoxes akin to our version of the Epimenides that involve no self-reference, such as the following intensional variant of Yablo's paradox: each person in the infinite queue of guests waiting to check in to Hilbert's Hotel says (or thinks) 'No one behind me in the queue is saying (or thinking) anything that is the case' (cf. Yablo 1993). For an explanation of why no one in the queue says (or thinks) anything, which coheres both with Mackie's account of how self-reference can block a sentence from saying anything and with the solution I am about to advance to the 'main problem', see Goldstein 2006.

that Epimenides utters no sentence or formula which, when standardly construed, expresses something that is the case. In other words, we have the following:

$$\forall Q(\delta Q \rightarrow (Q \leftrightarrow \neg\exists P(\delta P \wedge P)).$$

We can now set out an argument for the conclusion that Epimenides utters no sentence or closed formula which, when standardly construed, expresses anything. That conclusion may be formalized as $\ulcorner \neg\exists R\delta R \urcorner$, and we deduce it by deriving a contradiction from the contrary hypothesis $\ulcorner \exists R\delta R \urcorner$:

1.	$\forall Q(\delta Q \rightarrow (Q \leftrightarrow \neg\exists P(\delta P \wedge P))$	Premiss (see above)
2.	$\exists R\delta R$	Assumption, for *reductio*
3.	δR	2, existential instantiation
4.	$\delta R \rightarrow (R \leftrightarrow \neg\exists P(\delta P \wedge P))$	1, universal instantiation
5.	$R \leftrightarrow \neg\exists P(\delta P \wedge P)$	3, 4 modus ponens
6.	R	Assumption
7.	$\neg\exists P(\delta P \wedge P)$	5, 6 $\wedge$-elimination and modus ponens
8.	$\delta R \wedge R$	3, 6 $\wedge$-introduction
9.	$\exists P(\delta P \wedge P)$	8 existential generalization
10.	$\neg R$	7, 9 *reductio*, discharging assumption 6
11.	$\neg\neg\exists P(\delta P \wedge P)$	5, 10 $\wedge$-elimination and modus tollens
12.	$\exists P(\delta P \wedge P)$	Assumption
13.	$\delta S \wedge S$	12, existential instantiation
14.	δS	13, $\wedge$-elimination
15.	$\delta S \rightarrow (S \leftrightarrow \neg\exists P(\delta P \wedge P))$	1, universal instantiation
16.	$S \leftrightarrow \neg\exists P(\delta P \wedge P)$	14, 15 modus ponens
17.	S	13, $\wedge$-elimination
18.	$\neg\exists P(\delta P \wedge P)$	16, 17, $\wedge$-elimination and modus ponens
19.	$\neg\exists P(\delta P \wedge P)$	12, 18 *reductio*, discharging assumption 12

20.	$\neg\exists R\delta R$	11, 19 *reductio*, discharging assumption 2

So far, our derivation vindicates Mackie's claim that, in the unfavourable circumstances described, Epimenides's words (as standardly construed) express no thought. The problem, though, is that the logical rules used in reaching this conclusion may be further applied to show that things are as Epimenides's words (as standardly construed) say them to be, thereby ensnaring us in a revenge paradox. For, as it seems, our derivation may be validly extended as follows:

21.	$\exists P(\delta P \wedge P)$	Assumption
22.	$\delta P \wedge P$	21, existential instantiation
23.	δP	22, $\wedge$-elimination
24.	$\exists R\delta R$	23, existential generalization
25.	$\neg\exists P(\delta P \wedge P)$	20, 24, *reductio*, discharging assumption 21

Line 25 is just the formula λ. So, it seems, the argument that shows that Epimenides's words say nothing may be extend to establish the truth of what they say.

Commenting on Prior's use of 'Logic' to identify the circumstances in which our taking up certain mental attitudes is 'oddly blocked', Mackie remarks that Prior's derivations 'appeal, naturally enough, to the laws of non-contradiction and excluded middle *as applied to such formulae as would, when interpreted, become, e.g., "No Cretan statements are true"*. It is being implicitly assumed in the formal procedures that each such item must itself be either true or not true, and not both' (1973, p.279). Perhaps *reductio* presupposes some form of the law of non-contradiction, but the derivation just set out shows that any appeal to excluded middle is inessential: an intuitionist logician can accept all the rules that are applied in moving from line 1 to line 25.

Our derivation also shows the source of the difficulty lies in the apparatus on which Ramsey's definitions of truth and falsity rely—namely, sentential quantification into the scope of an intensional operator—rather than in the definitions themselves. Tweaking (*T*) or (*F*) will not help, for neither is used in deriving (20) or (25). In particular, the Bradwardinian idea sketched at the end of §III is of no avail here. On Bradwardine's view, λ signifies at least two different thoughts, so premiss (1) will be false if 'δP' is read as 'Epimenides utters a sentence or formula which signifies the thought that *P*'.

However, his theory presupposes a sharp distinction between the thought that a declarative utterance explicitly expresses *au pied de la lettre* (its primary signification, in his terminology) and those thoughts that it signifies in other ways—for example, by implicature. On Bradwardine's view, then, premiss (1) remains compelling if 'δP' is read as 'Epimenides utters a sentence or formula which explicitly expresses (i.e., primarily signifies) the thought that P'. Moreover, the derivation from lines 1 to 25 goes through under this interpretation of 'δ'. So Bradwardine's theory, while it may dispose of some versions of the Liar, is of no help with this one.

While demonstrably consistent in the hygienic environment that Williamson proposes, then, Prior's quantifiers lead to trouble when combined with a very weak assumption—viz., premiss (1)—involving an intensional notion. It is perhaps unsurprising that paradoxes should ensue from this combination: sentences involving quantifiers that range over (or otherwise relate to) those sentences themselves exhibit an impredicativity that often turns pathological. But Ramsey's definition of truth involves just such a combination, and Geach's warning seems apposite. Until the present 'nasty problem of intensionality' has been solved, we can hardly blame philosophers for shunning an account of truth built on such marshy ground.

V. A way out

Is there a solution that rehabilitates Ramsey's theory? Constraints of space mean that I can only gesture at problems in the approaches I find unpromising before sketching the solution I prefer.

One way out would be to ramify the sentential quantifiers: we cannot quantify over all the ways things might be said or thought to be; rather, a sentential quantifier at level n quantifies over the ways things might be said or thought to be at level n–1. An interesting paper by Dustin Tucker (2013) shows that this approach is less expressively restrictive than one might at first suppose, but it still faces the problem that bedevils any ramified theory: in saying 'we cannot quantify over all the ways things might be said or thought to be' we seem to be doing precisely what the ramified theory says cannot be done.

A more popular approach focuses on changes in context. Line 20 of our derivation says that Epimenides's utterance of λ says nothing in the context in which he uttered it. At line 25, we derive an

instance of the same formula; so λ had better be true, and hence say something, in the context of line 25. There is no immediate inconsistency here. The contexts of Epimenides's utterance and of my inscription are quite different, and a single sentence or formula might express a truth in one of those contexts while saying nothing in the other. More precisely, the inscription of a sentence in one context might truly say that an utterance of the same sentence in a different context says nothing.\[18]/

The last sentence describes a logical possibility, and contextual factors are crucial to some versions of the Liar.\[19]/ All the same, the contextualist approach does not get to the heart of the present paradox. The approach requires that the range of ways things might be said or thought to be expands between Epimenides's utterance of λ and my inscription of it, but no one has convincingly explained what this expansion consists in and why it has to occur.\[20]/ What is worse, the paradox-monger may take steps to forestall the expansion. Suppose Epimenides utters $\ulcorner\neg\exists P(\delta P \wedge P)\urcorner$, where '$\delta P$' now means 'Epimenides utters a sentence or formula which Rumfitt will interpret as expressing the thought that P'. As before, the derivation yields $\ulcorner\neg\exists R\delta R\urcorner$, so $\ulcorner\neg\exists P(\delta P \wedge P)\urcorner$, as I interpret it, says nothing. At line 25 of the derivation, though, I affirm the very same formula, $\ulcorner\neg\exists P(\delta P \wedge P)\urcorner$. That is, I am led in the course of the derivation to affirm a formula that, as I interpret it, says nothing.

One philosopher has embraced Mackie's conclusion—that the issue λ appears to raise is not genuine—while refusing to be worried by the fact that we seem to have a logical proof of λ. Timothy Smiley speaks of grammatically acceptable sentences *malfunctioning*: 'in a particular context, perhaps in any context, they fail to convey any coherent message' (Smiley 1993, p.23); and he invites us to

> consider *A*: '*A* is not true', where this is spelt out as '*A* is false or malfunctions'. Concluding that *A* malfunctions, how can we avoid the further conclusion that *A* is not true, with the consequent re-entry into paradox? Answer: this conclusion depends on the inference '*A* malfunctions. Therefore *A* is false or malfunctions'. But it's not that *A* is false or malfunctions; it's that '*A* is false or malfunctions' malfunctions. The fact that the conclusion of the inference fails to express the proposition which its form would suggest, undercuts the

---

[18] For this approach to the Liar, see Parsons 1974, and Glanzberg 2001 and 2004.

[19] Thus, in the situation of note 17, someone outside the queue can truly say (or think) of any person in the queue, 'No one behind him is saying (or thinking) anything that is the case'.

[20] Gauker (2006) persuasively criticizes the explanation advanced in Glanzberg 2004.

> appeal to form on which the inference relies for its validity. Sod's law trumps the law of or-introduction, just as it does when sentences are ambiguous or context-dependent…As Priest says of his own solution [to the Liar], 'If this is all disconcertingly non-algorithmic, this is just an unfortunate fact of life' (*op. cit.*, p.26).

On Smiley's view, the laws of logic are those of the classical system, but we have to recognize that the correct application of those laws to a true premiss will sometimes result, not in a true conclusion, but in a sentence or formula that does not say anything at all. There are no formal rules by following which we can be sure to avoid producing contentless utterance or inscriptions.

In a memorable clash at the 1993 Joint Session, Graham Priest protested that Smiley here 'waves goodbye to the project of formal logic, that is, of determining a (non-empty) class of inferences that are guaranteed to be truth-preserving in virtue of their form' (Priest 1993, p.43). Priest's characterization of the logical project is tendentious. On the present view, Liar-like paradoxes present no threat to laws of logic when these are conceived (as in §I) as laws of thought. Those laws lay down generalizations about ways things may be said or thought to be. On the present account, Epimenides's utterance of λ fails to say that things are any way, so its evaluation is irrelevant to the laws of thought.

The evaluation of λ does, though, bear on a related enterprise—that of laying down rules for constructing logical derivations in a given language, rules which guarantee that true premisses yield true conclusions. One moral of our discussion is that such rules cannot be purely formal or syntactic. Whether an utterance is paradoxical often depends on its context: Epimenides's original remark is paradoxical if no other utterance by a Cretan expresses a truth; otherwise, it is straightforwardly false.\[21]/ Even in a formalized language, we cannot expect to find a syntactic filter that excludes paradoxical inscriptions and thereby ensures that whenever we inscribe a well-formed formula we shall say something.

Even so, it would be good to have rules which at least warn us of what pitfalls to look out for as we take an inferential step. The quest for such rules may be thought quixotic. They try to spell out what *would* follow from a given utterance or inscription if it said anything—while holding open the possibility that those very consequences might reveal that it says nothing after all. The quest certainly involves skating on thin ice, but even in the absence of a syntactically specified safe route across the

[21] For another discussion of the Liar that stresses the role of context, see Kripke 1975.

frozen lake, it is worth trying to steer clear of particular danger zones. A logician who merely shrugs his shoulders and mutters 'Sod's Law' while thinkers disappear into the icy depths of nonsense is not really earning his keep.

In order to see what these rules might be, it helps to reflect on the approach to the Liar that is being recommended. Central to that approach is the distinction between a sentence or well-formed formula's expressing a falsehood and its failing to express a truth. In both classical and intuitionistic logic, we signify that a formula expresses a falsehood by asserting its negation. But how can we signify that it fails to express a truth?

One approach postulates two negation operators (see Bochvar 1939). We have the familiar 'internal' negation operator '¬', understood so that ⌜$\neg A$⌝ is true when *A* is false, is false when *A* is true, and fails to say anything when *A* so fails. Alongside this, some logicians have postulated an external negation operator '~'. As with '¬', ⌜$\sim A$⌝ is taken to be true when *A* is false and false when *A* is true. The external negation ⌜$\sim A$⌝, though, is counted as true when *A* says nothing. It is far from evident, however, that this attempt to endow '~' with a sense is coherent. If ⌜$\sim A$⌝ is true, it must surely say something, but it is quite unclear how it can do so when its component part, *A*, says nothing.\[22]/

For this reason, I recommend a different approach to the problem of formalizing the difference between a sentence's expressing a falsehood and its failing to express a truth. Let us use the notation +*A* to signify that *A* is *accepted as true*. Then we can express our belief that *A* is false by writing down +(¬*A*). Cognate to the mental act of accepting a sentence as true is that of *rejecting it as untrue*. We signify our rejection of *A* as untrue by writing down ─*A*. The difference between +(¬*A*) and ─*A* symbolizes the difference between rejecting *A* as false and rejecting it as untrue. Unlike those who postulate an external negation operator, I do not suppose that ─*A* expresses any propositional content when *A* says nothing. By writing down ─*A*, one simply expresses one's rejection of sentence *A* as untrue—an act one will wish to perform when *A* fails to say anything. As explained, the '─' symbol

[22] Smiley proposes a variant of this approach. Following Ducrot (e.g. 1984) and Horn (1989), he discerns a 'polemical' negation in natural language: 'Polemical negation signifies objection to an (actual or possible) utterance as inappropriate, whether because misleading, an understatement, untrue, unwarranted, meaningless, misspelt, not p*l*t*c*lly c*rr*ct, or for any other reason' (Smiley 1993, pp.20-21). As Priest observes, however, the trouble with this solution is that 'not_2 [the postulated polemical negation operator] will not do the job that is required of it. One may, in uttering not_2-α, be doing no more than rejecting certain connotations of conversational implicatures of α. This is quite compatible with the sentence negated expressing a truth' (Priest 1993, p.44).

may not be iterated: $-(\neg A)$ expresses one's rejection of (the well-formed sentence) $\ulcorner \neg A \urcorner$ as untrue. But $-A$ is not a sentence, so no meaning has been attached to $--A$.

What we shall then need is a system of rules that regulates inferential transitions between these *signed formulae*.\[23]/ We want the rules to preserve *correctness*: a positively signed formula $+A$ is correct if and only if A is true, while a negatively signed formula $-A$ is correct if and only if A is untrue. We take over the rules of the positive logic in the obvious way; thus *modus ponens* takes the form: From $+(A \rightarrow B)$ and $+A$, infer $+B$. But we also need rules that mix positively and negatively signed formulae. *Contra* the dialetheists, I suppose that deeming A to be false commits one to rejecting it as untrue. So I accept the rule: from $+(\neg A)$, infer $-A$. In general, though, the converse rule fails. From the premiss that A is untrue, we cannot always infer that A is false.

What form does the rule of *reductio* take in such a system? When the supposition that A is true yields (in tandem with background assumptions) the conclusions that B is true and that it is untrue, we may infer that A is untrue (given the same background assumptions). Thus we should accept the following rule of proof (in which X is an arbitrary set of signed formulae, perhaps empty):

(*Red*) From $X, +A \vdash +B$ and $X, +A \vdash -B$, infer $X \vdash -A$.\[24]/

From the fact that $+A$ yields a contradiction, we cannot always infer that A is false, but in special circumstances we can deduce instances of $+(\neg A)$. For present purposes, we may suppose that a sentence is untrue but not false only when it fails to say anything at all. Let us say that a set X of signed formulae *determines* a bare formula A when the correctness of all the members of X logically guarantees that A says something. Then we have the following rule for affirming a negated sentence:

$+\neg$-intro From $X \vdash -A$, infer $X \vdash +(\neg A)$ whenever X determines A.

---

[23] See Rumfitt 2000, where I gave a formalization of classical propositional logic in a calculus of signed formulae which has certain advantages over the familiar, purely affirmative formalization. In that paper, though, $-A$ signified that A was rejected *as false*—and not, as here, rejected *as untrue*. This difference plainly matters in the present context of discussion.

[24] Classical logicians, but not intuitionists, will also accept the corresponding rule with the signs reversed: from $X, -A \vdash +B$ and $X, -A \vdash -B$, infer $X \vdash +A$.

What light do these rules cast on our problematical derivation? What (*Red*) directly yields at line 10 is $-R$. However, we may apply $+\neg$-intro to deduce $+\neg R$: among the undischarged assumptions at this stage in the derivation is δR, which guarantees that R says something. As for the inference to line 11, (*Red*) and *modus ponens* combine to yield $-\neg\exists P(\delta P \wedge P)$. This combines with $+\neg\exists P(\delta P \wedge P)$ at line 18 to yield $-\exists P(\delta P \wedge P)$ at line 19. However, we cannot assert $+\neg\exists P(\delta P \wedge P)$, for the assumptions left undischarged at that stage of the derivation do not guarantee that $\ulcorner\exists P(\delta P \wedge P)\urcorner$ says anything. Correspondingly, the conclusion we reach at line 20 is simply $-\exists R\delta R$, not $+\neg\exists R\delta R$. That weaker conclusion, however, is quite sufficient to vindicate the present approach to the Liar. The supposition that Epimenides's words (as standardly construed) say something is demonstrably untrue; that is enough for us to be warranted in rejecting the hypothesis that they say something.

In this emended form, then, the argument goes through as far as line 20. What, though, of the next stage of the derivation? The positive logic takes us to $+\exists R\delta R$ at line 24, and (*Red*) duly yields $-\exists P(\delta P \wedge P)$ at line 25. We cannot, however, move forward to $+\neg\exists P(\delta P \wedge P)$. The remaining undischarged assumption of the derivation, viz. premiss (1), certainly does not determine $\ulcorner\exists P(\delta P \wedge P)\urcorner$. when this formula is standardly construed. Accordingly, the derivation's final conclusion $-\exists P(\delta P \wedge P)$ is entirely consistent with the thesis that Epimenides's utterance of $\ulcorner\neg\exists P(\delta P \wedge P)\urcorner$ fails to express a thought (as standardly construed).

This analysis vindicates an emended version of Mackie's solution to the liar's revenge. He wrote that λ's 'indeterminacy with respect to truth is of a kind which prevents our saying even that it is not true, and therefore from arguing, by a further step, that it is true' (1973, p.295). I say: λ's indeterminacy prevents us from asserting its negation and therefore from arguing that it is true after all; but we may, and should, reject λ as untrue.

The idea that we should deal with λ by rejecting is as untrue, rather than by affirming its negation, is not new. It was proposed thirty years ago by Terence Parsons (Parsons 1984). I hope to have contributed something Parsons did not provide—namely, a statement of the logical rules by which we can justify rejecting λ as untrue without falling into the trap of asserting its negation. We thereby improve on Smiley's insouciant acknowledgement that the laws of logic may be trumped by Sod's Law. The rule of $+\neg$-introduction states the condition which must be satisfied if we are to affirm a conclusion in the form $+(\neg A)$. For reasons already explained, there can be no syntactic test for satisfying that condition. All the same, our analysis shows us what we have to check for: we need to

verify that the correctness of the operative premisses and assumptions guarantees that *A* says something. In this way, we keep faith with the central project of formal logic: we do as much as can be done by formal methods to ensure that deductive arguments preserve truth. To check that *A* says something, we shall have to go beyond anything that those methods can accomplish. But we also need to do that when we check that the proponent of an argument has not equivocated between two senses of a term, or that the passage of time does not make a material difference to an argument formulated in the present tense.

We would be ensnared in paradox again if it were possible to introduce an external negation operator '~' for which $+\sim A$ were equivalent to $-A$. For in that case, Epimenides could cause renewed trouble by uttering $\ulcorner \sim\exists P(\delta P \wedge P)\urcorner$. As we saw, though, there is no good theoretical reason why it should be possible to introduce such a negation operator (see further Tappenden 1999). We would also get into trouble if we applied our rules to 'Everything a Cretan says should be rejected as untrue'. I think we just have to concede that problems would arise if the entire semantic machinery of the present paper were to be projected into the object language: the limitations on such projection mean (as in Kripke's theory) that 'the ghost of the Tarski hierarchy is still with us' (Kripke 1975, p.714). But the ghost is far less inhibiting than the hierarchy proper: the rules proposed enable us to do a great deal of semantic theorizing within our system.

Whether those rules solve all the 'nasty problems of intensionality' that afflict Ramsey's definition of truth remains to be determined. As we have seen, versions of that definition have attracted discerning philosophers, but I fear Strawson was too sanguine when he described it as 'innocuous' (1971, p.180): we have found a number of snakes lurking in the grass. Since Prior died, the task of driving them away and putting the definition on a firm foundation has been neglected. As we approach the centenary of his birth, it is high time the job was completed.\[25]/

[25] I have been presenting this material in lectures and classes for a good ten years, and have no record of those whose questions or remarks influenced its development. For comments that improved the final draft, though, I am indebted to Dorothy Edgington, Laurence Goldstein, Bob Hale, Keith Hossack, Daniel Isaacson, Nicholas Jones, James Levine, Guy Longworth, Hugh Mellor, Alex Silk, Hartley Slater, Dustin Tucker, Timothy Williamson, and Crispin Wright.

I finished this essay on 21 December 2013, the day Peter Geach died, aged 97. He was encouraging and kind to me when I was starting out, as he was to many tyro philosophers with logical interests. It was exciting to hear him recount his conversations with Wittgenstein about Frege. He would not have been fazed by the criticisms of §I: he enjoyed the cut and thrust of philosophical debate more than he liked being agreed with. I dedicate this essay to his memory.

REFERENCES

Bochvar, D.A. 1939. 'Ob odnom tréhznacnom isčislenii i égo priménii k analiza paradoksov klassičékogo răssirénnogo funkcional'nogo isčisléniá'. *Matématičeskij sbornik* **4**: 287-308. Translated by M. Bergmann as 'On a three-valued logical calculus and its application to the analysis of the paradoxes of the classical extended functional calculus' in *History and Philosophy of Logic* **2** (1981): 87-112.

Braithwaite, R.B. 1933. 'The nature of believing'. *Proceedings of the Aristotelian Society* **33**: 129-46.

Davidson, D.H. 1965. 'Theories of meaning and learnable languages'. In Y. Bar-Hillel. ed., *Logic, Methodology, and Philosophy of Science II* (New York: North Holland), pp.383-394. Reprinted in Davidson 1984, pp.3-15.

——. 1967. 'Truth and meaning'. *Synthese* **17**: 304-23. Page references are to the reprint in Davidson 1984, pp.16-36.

——. 1973. 'Radical interpretation'. *Dialectica* **27**: 313-28. Page references are to the reprint in Davidson 1984, pp.125-39.

——. 1976. 'Reply to Foster'. In Evans and McDowell, eds., 1976, pp.33-41. Page references are to the reprint in Davidson 1984, pp.171-79.

——. 1984. *Inquiries into Truth and Interpretation.* Oxford: Clarendon Press.

——. 1990. 'The structure and content of truth'. *The Journal of Philosophy* **87**: 279-328. Page references are to the reprint in Davidson, *Truth and Predication* (Cambridge, Mass.: Harvard University Press, 2005), pp.7-75.

Davies, M.K. 1981. *Meaning, Quantification, and Necessity*. London: Routledge.

Ducrot, O. 1984. *Le Dire et le Dit*. Paris: Editions de Minuit.

Evans, M.G.J. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.

Evans, M.G.J. and J.H.McDowell, eds. 1976. *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press.

Foster, J.A. 1976. 'Meaning and truth theory'. In Evans and McDowell, eds., 1976, pp.1-32.

Gauker, C. 2006. 'Against stepping back: a critique of contextualist approaches to the semantic paradoxes'. *Journal of Philosophical Logic* **35**: 393-422.

Geach, P.T. 1982. 'Truth and God'. *Proceedings of the Aristotelian Society, Supplementary Volumes* **56**: 83-97.

Glanzberg, M. 2001. 'The Liar in context'. *Philosophical Studies* **103**: 217-51.

——. 2004. 'A contextual-hierarchical approach to truth and the Liar Paradox'. *Journal of Philosophical Logic* **33**: 27-88.

Goldstein, L. 2006. 'Fibonacci, Yablo, and the cassationist approach to paradox'. *Mind* **115**: 867-89.

Higginbotham, J.T. 1992. 'Truth and understanding'. *Philosophical Studies* **65**: 3-16.

Horn, L.R. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.

Kripke, S.A. 1975. 'Outline of a theory of truth'. *The Journal of Philosophy* **72**: 690-716.

McDowell, J.H. 1982. 'Truth-value gaps'. In L.J. Cohen et al., eds., *Logic, Methodology, and Philosophy of Science VI* (New York, New York: North Holland), pp.299-313.

Mackie, J.L. 1973. *Truth, Probability, and Paradox: Studies in Philosophical Logic*. Oxford: Clarendon Press.

Mellor, D.H. 2102. ‘Successful semantics’. In his *Mind, Meaning, and Reality: Essays in Philosophy* (Oxford: Clarendon Press), pp.60-77.

Moore, J. 1999. ‘Propositions without identity’. *Noûs* **33**: 1-29.

Parsons, C.D. 1974. ‘The Liar Paradox’. *Journal of Philosophical Logic* **3**: 381-412.

Parsons, T. 1984. ‘Assertion, denial, and the Liar Paradox’. *Journal of Philosophical Logic* **13**: 137-52.

Priest, G. 1993. ‘Can contradictions be true? II’ *Proceedings of the Aristotelian Society Supplementary Volumes* **67**: 35-54.

Prior, A.N. 1958. ‘Epimenides the Cretan’. *The Journal of Symbolic Logic* **23**: 261-6.

——. 1961. ‘On a family of paradoxes’. *Notre Dame Journal of Formal Logic* **2**: 16-32.

——. 1971. *Objects of Thought*, ed. P.T. Geach and A.J.P. Kenny. Oxford: Clarendon Press.

Ramsey, F.P. 1927. ‘Facts and propositions’. *Proceedings of the Aristotelian Society Supplementary Volumes* **7**: 153-70. Page references are to the reprint in Ramsey, ed. D.H. Mellor, *Philosophical Papers* (Cambridge: Cambridge University Press, 1990), pp.34-51.

——. 1991. *On Truth*. Dordrecht: Kluwer.

Read, S.L. 2009. ‘Plural signification and the Liar paradox’. *Philosophical Studies* **145**: 363-75.

Rumfitt, I. 2000. ‘“Yes” and “No”’. *Mind* **109**: 781-823.

——. 2011. ‘Ramsey on truth and meaning’. In B.C.A. Morison and K. Ierodiakonou, eds., *Episteme, etc: Essays in Honour of Jonathan Barnes* (Oxford: Clarendon Press), pp.213-45.

——. Forthcoming. ‘Objects of thought’. In G. Ostertag, ed., *Meanings and Other Things: Essays on the Work of Stephen Schiffer* (New York: Oxford University Press).

Smiley, T.J. 1993. 'Can contradictions be true? I' *Proceedings of the Aristotelian Society Supplementary Volumes* **67**: 17-33.

Soames, S. 1999. *Understanding Truth*. New York: Oxford University Press.

Strawson, P.F. 1971. 'Meaning and truth'. In his *Logico-Linguistic Papers* (London: Methuen), pp.170-89.

Tappenden, J. 1999. 'Negation, denial and language change in philosophical logic'. In D.M. Gabbay and H. Wansing, eds., *What is Negation?* (Dordrecht: Kluwer), pp.261-98.

Tarski, A. 1935. 'Der Wahrheitsbegriff in den formalisierten Sprachen'. *Studia Philosophica* **I**: 261-405. Translated by J.H. Woodger as 'The concept of truth in formalized languages' in Tarski, eds. Woodger and J. Corcoran, *Logic, Semantics, Metamathematics* (Indianapolis: Hackett, 1983), pp.152-278.

——. 1944. 'The semantic conception of truth and the foundations of semantics'. *Philosophy and Phenomenological Research* **4**: 341-76.

Tucker, D. 2013. 'Outline of a theory of quantification'. In N. Griffin and B. Linsky, eds., *The Palgrave Centenary Companion to Principia Mathematica* (New York: Palgrave Macmillan), pp.337-65.

Williamson, T. 1998. 'Indefinite extensibility'. *Grazer Philosophische Studien* **55**: 1-24.

Yablo, S.J. 1993. 'Paradox without self-reference'. *Analysis* **53**: 252-3.