Grazer Philosophische Studien 55 (1998): 1-24

INDEFINITE EXTENSIBILITY*

Timothy WILLIAMSON
University of Edinburgh

§ 1

Of all the cases made against classical logic, Michael Dummett's is the most deeply considered. Issuing from a systematic and original conception of the discipline, it constitutes one of the most distinctive achievements of twentieth century British philosophy. Although Dummett builds on the work of Brouwer and Heyting, he provides the case against classical logic with a new, explicit and general foundation in the philosophy of language. Dummett's central arguments, widely celebrated if not widely endorsed, concern the implications of the relation between meaning and use for both the inference rules that govern logical connectives and the relation between truth and its recognition. It is less often noted that Dummett has a further argument against classical logic, one based on the semantic and set-theoretic paradoxes. That is the topic of this paper.

Dummett writes:

> The paradoxes – both the set-theoretic and the semantic paradoxes – result from our possessing indefinitely extensible concepts [...]. An indefinitely extensible concept is one for which, together with some determinate range or ranges of objects falling under it, we are given an intuitive principle whereby, if we have a sufficiently definite grasp of any one such range of objects, we can form, in terms of it, a conception of a more inclusive such range. [...] By the nature of the case, we can form no clear conception of the extension of an indefinitely extensible concept; any attempt to do so is liable to lead us into contradiction.[1]

* Thanks to Michael Glanzberg and participants at the 1997 Stirling conference on the philosophy of Michael Dummett for helpful comments.

1. Dummett (1993): 454. See also Dummett (1963) and (1991): 316-319, which attributes the idea to Russell.

This account surely reflects *something* in our experience of the paradoxes. Dummett goes on to argue that the logic of a language in which we can express indefinitely extensible concepts will be non-classical, because not all statements expressible in it will be determinately true or false; bivalence and excluded middle fail. He proposes intuitionistic logic for such languages.[2]

I will suggest that although the paradoxes do involve something like indefinite extensibility, it poses no threat to classical logic or two-valued semantics. Like Dummett, I will discuss philosophical motivation rather than technical detail.

From a purely technical perspective, intuitionistic logic presents no obvious advantage. In the simplest paradoxes, a plausible general principle turns out to have a substitution instance of the form $P \equiv \sim P$, which is inconsistent in both intuitionistic and classical logic. Adopting intuitionistic logic would not enable us to retain the plausible general principle while blocking the inference to a contradiction. Rather, Dummett's idea must be that a proper understanding of indefinite extensibility forces us to reject the philosophical picture that lends plausibility both to the general principle and to classical logic and two-valued semantics.

§2

Discussions of indefinite extensibility by Dummett and others have usually concentrated on the philosophy of mathematics. I will concentrate on the semantic paradoxes. Dummett applies his idea to them by classifying the concept of a statement as indefinitely extensible.[3] Presumably, it is so because, given a 'sufficiently definite grasp' of a range R of statements, we can make a statement not itself in R about the statements in R. Perhaps we could retain classical logic and avoid the semantic paradoxes by using implicit indices to restrict the application of each occurrence of 'true' to a definite range of statements. But Dummett allows that we can also make unrestricted true generalizations about all objects falling under an in-

2. Dummett (1993): 455; (1963); (1991): 319.
3. Dummett (1993): 454.

definitely extensible concept.[4] For example, we might make the unrestricted true generalization that no statement is both true and false. On Dummett's view, it is for unrestricted uses that classical logic fails. By contrast, intuitionistic logic is supposed to be unproblematic for them. Of course, the semantic paradoxes still arise for the unrestricted uses within intuitionistic logic; exactly how one is supposed to avoid them remains unclear.

The distinction between restricted and unrestricted uses of 'true' can be compared with Tyler Burge's claim that there are

> two uses of 'true' in natural language: indexical uses and schematic uses. A predicate is *indexical* on an occasion of use if and only if it has a definite, fixed extension (or extensional application) on that occasion that depends not only on the contextually appropriate conventional meaning of the predicate, but further on the immediate context of its use. A predicate is *schematic* on an occasion of use if and only if it lacks a definite extension on that occasion, but through its conventional use on that occasion provides general systematic constraints on the extension(s) of the same predicate (or importantly related ones) on other occasions of use.[5]

Burge diagnoses the semantic paradoxes as illicitly exploiting shifts of context. He postulates a system of levels; for each level i, there is a set of true$_i$ sentences and a set of pathological$_i$ sentences, and the Tarskian disquotational T-schema holds with respect to 'true' for all non-pathological$_i$ sentences. When 'true' is indexical, its level is determined contextually; as Kripke emphasized, the level cannot be determined just by the sentence-type in which 'true' occurs. To validate individual steps in the paradoxical derivations, one must assign different levels to different occurrences of 'true', and the contradictions vanish. The schematic 'true' is used to formulate general principles about the indexical 'true' To read 'true' in a paradox schematically would be to formulate a paradox-schema, each instance of which could be solved as before.

4. Dummett (1993): 455. Of course, other features of the statements in a definite range might require intuitionistic logic, on Dummett's view.

5. Postscript to Burge (1979) at Martin (1984): 116. What follows extrapolates significantly from Burge's account. See Parsons (1974) for a related approach.

How does the schematic 'true' compare with an indefinitely extensible concept of truth, and the indexical 'true' with restrictions of that concept? 'True' is indexical in its tacitly restricted uses, because the envisaged restrictions are to definite, fixed extensions which depend on the immediate context of use. Could 'true' be schematic in Burge's sense when it is used unrestrictedly in the way Dummett regards as legitimate to formulate generalizations such as 'No statement is both true and false'? Using 'true' schematically involves reflection on other occasions when it is used indexically; the latter uses are conceptually more primitive. For Burge, 'true' is indexical when used naively. By contrast, Dummett does not imply that naive uses are restricted to definite extensions; such restrictions are introduced later, after reflection on the paradoxes, to preserve consistency.

When using 'true' schematically, one lays down a schema; its instances come from interpreting 'true' indexically and assigning it levels. A schema commits one to each of its instances. Thus the schematic 's is true' is valid just in case s is true$_i$ and true$_j$ and .... (the dots are essential). By contrast, if 'true' can be restricted to F-statements or G-statements or ..., the unrestricted 's is true' is valid just in case s is either a true F-statement or a true G-statement or (the dots are again essential). Roughly: a schematic predication of 'true' behaves like a conjunction of indexical predications, an unrestricted predication of 'true' like a disjunction of restricted predications. The difference disappears in negative occurrences: the schematic 's is not true' is valid just in case s is neither true$_i$ nor true$_j$ nor ....; the unrestricted 's is not true' is valid just in case s is neither a true F-statement nor a true G-statement nor .... ('true' occurs negatively in 'No statement is both true and false' and when 's is true' is the antecedent of a conditional).

When s is true$_i$ but not true$_j$, neither the schematic 's is true' nor the schematic 's is not true' is valid. But this is not a genuine failure of bivalence. It resembles the fact that neither 'n is prime' nor 'n is not prime' is a valid generalization about natural numbers. The implicit generality of a schema means that its apparent negation is not its real negation. On Burge's view, classical logic holds for languages in which 'true' is indexical. Since 'true' is indexical in all instances of a schema, the latter, if a classical theorem, remains valid.

Thus the schematic '*s* is true or *s* is not true' is valid (compare supervaluationist semantics). By contrast, '*s* is true or *s* is not true' is not universally valid on Dummett's account when 'true' is unrestricted.

For Dummett, any attempt to form a definite conception of the extension of an indefinitely extensible concept is defeated by an intuitive principle that allows us to form a more inclusive conception. For Burge, the appearance of such an intuitive principle masks a context-shifting device that transforms a context in which the indexical 'true' has one definite extension into another context in which it has a more inclusive but still definite extension. Dummett does say of a concept he regards as indefinitely extensible:

> Better than describing [it] as having a hazy extension is to describe it as having an increasing sequence of extensions; what is hazy is the length of the sequence, which vanishes in the indiscernible distance.[6]

But he does not say that in each context of use the concept has exactly one of these extensions. Doing so would violate his account of the extension-enlarging intuitive principle and undermine his case against classical logic for indefinitely extensible concepts.

That the semantic paradoxes involve context-shifting has some plausibility, as Burge shows. Consider a simple formal version of the Liar. Let 'L1' name the sentence ~True(L1) ('L1 is not true'), so:

(1) L1 = '~True(L1)'

The Tarskian disquotational equivalence for '~True(L1)' is:

(2) True('~True(L1)') ≡ ~True(L1)

Using (1), we substitute its left-hand side for its right-hand side in (2):

(3) True(L1) ≡ ~True(L1)

(3) is logically inconsistent.

6. Dummett (1991): 316-317. The concept is that of ordinal number.

When we reflect on the paradox, it seems right to assert 'Whatever the semantic status of L1, it cannot be true, given this paradox', and therefore 'L1 is not true', i.e. L1. But in originally propounding the paradox, it was not right to assert L1, because it was paradoxical. Indexicality in 'true' would explain L1's change in assertibility.

Context-shifting is particularly vivid in a variant of König and Berry's paradoxes. Only finitely many natural numbers will ever be named by me. Thus some natural numbers will never be named by me. By the least number principle, there is the least natural number I shall never name; name it 'Minimus'. I have just named Minimus; contradiction. This is no *a priori* proof that I shall eventually name all natural numbers. Rather, reflection on an initial use of 'name' seems to yield a more inclusive use of the term, and a correspondingly higher value of 'Minimus', a process which can be iterated indefinitely. A semantic account of 'name' and 'Minimus' should not ignore such contextual effects.

Why, therefore, attribute indefinite extensibility rather than indexicality to semantic terms? Couldn't we explain the phenomena like Burge, and retain classical logic? Unfortunately, there appear to be indexicality-free semantic paradoxes with which Burge's account cannot cope.

§3

Burge replies to the objection that a Strengthened Liar Paradox infects his approach:

> one might suggest a sentence like (a), '(a) is not true at any level'. But this is not an English reading of any sentence in our formalization. Our theory is a theory of 'true', not 'true at a level'. From our viewpoint, the latter phrase represents a misguided attempt to quantify out the indexical character of 'true'; it has some of the incongruity of 'here at some place' No relativization will "deindexicalize" 'true'.[7]

This is unsatisfying. To try to deindexicalize 'here' by saying 'here at some place' is indeed confused, but English has a perfectly good word for what it is an attempt to say: 'somewhere' Quantification

7. Burge (1979) at Martin (1984): 108.

over levels certainly *appears* to make sense; if it does, we can give Burge's sentence (a) a paradoxical reading, even if that is not the reading it receives in current English, and we must confront that paradox. If that reading is non-indexical, as it apparently is, the Strengthened Liar requires a different solution, which suggests that Burge's diagnosis does not go deep enough. Burge contests these appearances. In his meta-language, the subscript in 'true$_i$' is not a variable of quantification but a schematic letter; if we say that every instance of a schema is true, that use of 'true' is itself either indexical or schematic over indexical uses, and a version of Burge's original diagnosis applies.[8]

We can consider the problem within a rigorous framework for treating indexicality, provided by David Kaplan. On Kaplan's account:

> If c is a context, then an occurrence of [a sentence] ϕ in c is true iff the content [proposition] expressed by ϕ in this context is true when evaluated with respect to the circumstances of the context.[9]

An appropriate level for 'true' would constitute one aspect of the context. For Kaplan, the contextual variable 'c' *is* open to quantification; he quantifies on it in defining logical truth as truth in every context (this is not Burge's meta-meta-linguistic quantification over the instances of a schema).[10] How else could we generalize about indexicality? We can therefore construct a paradox about logical truth.

Let True(s,c) say that the sentence s is true as uttered (with its present meaning) in the context c. Let 'L2' name the sentence $\forall c{\sim}\text{True}(\text{L2},c)$ ('L2 is true in no context'), so:

(1*) L2 = '$\forall c{\sim}\text{True}(\text{L2},c)$'

Let c_0 be a context. We cannot generally assert homophonic Tarskian disquotational equivalences for truth in c_0; unless c_0 is the very context of our theorizing, an indexical sentence may express a different content in c_0 from the one it expresses in the context of our theorizing, so the two sides of the biconditional may differ in truth-

8. Postscript to Burge (1979) at Martin (1984): 116.
9. Kaplan (1989): 522.
10. Kaplan (1989): 539; note also the definition of 'valid' at 547.

value. Thus I do not assert: 'I live in Edinburgh' is true as uttered by you if and only if I live in Edinburgh. However, let us assume that L2 expresses the same content in every context, as it appears to do. Then we have the homophonic equivalence:

(2*) $\text{True}(`\forall c{\sim}\text{True}(\text{L2},c)\text{'},c_0) \equiv \forall c{\sim}\text{True}(\text{L2},c)$

Using (1*), we substitute its left-hand side for its right-hand side in (2*):

(3*) $\text{True}(\text{L2},c_0) \equiv \forall c{\sim}\text{True}(\text{L2},c)$

The left-to-right direction of (3*) and universal instantiation give:

(4*) $\text{True}(\text{L2},c_0) \supset {\sim}\text{True}(\text{L2},c_0)$

By propositional logic, (4*) gives:

(5*) ${\sim}\text{True}(\text{L2},c_0)$

Since c_0 was an arbitrary context, we can generalize on (5*):

(6*) $\forall c{\sim}\text{True}(\text{L2},c)$

We combine (6*) with the right-to-left direction of (3*):

(7*) $\text{True}(\text{L2},c_0)$

(5*) and (7*) are contradictories.

The contentious premise is (2*). It could fail if L2 expressed different contents in different contexts. Like Kaplan, we are not concerned with what the sentence might express if its conventional linguistic meaning were different; what matters is what the sentence with its present meaning expresses with respect to various contexts. In Kaplan's terminology, the question is whether the character of L2 is a constant function, where the character of *e* maps each context to the content that *e* (with its given meaning) expresses in that context. It is hard to see where variation in the character of L2 could come from. The name 'L2' itself is not indexical; it names the same thing with respect to each context.[11] If we build enough into contexts, including local standards for assigning levels to occurrences of semantic terms, why should 'true in c' be indexical? Equally, why should quantification over absolutely all contexts be impossible, or irredeemably indexical?

We can discharge the premise (2*), and turn the argument into a proof by *reductio ad absurdum* of (8*), with (1*) the only premise:

(8*) $\sim\forall c_0(\text{True}(`\forall c{\sim}\text{True}(\text{L2},c)\text{'},c_0) \equiv \forall c{\sim}\text{True}(\text{L2},c))$

The homophonic equivalence (2*) does not hold in every context. No mere practical impossibility in fully deindexicalizing 'true in no context' would explain (8*); the impossibility would have to be one of deep logical principle. But then the real diagnosis would be whatever explained the impossibility of full deindexicalization, not the mere fact of indexicality, for the latter is consistent with full deindexicalization. If the problem with 'in no context' were indefinite extensibility, Dummett's alternative diagnosis would be vindicated.

§4

In spite of L2, shifts in the reference of semantic terms may be crucial to the semantic paradoxes. For shifts in the reference of a term constitute indexicality only when the term retains the same (conventional linguistic) meaning throughout. 'I' is indexical because we exploit the same meaning when you use it to refer to you and I use it to refer to me. Mere ambiguity does not constitute indexicality, because the shift in reference follows a shift in meaning. In principle, an ambiguous expression like 'bank' can have several non-indexical meanings. The paradox of L2 presupposes that 'true' does not shift in meaning across contexts, for the justification of (2*) is that since

11. On the non-standard account of proper names in Burge (1973), they do have a concealed indexical element. But the argument in the text requires only that we *could* use 'L2' as a genuine non-indexical constant to refer to L2, as we surely can, whether or not we call such a constant a 'proper name' In any case, contexts in which 'L2' refers to (e.g.) a penguin are irrelevant to the paradox.

$\forall c{\sim}\mathrm{True}(L2,c)$ is non-indexical, it expresses the same content in an arbitrary context c_0 as in the present context; but non-indexicality entails that the same content is expressed in different contexts only with respect to a fixed meaning. This suggests a more radical diagnosis of the semantic paradoxes: perhaps 'true' shifts its meaning, not just its reference.[12]

Shifts of meaning occur anyway in the case of 'Minimus' above. For since it is a name, shifts in its reference constitute shifts in its meaning; in Kaplan's terminology, names have constant non-indexical characters. The point does not essentially depend on the concept of a name. Let a *constant* be any term of constant character. We could introduce 'Minimus' as a constant by the stipulation that its character maps everything to the least natural number for which I shall never have a constant. Why not accept shifts of meaning for 'true' too?

At first sight, the revised diagnosis is very implausible. If 'true' shifts its meaning in paradoxical derivations, where does the new meaning come from, and how do we recognize it? I understand your use of 'I' to refer to you because I already know what 'I' means; I know the general rule that when someone produces 'I' it refers to

12. In his (1996) and elsewhere, Charles Travis argues on quite different grounds for ubiquitous context-dependence of a radical kind. According to Travis, meaning and reference underdetermine what is said ((1996): 455); truth depends on what is said. However, his account of what underdetermines what is said seems to conflate use and mention. He assumes that someone who says 'That is round' of a ball, using the words with their usual meanings, describes the ball as round, and argues that, given the way the ball is, whether it is true to describe it as round depends on the context of utterance. This commits Travis to rejecting the plausible principle that it is true to describe the ball as round if and only if it is round. But once we accept that 'round' is context-dependent, we should reject Travis's assumption, just as we should reject the assumption that if you say 'That is mine' of something, you describe it as mine. Quotation marks are needed. 'Mine', with its usual meaning, does not always refer to the property of being mine; why should 'round', with its usual meaning, always refer to the property of being round? With this correction, Travis's argument can be reconstructed as leading to the less radical conclusion that meaning alone ubiquitously underdetermines what is said. However, he simply takes the sameness in meaning of the words in his examples as obvious, and does not discuss the possibility of slight meaning changes, as postulated below.

them, so when you produce 'I' it refers to you (if you are using 'I' with its conventional English meaning). But if you used 'I' with a different meaning from mine, how could I know what you meant by it? Surely we can understand 'true' throughout the paradoxical derivations without anything like the *gestalt* shift we experience when we switch from one meaning of an ambiguous expression to another.

However, many shifts of meaning are far less stark than that between two unrelated meanings of an ambiguous word. Consider shifts from literal to metaphorical meanings. Even if, like Davidson, we deny that there is any such thing as metaphorical meaning, we can still consider the relation between the original literal meaning of a word such as 'pedestrian' [on foot] and its dead metaphorical meaning [uninspired], a new literal meaning.[13] Competent speakers of a language regularly understand metaphors on first hearing, without any special instruction, on the basis of context and their understanding of the words in their original literal meanings. Context guides them in modifying the original meaning. The context of following a paradoxical derivation might guide competent speakers to a new meaning of 'true' Still, it does not feel that way; it feels as though we are using 'true' in its strictest, most literal sense throughout.

§5

Lay the question of semantic shifts temporarily aside, and consider the meaning of 'true' in more detail. To simplify the formulation of the paradoxes, treat 'true' as a predicate of sentences in contexts, without prejudice to the nature of primary truth-bearers. One might initially suppose that the homophonic disquotational T-schema is

13. Davidson (1978). Davidson criticizes attempts to use similes to state metaphorical truth-conditions on the grounds that similes are trivially true, because 'everything is like everything, and in endless ways' ((1984): 254). On Davidson's view, the Nazi who says to the rabbi 'You are like dirt' says something trivially true. Davidson neglects the possibility that the reference of 'like' is context-dependent. If some respects of similarity are irrelevant in some contexts, a simile can be false or non-trivially true. Thus Davidson's argument leaves at least one candidate account of metaphorical truth-conditions standing.

central to the meaning of 'true'. However, as already noted, for sentences with indexicals the homophonic schema holds only of truth in the present context. Since we talk about truth in contexts other than that in which we are talking, we need something more general. This will do:

(T) $\text{Say}(s,c,\text{P}) \supset (\text{True}(s,c) \equiv \text{P})$

Informally: in a fixed context, a sentence, saying that P, is true if and only if P. In instances of (T), names of sentences of any language replace '*s*', names of contexts replace '*c*' and declarative sentences of the meta-language replace 'P'. One could also read Say(s,c,P) as 'In *c*, *s* expresses the proposition that P'. Given Say('P',c,P), we can recover the homophonic biconditional True('P',c) ≡ P. Thus 'I live in Edinburgh' as presently uttered by me is true if and only if I live in Edinburgh, because 'I live in Edinburgh' as presently uttered by me says that I live in Edinburgh. We cannot recover the unwanted homophonic biconditional that 'I live in Edinburgh' as uttered by you is true if and only if I live in Edinburgh, because 'I live in Edinburgh' as uttered by you does not say that I live in Edinburgh.

We can treat falsity likewise:

(F) $\text{Say}(s,c,\text{P}) \supset (\text{False}(s,c) \equiv \sim\text{P})$

For example, since 'I live in Edinburgh' as presently uttered by me says that I live in Edinburgh, 'I live in Edinburgh' as presently uttered by me is false if and only if I do not live in Edinburgh. (F) gives a better account of falsity than does its equation with non-truth or with having a true negation. The former equation wrongly implies that a sentence is false in a context in which it fails to say anything at all. The latter equation wrongly implies the impossibility of a simple language in which sentences can be false but negation cannot be expressed. An equation of falsity with having a true negation in an expanded language would put the weight on what was meant by 'negation'; if the negation of *s* is just a sentence true if and only if *s* is not true, this account is equivalent to the equation of falsity with non-truth. An account of falsity should use negation without mentioning it, as (F) does.

Given classical logic, (T) and (F) together imply the principle of bivalence in an appropriate form:

(B) $\mathrm{Say}(s,c,\mathrm{P}) \supset (\mathrm{True}(s,c) \vee \mathrm{False}(s,c))$

In a given context, a sentence that says something is either true or false. We should like a modified converse:

(BC) $\sim\exists \mathrm{P}\ \mathrm{Say}(s,c,\mathrm{P}) \supset \sim(\mathrm{True}(s,c) \vee \mathrm{False}(s,c))$

In a given context, a sentence that says nothing is neither true nor false. But (BC), unlike (B), quantifies over what is said. If such quantification is objectual, the variable 'P' must be in name position; but it is not, for it flanks ≡ in (T). If the quantification is substitutional, the sentences substituted for 'P' must be in the language to which the substitutional quantifiers belong, our language, for the semantics of the language gives universal substitutional quantification the truth-conditions of an infinite conjunction of the substitution instances, and the semantics can handle a conjunction only if it can handle each conjunct. Thus, in *c*, *s* would count as saying nothing if it says nothing sayable in our language; but surely a sentence of an alien language could be true or false by saying something not sayable in our language. For example, other languages have natural kind terms for kinds for which English has no such term. Although our language might be expanded to cover a given expressive deficiency, that would not answer the objection, for the semantics of our present language handles substitutional quantification only by handling each substitution instance, so the instances must be in our present language.

To say what we want to say by (BC), we need a kind of quantification that is neither objectual nor substitutional. It must be into sentence rather than name position, yet its truth-conditions must not be tied to the sentences of our language. Our use of expressions such as 'say nothing' strongly suggests that we already understand such a form of quantification, even though English forces it to share the vocabulary of objectual quantification ('no*thing*'). If we are permitted such quantification, we can give (BC) its intended meaning. We can also quantify (T) and (F) universally on 'P', in addition to the objectual variables 's' and 'c' We can then derive explicit definitions of truth and falsity in familiar form:

(T_{Def}) $\text{True}(s,c) \equiv \exists P(\text{Say}(s,c,P) \wedge P)$

(F_{Def}) $\text{False}(s,c) \equiv \exists P(\text{Say}(s,c,P) \wedge \sim P)$.

The right-to-left direction of (T_{Def}) follows from the universally quantified form of (T). (BC) entails $\text{True}(s,c) \supset \exists P\ \text{Say}(s,c,P)$, whence the quantified form of (T) gives the left-to-right direction of (T_{Def}). (F_{Def}) is derived similarly.

Non-substitutional quantification into sentence position (propositional quantification) has the further advantage of automatically excluding instances of (T) and (F) in which the sentence substituted for 'P' says nothing in the context of the theoretician's utterance; such instances would themselves say nothing. In the rule of universal instantiation, substituting a sentence that says nothing for 'P' is no more legitimate than substituting an empty name for an objectual variable. If we are not permitted propositional quantification, we must leave (T) and (F) as schemata, and express the corresponding restriction on their instances less directly. However, my main conclusions will not depend on the controversial legitimacy of propositional quantification; we can work with the unquantified schematic forms of (T) and (F).

No contradictions arise in a system of classical logic with propositional quantification, quotation marks (quantification into which is forbidden), the theorems (T_{Def}) and (F_{Def}), universally quantified forms of (T) and (F), and identities such as (1) and (1*). We can prove consistency by constructing an unintended model of the system in which formulas are treated as referring to truth-values, the propositional quantifiers range over truth-values and all formulas of the forms $\text{Say}(s,c,P)$, $\text{True}(s,c)$ and $\text{False}(s,c)$ are treated as false.

If we add substitution instances of $\text{Say}(\text{'P'},c,P)$ as axioms to the system, semantic paradoxes arise. Of course, we do not want all such formulas as axioms, independently of the semantic paradoxes, if the language contains indexicals. But, by (T) and (8*), we also must not add $\forall c_0\ \text{Say}(\text{'}\forall c{\sim}\text{True}(L2,c)\text{'},c_0,\forall c{\sim}\text{True}(L2,c))$, even if $\forall c{\sim}\text{True}(L2,c)$ is non-indexical. Similarly, the Liar shows that we must not add $\exists c \exists s$ Say(s,c,~True(s,c)); no sentence can say in a context that it is not true in that context. $\exists c \exists s$ Say(s,c,False(s,c)) leads equally directly to paradox, by (F); no sentence can say in a context that it is false in that context. But one can consistently add substitution instances of Say('P',c,P) as axioms for large classes of non-pathological formulas in place of 'P'; to this end, one can adapt any of several inductive techniques in the literature for capturing non-pathologicality. The semantic paradoxes are transformed into sound arguments for constraints on what can say what in what contexts.

§6

We still want to know *why* 'L2 is true in no context' does not say in every context that L2 is true in no context. Indeed, given that the sentence is non-indexical, we need to know why it does not say in *any* context that L2 is true in no context. For it appears to say exactly that.

Encountering a semantic paradox might prompt us to enlarge what we mean by 'say' We start with one set of correlative meanings for 'say', 'true' and 'false'; we use them to construct a sentence that says nothing in that sense of 'say'; but reflection on that sentence causes normal speakers to give 'say', 'true' and 'false' a new set of correlative meanings, much like the previous ones except that the sentence in question says something in the new sense of 'say'; the process can be repeated indefinitely. Normal speakers are not aware of the change, just as they are not aware of many ordinary processes of gradual semantic change. They feel themselves to be going on in the same way, but they are not. They are not applying the old rule to a new case; they are applying a very slightly different rule. As L2 indicates, this is no mere change of context.

Disquotational principles can appear to fix unique conceptual roles for 'true' and 'false', with no room for changes in meaning small enough to be missed by normal speakers. But (T) and (F) (or (T_{Def}) and (F_{Def})) show that the conceptual roles of 'true' and 'false' are fixed only relative to 'say'. It is much easier to understand how 'say' can undergo changes in meaning small enough to be missed by normal speakers. But small changes in the meaning of 'say' induce small changes in the meaning of 'true' and 'false', for a sentence that says something in one sense and not in another will be true or false in a sense correlative with the former and not in a sense correlative with the latter. What remains fixed is only the relation between 'say' and 'true' and 'false'

Sentences will be treated as sequences of sounds or marks, which have no intrinsic interpretation but receive different interpretations in different languages; for these purposes, a change of meaning is *ipso facto* a change of language. One could make the linguistic parameter explicit in (T), (F), (T_{Def}) and (F_{Def}). Although the contextual parameter is still needed to handle indexicality, we can suppress it for simplicity since we are no longer blaming the semantic paradoxes on indexicality.

Does a Strengthened Liar confound this approach, just as L2 confounded Burge's? We might let 'L3' name the sentence 'L3 is true in no language' To mimic the argument of §3, we would need to assume that 'L3 is true in no language' says in an arbitrary language that L3 is true in no language. But that assumption is false for many languages, in which the constituent words of L3 have quite different meanings. Context-dependence is intralinguistic; language-dependence is interlinguistic. Context-dependence is a feature of individual meanings; we can stipulate a non-indexical meaning for a term by using it to name a given object. By contrast, nothing we do with a given sequence of marks or sounds can ensure that it will not be used to mean something utterly different in another community. We can in principle eliminate mere indexicality on a given dimension by quantifying along that dimension, but the sequence of marks or sounds we make when we produce a quantified sentence is just as open as any other to alternative interpretations in other languages.

We could try to limit the variation by considering only those languages that we shall actually use while considering the paradoxes in what would ordinarily be called 'English'; call them 'versions'. Let L4 name the sentence 'L4 is true in no version' To mimic the argument of §3, we must assume that L4 says in an arbitrary version that L4 is true in no version. We should not accept that assumption. Consider first the counterfactual case in which there is only one version (perhaps nobody encounters the paradox). Use the subscript '0' to assign words their meanings in that version. Then, by (T), the argument shows that L4 does not say_0 that L4 is $true_0$ in no version. Presumably, L4 does not say_0 anything. In a second case, encounter with the paradox causes us to assign 'say' an enlarged meaning; use the subscript '1' to assign words their meanings in this new version. Then L4 $says_1$ that L4 is $true_0$ in no version, but it neither $says_0$ nor $says_1$ that L4 is $true_1$ in no version. The argument iterates as further versions arise. The role of the subscripts is to disambiguate 'true' and 'say', as we might disambiguate 'bank' by writing '$bank_1$' and '$bank_2$' The subscripts do not represent semantic complexity; '1' is no more an argument of 'true' in '$true_1$' than it is of 'bank' in '$bank_1$' Thus we cannot quantify on the subscripts; 'This sentence is $true_i$ for no i' is meaningless. We can understand the conjunction or disjunction of '$true_1$', ..., '$true_n$' if we already understand them individually, but not otherwise. Whatever our present understanding, encountering a paradox can cause us to reach a new understanding. We cannot construct a Strengthened Liar for this approach because we cannot anticipate our future understanding in our present meaning.

§7

Is the present approach a theory of levels? The possible meanings of 'true' do not constitute a systematic hierarchy; 'true' could mean anything in an alien language. Nor do its actual present and future meanings constitute a systematic hierarchy; they are a matter of historical accident, truncated within a finite time by the extinction of the human species. Even the relation between one meaning of 'true' and its successor is in part psychological; encountering a Liar sentence causes us to develop the latter from the former. This is not to deny that there are hierarchies of some possible meanings of 'true', as described by various theories in the literature on the paradoxes. Many different hierarchies exist, corresponding to different relations between possible meanings of 'true'. Read this way, different theories of levels are not rivals; they may be equally true by describing different hierarchies. Some of these hierarchies may even have significant applications. In a given context, we can choose to conform our use of 'true' and 'false' to one of these hierarchies. But reflection on a given hierarchy of meanings enables us to grasp new meanings outside that hierarchy, and thereby to recognize a new hierarchy. No single hierarchy gives an adequate treatment of all the paradoxes.

Even the term 'meaning' is presumably unstable, given its connections with 'say', 'true' and 'false'; the meaning of a sentence determines what it says in a context. These instabilities preclude us from stating in one go principles general enough to provide an exhaustive treatment of all semantic paradoxes. Further reflection always uncovers new paradoxes and new solutions. Our ability to conform our use of 'true' to different hierarchies in any case makes a unitary solution unlikely. The Protean nature of the material leaves us no alternative to a kind of intellectual opportunism. The present approach offers no more than a loose strategy in handling the paradoxes; if it is right, there is no better to be had, contrary to reasonable prior expectations. It does preserve some general formulas: classical logic, (T), (F), (T_{Def}), (F_{Def}). For most purposes, they are what matter. But their application to particular paradoxes is not mechanical; to some extent it must be determined *ad hoc*. For we may not grasp the relevant concepts until we confront the paradox.

We can call this approach a theory of levels in the minimal sense that it is naturally formulated using an unlimited supply of subscripts on 'say', 'true' and 'false' Methodologically, theories of levels start at a disadvantage. They postulate massive complexity not apparent in the data. Other things being equal, we should prefer a theory without levels. But levels have a habit of returning through the back door when they are kicked out of the front. For example, Kripke concedes that his theory of truth requires a meta-language more expressive than the object-language: 'The ghost of the Tarski hierarchy is still with us'.[14] Similarly, Gupta and Belnap criticize theories of levels for useless complexity, but later admit:

> the concept [...] that is appropriate for describing the behavior of the Ordinary Liar is not appropriate for the Strengthened Liar. To correctly describe the behavior of the latter we need to appeal to a higher-level notion [...]. This higher-level notion would itself manifest paradoxical behavior in the presence of vicious reference. And we would account

14. Kripke (1975): 714.

> for it in the same way. The higher-level paradoxes would demand a still higher-level notion for their description.[15]

The repeated failure of attempts to do without levels suggests that they are a genuine phenomenon, not a theoretical fiction.[16]

In this paper there is no room to survey the wide variety of purported solutions to the paradoxes. Brave attempts to do without levels continue to be made. We cannot be sure that none of them will succeed at less than exorbitant cost. Nevertheless, we have extensive grounds for pessimism. This paper is written on the hunch that levels are inescapable.

What of the set-theoretic paradoxes? Gödel's second incompleteness theorem notwithstanding, we have no serious reason to doubt the consistency of several current theories of sets with a classical background logic, including ZFC. However, their consistency does not immediately entail the consistency of the intuitive conceptions standardly used to motivate them, such as the iterative conception of sets.[17] For even if an intuitive conception σ entails the truth of every theorem of a consistent formal system Σ on its intended interpretation, σ might also entail the truth of every theorem of an inconsistent extension Σ^+ of Σ on its intended interpretation, in which case σ would itself be inconsistent. That the consistency of ZFC implies the consistency of the iterative conception of sets is far from obvious.

Isn't the consistency and coherence of the iterative conception of sets intuitively obvious? Someone may argue otherwise, on the grounds that even the iterative concept of set is indefinitely extensible. In particular, it may be suggested, however large we suppose the universe of sets to be, our ability to enlarge it by adding proper classes shows that we have not really exhausted the full import of the iterative conception, and therefore that that conception is incoherent after all. This argument assumes that, on the iterative conception, proper classes are really sets under another name, so that denying them the name 'set' is an *ad hoc* verbal evasion, inconsistent with

15. Gupta and Belnap (1993): 256; the general criticism of theories of levels is at 103-104.

16. Priest (1987) and (1995) criticizes other attempts to treat the paradoxes consistently without a theory of levels.

17. The classic account is Boolos (1971).

the underlying conception. The Burali-Forti paradox can be used likewise against the concept of ordinal, which the iterative conception needs to index stages of iteration.

The present view has the resources to interpret the iterative conception more charitably, while helping to accommodate the intuitions behind the argument against it. For given any reasonable assignment of meaning to the word 'set' we can assign it a more inclusive meaning while feeling that we are going on in the same way, and make correlative changes to the words of an iterative account of sets, to preserve it too. The inconsistency is not in any one meaning we assign the iterative account; it is in the attempt to combine all the different meanings that we could reasonably assign it into a single super-meaning.

This is not to say that the present view is committed to the iterative conception of sets, or to any other particular conception of them. Rather, it provides similar defences of several different conceptions of sets against the charge of incoherence. The choice between those conceptions must be made on more specific grounds.

§8

The present approach postulates a phenomenon something like indefinite extensibility, in the expressive power of languages. But there are no indefinitely extensible concepts. Rather, a word such as 'true' stands for slightly different concepts (has slightly different meanings) in successive languages. We no more grasp a single super-concept that encapsulates all the concepts for which 'true' might come to stand than we can give a word a single super-meaning that encapsulates all its possible metaphorical extensions. Since the extensibility is not internal to the meanings, any objection to classical logic from indefinite extensibility lapses, and we can accept the principle of bivalence: if a sentence says something, it is true or false. Naturally, those words will shift in meaning when 'say', 'true' and 'false' do. We can freely formulate absolutely unrestricted generalizations about everything said, everything true or everything false; but our words will soon receive new meanings.

Since 'say', 'true' and 'false' are subject to slight unnoticed shifts in meaning, we should expect them to be *vague*. For example, The Liar will sometimes be a borderline case for truth, neither clearly true nor clearly not true. Perhaps The Liar is clearly true$_i$ and clearly not true$_j$, yet it is unclear whether truth (in the current sense of the term) is truth$_i$, truth$_j$ or neither. However, I have argued elsewhere that vagueness is an epistemic phenomenon quite consistent with classical logic and the principle of bivalence.[18] In a borderline case, a vague concept applies or fails to apply, but those who grasp the concept are in no position to know which, because they are not sensitive to every small difference between the boundaries of concepts. Thus vagueness ubiquitously involves just the kind of limit on our powers of conceptual discrimination that explains our failure to notice the slight shifts in meaning postulated by the present account of the semantic and set-theoretic paradoxes.

When Dummett postulates an indefinitely extensible concept, he faces a methodological challenge. The paradoxes suggest that the appearance of a single concept may be an illusion. Although we use a single word, we may be equivocating. We appear to have an intuitive principle whereby, if we have a 'sufficiently definite grasp' of any one 'determinate range' of objects falling under the supposed concept, 'we can form, in terms of it, a conception of a more inclusive such range'; but perhaps we merely have a tendency to replace concepts by more inclusive concepts. How can Dummett allay this suspicion? Certainly, it is not a case in which one word has several unrelated meanings; nobody suggested that it was. The suspicion to be allayed is that the meanings are closely related yet distinct, and that the new meanings are grasped by a process similar to metaphorical extension. Ordinarily, we might answer a charge of equivocation by showing that a single explicit rule could generate all the uses in question of the term. In the present case, however, such rules generate paradoxes. A concept governed by a paradoxical rule is incoherent. But, on Dummett's view, we can reason soundly with indefinitely extensible concepts, using intuitionistic logic. Such concepts are therefore not genuinely incoherent, and so are not governed by genuinely paradoxical rules. The attempt to answer the charge of equivocation by citing an explicit rule undermines itself.

18. Williamson (1994) (at 197 the semantic paradoxes are briefly discussed).

Could indefinitely extensible concepts be characterized by non-paradoxical rules? All previous attempts to formulate such rules have failed; the suggestion is unpromising. Suppose, however, that a non-paradoxical rule can be formulated. For Dummett's purposes, it must be non-paradoxical in intuitionistic logic. If it has a non-paradoxical analogue in classical logic, Dummett's argument for indefinite extensibility lapses. But we are unlikely to find a non-paradoxical rule in intuitionistic logic with no non-paradoxical analogue in classical logic, for we can expect to use the formal correspondences between intuitionistic logic and classical modal logic to construct such an analogue. Specifically, there are recursively defined syntactic transformations of every formula α of quantified intuitionistic logic IQC into a formula α' of the quantified classical modal logic QS4 such that $\vdash_{IQC} \alpha$ if and only if $\vdash_{QS4} \alpha'$.[19] Moreover, $(\neg\alpha)'$ can be chosen as $\sim\Box\alpha'$, where $\sim$ and $\neg$ are classical and intuitionistic negation respectively, so $\vdash_{QS4} \sim\Box\alpha'$ only if $\vdash_{IQC} \neg\alpha$. Since $\vdash_{QS4} \Box\alpha' \supset \alpha'$, $\vdash_{QS4} \sim\alpha'$ only if $\vdash_{IQC} \neg\alpha$. Thus if α is the intuitionistic principle governing the indefinitely extensible concept, the corresponding classical principle α' is inconsistent in classical logic only if α was already inconsistent in intuitionistic logic. Contrapositively, if α is intuitionistically non-paradoxical, then α' is classically non-paradoxical.

The notion of an indefinitely extensible concept is an insufficiently radical response to the paradoxes. They involve a kind of indefinite extensibility so far-reaching that it overflows any single concept. The nature of a concept is to be graspable, and therefore limited. In the paradoxes, we use a limit to grasp a concept that transcends that limit, a concept with a new limit of its own. Indefinite extensibility is a way of generating new concepts, not something internal to a single concept. The limitless possibility of new concepts never yields limitless concepts. It never yields concepts that violate classical logic and the principle of bivalence.

19. Troelstra (1986) briefly summarizes the relevant results.

REFERENCES

Boolos, G. (1971): 'The iterative conception of set', *Journal of Philosophy* 68: 215-232.

Burge, T. (1973): 'Reference and proper names', *Journal of Philosophy* 70: 425-439.

— (1979): 'Semantical paradox', *Journal of Philosophy* 76: 169-198. Reprinted with postscript in Martin (1984).

Davidson, D. (1978): 'What metaphors mean', *Critical Inquiry* 5: 31-47. Reprinted in Davidson (1984).

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

Dummett, M. (1963): 'The philosophical significance of Gödel's Theorem' *Ratio* 5: 140-155. Reprinted in Dummett (1978).

— (1978): *Truth and Other Enigmas* (London: Duckworth).

— (1991): *Frege: Philosophy of Mathematics* (London: Duckworth).

— (1993): *The Seas of Language* (Oxford: Clarendon Press).

Gupta, A., and Belnap, N. (1993): *The Revision Theory of Truth* (Cambridge, MA: MIT Press).

Kaplan, D. (1989): 'Demonstratives', in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan* (Oxford: Oxford University Press).

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

Martin, R. (ed.) (1984): *Recent Essays on Truth and The Liar Paradox* (Oxford: Clarendon Press).

Parsons, C. (1974): 'The Liar Paradox', *Journal of Philosophical Logic* 3: 381-412. Reprinted with a postscript in Martin (1984).

Priest, G. (1987): *In Contradiction* (Dordrecht: Kluwer Academic Publishers).

— (1995): *Beyond the Limits of Thought* (Cambridge: Cambridge University Press).

Travis, C. (1996): 'Meaning's role in truth', *Mind* 105: 451-466.

Troelstra, A.S. (1986): 'Introductory note to *1933f*', in Kurt Gödel, *Collected Works Volume I: Publications 1929-1936*, ed. S. Feferman *et al.*

(Oxford, Clarendon Press).
Williamson, T. (1994): *Vagueness* (London: Routledge).