

SORITES

σωρίτης

An International Electronic Magazine of Analytical Philosophy
Indexed and Abstracted in *THE PHILOSOPHER'S INDEX*

ISSN 1135-1349

Legal Deposit Registration: M 14867-1995

Editor: Lorenzo Peña

(Spanish Institute for Advanced Studies)

Associate Editor: Txetxu Ausín

(European University, Madrid)

Board of Editorial Consultants:

Jean-Yves Béziau, Enrique Alonso, Guillermo Hurtado,
Manuel Liz, Raymundo Morado, Raúl Orayen

Regular-Mail Address:

Prof. Lorenzo Peña

CSIC [Spanish Institute for Advanced Studies]

Department of Theoretical Philosophy

Pinar 25

E-28006 Madrid

Spain

Fax +3491 564 52 52

Voice Tph +3491 411 70 60, ext 18

INTERNET ACCESS:

<<http://sorites.net/>>

<<http://www.ifs.csic.es/sorites/>>

<<http://www.filosoficas.unam.mx/~sorites/>>

Editorial e-mail inbox (esp. for submissions): <sorites@ifs.csic.es>

Inquiries and subscription-requests: <sorites@pinar2.csic.es>

Issue #13 — October 2001

SORITES (ΣΩΡΙΤΗΣ)

ISSN 1135-1349

Issue #13 — October 2001

Copyright © by SORITES and the authors

MAIN INTERNET ACCESS:

<<http://sorites.net/>>

<<http://www.ifs.csic.es/sorites/>>

<<http://www.filosoficas.unam.mx/~sorites>>

<sorites@ifs.csic.es> (Editorial e-mail inbox, esp. for submissions)

<sorites@pinar2.csic.es> (Inquiries and subscription-requests)

SORITES

ISSN 1135-1349

ROLL OF REFEREES

Rainer Born Johannes-Kepler Universitaet Linz (Austria)
Amedeo Conte University of Pavia (Italy)
Newton C.A. da Costa University of São Paulo (Brazil)
Marcelo Dascal University of Tel Aviv (Israel)
Dorothy Edgington Birbeck College (London, UK)
Graeme Forbes Tulane University (New Orleans, Louisiana, USA)
Manuel García-Carpintero University of Barcelona (Spain)
Laurence Goldstein University of Hong Kong (Hong Kong)
Jorge Gracia State University of New York, Buffalo (USA)
Nicholas Griffin McMaster University (Hamilton, Ontario, Canada)
Rudolf Haller Karl-Franzens-Universitaet Graz (Austria)
Terence Horgan University of Memphis (Tennessee, USA)
Victoria Iturralde Univ. of the Basque Country (San Sebastian, Spain)
Tomis E. Kapitan Northern Illinois University (USA)
Manuel Liz University of La Laguna (Canary Islands, Spain)
Peter Menzies Australian National University (Canberra, Australia)
Carlos Moya University of Valencia (Spain)
Kevin Mulligan University of Geneva (Switzerland)
Jesús Padilla-Gálvez Johannes-Kepler Universitaet Linz (Austria)
Philip Pettit Australian National University (Canberra, Australia)
Graham Priest University of Queensland (Brisbane, Australia)
Eduardo Rabossi University of Buenos Aires (Argentina)
David-Hillel Ruben London School of Economics (London, UK)
Mark Sainsbury King's College (London, UK)
Daniel Schulthess University of Neuchâtel (Switzerland)
Peter Simons University of Leeds (Leeds, UK)
Ernest Sosa Brown University (Providence, Rhode Island, USA)
Friedrich Stadler Institut «Wien Kreis» (Vienna, Austria)

SORITES

ISSN 1135-1349

Issue #13 — October 2001

TABLE OF CONTENTS

— Abstracts of the Papers	03
— «Non-Conceptually Contentful Attitudes in Interpretation» by Daniel Laurier	06
— «Metaphor and Meaning» by Alec Hyslop	23
— «The Justification of Deduction» by Silvio Pinto	33
— «New Systemic Hypothesis of Ageing» by A. Kolomytsev	48
— «Truth Conditions Without Interpretation» by John Collins	52
— «Identity, Analyticity and Epistemic Conservatism» by Paul Tomassi	72
— «Gorgias the Sophist on Not Being. A Wittgensteinian Interpretation» by Michael Bakaoukas	80
— «Privacy, Individuation, and Recognition» by Michael Ming Yang	90
— Copyright Notice and Legal Disclaimer	99
— Release Notice	101

SORITES (ΣΩΡΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 3-5.

Abstracts of the Papers

Copyright © by SORITES and the authors

ABSTRACTS OF THE PAPERS

Non-Conceptually Contentful Attitudes in Interpretation

Daniel Laurier

<Daniel.Laurier@umontreal.ca>

Brandom's book *Making It Explicit* defends Davidson's claim that conceptual thought can arise only on the background of a practice of mutual interpretation, without endorsing the further view that one can be a thinker only if one has the concept of a concept. This involves (*inter alia*) giving an account of conceptual content in terms of what Brandom calls practical deontic attitudes. In this paper, I make a plea for the conclusion that these practical attitudes are best seen as intentional, but non-conceptually contentful. In particular, I argue that the hypothesis that Brandom's practical deontic attitudes are non-conceptually contentful wouldn't conflict with his view that non-conceptual intentionality is merely derivative. I then explore some of the implications which this hypothesis might have with respect to various forms of «intentional ascent».

Metaphor and Meaning

Alec Hyslop

<A.Hyslop@latrobe.edu.au>

The paper argues, against Davidson, that metaphorical utterances involve meaning other than literal meaning. The kind of meaning is a particular case of contextual meaning. It is argued that metaphorical meaning is not a case of speaker's meaning (Searle), nor is it occasion meaning (Beardsley). I offer an explanation of why those metaphors that are not paraphrasable cannot be paraphrased.

The Justification of Deduction

Silvio Pinto

<silvio@filosoficas.unam.mx>

According to Michael Dummett, deductive inference stands in need of justification which must be provided by the theory of meaning for natural language. Such a theory, he insists,

should deliver an explanation for the two essential features of deduction: validity and fruitfulness. Dummett claims that only a molecularist theory of meaning could offer the desired justification. In this paper, I will consider and criticize his solution to the problem of the justification of deduction: the so-called molecular verificationist explanation. My aim here will be to show that Dummett's solution does not succeed in reconciling the conflicting demands of the respective explanations of validity and fruitfulness.

New Systemic Hypothesis of Ageing

Alexey Kolomiytsev

<betrg@pbox.ttn.ru>

A new evolutionary hypothesis on ageing is presented. The developmental process is determined by the kinetic curve of population growth which is typical for every cell association. Multicellular organisms are considered systems that consist of cellular associations in symbiotic interaction. One of those associations is dominating and determines the developmental kinetics of the whole organism and its lifespan.

Truth Conditions Without Interpretation

John Collins

<JCollins42@compuserve.com>

Davidson has given us two theses: (i) Tarski's format for truth definitions provides a format for theories of meaning and (ii) that the justification for a theory of language **L** as one of meaning is based upon the theory affording an informative interpretation of **L**-speakers. It will be argued, on the basis of a consideration of compositionality, that the Tarski format can indeed be re-jigged in line with (i). On the other hand, in opposition to (ii), I shall commend a cognitive understanding of semantic competence under which interpretation is at best of marginal interest. In defence of this move, it will be argued that once compositionality and other commonly held constraints are recognised as being in fact empirical ones, then no principled distinction can be made between them and other constraints which militate against semantic competence being socially constituted.

Identity, Analyticity and Epistemic Conservatism

Paul Tomassi

<p.tomassi@abdn.ac.uk>

In the first instance, the paper proposes a response to W.V.O. Quine's infamous attack on the analytic-synthetic distinction which attempts to carve out a core notion of analyticity by strictly delimiting the extension of that concept. The resulting position — *epistemic conservatism* — provides a platform for a significant epistemic challenge to essentialist positions of the kind proposed by Saul Kripke and Hilary Putnam: under exactly which kinds of circumstance are we warranted in asserting that we have grasped the truth of an identity-

statement of the requisite kind? In the absence of a clear and complete response to that epistemic challenge, the paper concludes that the Kripke-Putnam case remains not proven.

Gorgias the Sophist on not being. A Wittgensteinian Interpretation

Michael Bakaoukas

<bakaoukas_michael@lycos.com>

Provocative aphorisms of the most notable fifth century Sophist, Gorgias, such as «Nothing actually exists» or his realist tenet that «it is not speech (logos) that serves to reveal the external object, but the external object that proves to be explanatory of speech» have been subject to endless cycles of interpretation. I do not propose to offer here a full analysis of Gorgias' thought, but rather some new suggestions as to how to interpret Gorgias on the basis of how scholars have interpreted Gorgias. I shall put special emphasis on the analytical approaches to Gorgias' thought.

Privacy, Individuation, and Recognition

Michael Ming Yang

<mchly@juno.com>

In this paper I examine Wittgenstein's private language argument and Ayer's counter argument. (1) I argue that the language-game approach is of transcendental character in the sense that it concerns the logical structure of human activity, which underlies concrete linguistic practices as well as operations of consciousness. Failure in recognizing this results in much confusion. (2) I demonstrate that the key issue concerning private language is not the problem of correctness of identification as commonly believed, but the social nature of individuation. (3) I conclude that there is no such primary recognition as Ayer assumed, and sensation can only be recognized through the network of human action. If one believes that sensation (assisted by memory) can be the sole basis upon which language and knowledge are maintained, then certain preference on privacy seems inevitable, and then skepticism as well seems unavoidable.

SORITES (ΣΩΡΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 6-22

Non-Conceptually Contentful Attitudes in Interpretation

Copyright © by SORITES and Daniel Laurier

Non-Conceptually Contentful Attitudes in Interpretation

Daniel Laurier

Abstract

Brandom's book *Making It Explicit* defends Davidson's claim that conceptual thought can arise only on the background of a practice of mutual interpretation, without endorsing the further view that one can be a thinker only if one has the concept of a concept. This involves (*inter alia*) giving an account of conceptual content in terms of what Brandom calls practical deontic attitudes. In this paper, I make a plea for the conclusion that these practical attitudes are best seen as intentional, but non-conceptually contentful. In particular, I argue that the hypothesis that Brandom's practical deontic attitudes are non-conceptually contentful wouldn't conflict with his view that non-conceptual intentionality is merely derivative. I then explore some of the implications which this hypothesis might have with respect to various forms of «intentional ascent».

§1.— Practical Attitudes as Non-Conceptually Contentful

It is well known that for Davidson, for any individual to have conceptual thoughts at all, he must have the concept of belief, which is tantamount to saying that he must have the concept of a concept (the concept of correct and incorrect application of a concept). Since (on Davidson's account) one can only have this concept if one is an interpreter of the speech of others, it follows that only a community whose members are capable of mutually interpreting (i.e. attributing thoughts to) each other could count as a community of thinkers (and that there can be no such thing as a solitary thinker).

One useful way of reading Brandom's work is as propounding an extensive defense of Davidson's claim that conceptual thought can arise only on the background of a practice of mutual interpretation (and thus as endorsing a fundamentally interpretationist view of content, or as he would probably prefer to say, a «phenomenalist» view of content), *but without endorsing* the further view that one can be a thinker (and hence an interpreter) only if one has the concept of a concept (and hence the capacity to *conceptually think that* someone else thinks anything).

The leading intuition behind Davidson's and Brandom's common attitude towards matters of meaning and content is that an expression can mean anything only insofar as, and in virtue of the fact that, it is *taken* to mean it. In Brandom's favored words, this becomes the view that what one is committed or entitled to, and hence what it is appropriate or inappropriate for one to do or not to do, depends on what one is *taken* to be committed or entitled to. This intuition, however, is counterbalanced by the opposite intuition that one may *objectively* be committed or entitled to something, even when one is not taken to be so committed or entitled by anyone (and conversely, that one may objectively not be committed or entitled to what everyone, including oneself, takes one to be committed or entitled). The challenge is thus to provide a constructive account of norms that is yet capable of sustaining a real distinction between being

correct or appropriate, and being treated as correct or appropriate (by anyone, including the community as a whole). It is one of Brandom's grand contention to have taken up this challenge and shown how these two conflicting intuitions can be reconciled and how social practices of a certain kind can *institute* objective norms and *confer* objective conceptual contents on expressions and performances. As Brandom puts it in his preface (1994: xviii):

A fundamental criterion of adequacy of the account [to be propounded here] is that the theorist not attach semantic contents to expressions by *stipulation*; it must always be shown how such contents can be *conferred* on expressions by the scorekeeping activities the theorist attributes to the linguistic practitioners themselves. That is, the aim is to present conditions on an interpretation of a community as discursive scorekeepers that are sufficient (though perhaps not necessary) to ensure that interpreting the community as engaged in *those* implicitly normative practices *is* interpreting them as taking or treating their speech acts as expressing the sorts of semantic contents in question.¹

As far as I understand what Brandom is getting at here, his aim is to specify a set of conditions such that, if a community would engage in practices meeting these conditions, then its members would, thereby, be taking some of their performances as having content, and these performances would thereby have been endowed with content. This raises two questions: (i) what must a *practice* look like, in order for its practitioners to count as taking or treating something as having content, and (ii) how is the fact that something is *taken* (in practice) to have content supposed to make it the case that it *has* content (in some more robust sense)? In what follows, I'm going to ignore the second question (deferring it to another paper), and focus on some of the issues pertaining to the first.

The important thing to keep in mind for this discussion is that according to Brandom, a discursive practice is essentially one in which some performances have the power to alter the normative statuses of the individuals involved in it, i.e., to alter what they are committed or entitled to do. In this «deontic» perspective, to make an assertion is to acquire (and to believe something is to have) a (discursive) normative status of a certain kind, which is called a «doxastic commitment». Both the fact that a given performance has the force of an assertion, and the content of this assertion, depend on the inferential role of this performance, which is itself determined by the set of further performances (of the same type) to which the participants are committed or entitled, in virtue of the fact that this performance has been issued. I cannot rehearse the details of Brandom's specific brand of inferential semantics here; it should suffice to recall that he conceives of conceptual content in terms of inferential articulation, and inferential articulation in terms of inheritance (and exclusion) relations among deontic statuses.

From this point of view, the notion of conceptual content is to be explained in terms of deontic statuses (i.e., in terms of what the members of the community are committed or entitled to do), and the notion of deontic status is, in turn, to be explained in terms of *practical deontic attitudes*, that is to say, in terms of attitudes of acknowledging or attributing, *in practice*, such and such deontic statuses.

As Brandom strongly emphasizes, for a normative account of content to avoid regress or circularity, a performance's being correct or incorrect cannot depend on its being *judged or conceived* to be correct or incorrect (or on anyone's having the capacity to judge or conceive it to be correct or incorrect); but since it must nonetheless depend somehow on the activities of those who produce and consume it, there arises the need to construe what it is to take a

¹ See also Brandom (1994: xiii, xxii, 7, 61, 155, 190) for some other significant statements.

performance to be correct or incorrect (or to attribute a deontic status) in a way that doesn't equate it with any *conceptually contentful* state or attitude (or more generally, in such a way that the capacity to take a performance to be correct or incorrect doesn't presuppose any capacity for conceptual thought). In this respect at least, Brandom's practical deontic attitudes (taking a performance as correct, acknowledging or attributing a deontic status) are strikingly analogous to Davidson's non-individuative attitudes² of holding-true and preferring-true (i.e., those in terms of which the data of radical interpretation are to be describable), and raise the same kind of worries. These are worries concerning the legitimacy of appealing to non-conceptual intentionality in an account of conceptual intentionality.

In order to be in a position to deal with these issues, I must first be more explicit about the impact that the fact that some type of phenomenon «depends» on another is supposed to have on the way in which it can be explained or accounted for. I submit that, in the present context, the intuition behind such talk of «dependence» can be captured by taking the claim that some type of phenomenon X depends on another type of phenomenon Y as equivalent to the claim that any explanation of X involves either Y or an explanation of Y (in the sense that it contains ingredients sufficient to provide an explanation of Y). In other words, I submit that X depends on Y if and only if it is impossible to account for X without either invoking Y or being in a position to provide an account of Y as well.

Let's now ask which dependence relations could plausibly hold between conceptual and non-conceptual content. It may be useful, here, to observe that the idea that thought depends on language is naturally understood as implying, at least, that any system capable of having thoughts has mastered some system of symbolic communication. This strongly suggests that, in the same way, (i) the claim that conceptual content depends on non-conceptual content implies that no system can have the capacity to be in any conceptually contentful state unless it also has the capacity to be in some non-conceptually contentful state, and (ii) the claim that non-conceptual content depends on conceptual content implies that no system can have the capacity to be in any non-conceptually contentful state unless it also has the capacity to be in some conceptually contentful state. But then, it goes almost without saying that whoever is willing to grant that there are non-conceptual contents will be likely to hold that conceptual content depends on non-conceptual content, and that whoever admits further that non-linguistic animals and infants can be in non-conceptually contentful states, will assume that non-conceptual content does not depend on conceptual content. This implies that it must be possible to account for non-conceptual content without relying on conceptual content or being in a position to account for conceptual content, and precludes the coherentist strategy of simultaneously accounting for both conceptual and non-conceptual content.

But one must be very careful here. For the fact that some system may have the capacity to be in *some* non-conceptually contentful states while lacking the capacity to be in any conceptually contentful state obviously *does not* entail that *all* non-conceptually contentful states are accessible to some system lacking the capacity to be in conceptually contentful states. That non-conceptual content doesn't depend on conceptual content doesn't preclude the possibility that *some* non-conceptual thoughts can be had only by systems which are capable of conceptual thought, or even the possibility that some types of non-conceptual contents

² It seems plausible to assume that these attitudes can only have non-conceptual content, despite the fact that Davidson arguably doesn't allow for (literal) non-conceptual content. As I suggest below that this may also be the case of Brandom's practical deontic attitudes, the analogy seems warranted.

depend on conceptual content (i.e., on the capacity to be in some conceptually contentful states).

In other words, the claim that non-conceptual content does not depend on conceptual content should *not* be confused with the claim that no type of non-conceptual content depends on conceptual content. The first claim implies that it be possible to account for the capacity to be in a non-conceptually contentful state without relying on (or be in a position to provide) an account of the capacity to be in a conceptually contentful state, *but not* that it should be possible, for any given non-conceptual content, to account for the capacity to be in a state with this content, without relying on (or be in a position to provide) an account of the capacity to be in a conceptually contentful state.

What I want to suggest, at this point, is that there is nothing to prevent one from holding that Brandom's basic practical attitudes, if not conceptually contentful, may yet turn out to be *non-conceptually* contentful, and thus intentional, in some weaker sense. In particular, no incoherence or circularity would ensue from this assumption, if it is granted that admitting the notion of non-conceptual content almost forces one to assume that non-conceptual content doesn't depend on conceptual content. Moreover, even if it must be acknowledged that the capacity to practically attribute *discursive* normative statuses entails the capacity to be in conceptually contentful states (or in other words, that a system cannot have the capacity to practically attribute discursive statuses unless it also has the capacity to undertake such statuses), this doesn't conflict with the view that non-conceptual content doesn't depend on conceptual content. For this implies, at most, that *some* practical (and non-conceptually contentful) deontic attitudes (namely, those which consist in practically attributing or undertaking *discursive* statuses) depend on the capacity to be in conceptually contentful states, which is exactly what one would expect, if these practical deontic attitudes were «constitutive» of conceptual content, as Brandom's analysis requires. It may, however, be feared that Brandom is precluded from holding both that practical deontic attitudes are non-conceptually contentful and that they can be used to account for conceptual content, by his view³ that non-conceptual content is «derived» from conceptual content, which alone can be «original». This the worry I want to address in the next section.

§2.— Original, Yet Dependent Intentionality?

To deal with this worry, we should first ask what the distinction between original and derivative intentionality consists in, and especially whether the fact that a kind of intentionality or contentfulness is derivative, while some other kind is original, is supposed to entail that the former depends on the latter, in the sense that any explanation of the first would involve either the second or an explanation of the second⁴. One obvious suggestion is that something counts as derivatively intentional when its being intentional depends on something else's being intentional, and as originally intentional when its being intentional doesn't depend on anything else's being intentional. Now, this can generalize in at least two different ways:

³ A view which he shares with Davidson, and generally, with all those who hold that thought depends on language.

⁴ In other words, I submit that X depends on Y if and only if it is impossible to account for X without either invoking Y or being in a position to provide an account of Y as well.

- (1) an intentional system counts as derivatively intentional when its being intentional (its capacity to be in contentful states) depends on some other system's being intentional, and as originally intentional when its being intentional doesn't depend on any other system's being intentional
- (2) a (conceptually or non conceptually contentful) state/performance counts as derivatively contentful when its being contentful depends on another state/performance's being (either conceptually or non-conceptually) contentful, and as originally contentful when its being contentful doesn't depend on any other state/performance's being (conceptually or non-conceptually) contentful.

And none of these seems to imply that if non-conceptual content is derivative and conceptual content, original, then non-conceptual content depends on conceptual content; for none implies that if non-conceptual content is derivative and conceptual content, original, then no intentional system can have the capacity to be in non-conceptually contentful states unless it also has the capacity to be in some conceptually contentful state⁵. However, this may not be a very significant observation, given that neither (1) nor (2) seems to capture exactly what is involved in saying that non-conceptual intentionality is merely derivative (and conceptual intentionality, original).

According to (1), to claim that non-conceptual intentionality is derivative is to say that all intentional systems which have the capacity to be in non-conceptually contentful states are such that their having this capacity depends on some other system's having the capacity to be in (conceptually or non-conceptually) contentful states, and to claim that conceptual intentionality is original is to say that no intentional system which has the capacity to be in conceptually contentful states is such that its having this capacity depends on any other system's having the capacity to be in (conceptually or non-conceptually) contentful states. But this cannot possibly be what Brandom means by «derivative» and «original»; for on this understanding, it would follow from the social character of intentionality, that both conceptual intentionality and non-conceptual intentionality are derivative, and that neither is original.

According to (2), to claim that non-conceptual intentionality is derivative is to say that all non-conceptually contentful states/performances are such that their contentfulness depends on the (conceptual or non-conceptual) contentfulness of some other state/performance, and to claim that conceptual intentionality is original is to say that no conceptually contentful state/performance is such that its contentfulness depends on the (conceptual or non-conceptual) contentfulness of any other state/performance. But on this understanding, conceptual intentionality will probably turn out not to be original, since (unless content turns out to be thoroughly atomistic) it is likely that many conceptually contentful states/performances are such that their contentfulness depends on the (conceptual or non-conceptual) contentfulness of some other state/performance. Moreover, this same reading would lead one to hold that non-conceptual intentionality is derivative even if it should turn out that the contentfulness of all non-conceptually contentful states/performances depends on the *non-conceptual* contentfulness of some other state/performance (and never on the conceptual contentfulness of any state/performance).

⁵ And we have agreed above to understand the claim that non-conceptual content depends on conceptual content as implying just that.

It is becoming apparent that these suggestions fail to take account of the fact that the idea behind saying that non-conceptual intentionality is derivative, is that it is somehow derived *from conceptual* intentionality. A better way to capture what is intended would thus be to say either that

- (3) non-conceptual intentionality is derivative (with respect to conceptual intentionality) if and only if all intentional systems which have the capacity to be in non-conceptually contentful states are such that their having this capacity depends on some other system's having the capacity to be in conceptually contentful states

and

- (4) conceptual intentionality is original (with respect to non-conceptual intentionality) if and only if no intentional system which has the capacity to be in conceptually contentful states is such that its having this capacity depends on any other system's having the capacity to be in non-conceptually contentful states,

or that

- (5) non-conceptual intentionality is derivative (with respect to conceptual intentionality) if and only if all non-conceptually contentful states/performances are such that their contentfulness depends on the conceptual contentfulness of some other state/performance

and

- (6) conceptual intentionality is original (with respect to non-conceptual intentionality) if and only if no conceptually contentful state/performance is such that its contentfulness depends on the non-conceptual contentfulness of any other state/performance.

Since it has been granted above that conceptual content depends on non-conceptual content, in the sense that no system can have the capacity to be in conceptually contentful states unless it also has the capacity to be in non-conceptually contentful states, and given the social character of intentionality, it follows that conceptual intentionality isn't original either in the sense provided by (4), or in the sense provided by (6). On the contrary (on the assumptions just mentioned), conceptual intentionality turns out to be derivative with respect to non-conceptual intentionality (when «to be derivative with respect to non-conceptual intentionality» is understood in the way suggested by (3) or (5)).

As far as I can see, however, the fact that non-conceptual content doesn't depend on conceptual content wouldn't prevent it from being derivative (with respect to conceptual content) either in the sense provided by (3) or the sense provided by (5). Actually, (5) boils down to (3), on the assumption that non-conceptual content doesn't depend on conceptual content.

On these readings, then, it would turn out that (i) conceptual intentionality is derivative with respect to non-conceptual intentionality, in virtue of the fact that it depends on non-conceptual content (and that content is essentially social), and that (ii) non-conceptual intentionality *could be* derivative with respect to conceptual intentionality, even though it doesn't depend on conceptual content. This may help to make sense of the suggestion that it would not necessarily be inconsistent to hold that non-conceptual intentionality is derivative with respect to conceptual intentionality while denying that non-conceptual content depends

on conceptual content. But there still remains to make sense of the claim that conceptual intentionality is original.

This is easier than the preceding remarks might suggest. The only thing to do is to restrict exclusion of the dependence relation to intentional systems which are at once capable of being in non-conceptually contentful states but incapable of being in conceptually contentful states, thus:

- (7) conceptual intentionality is original (with respect to non-conceptual intentionality) if and only if no intentional system which has the capacity to be in conceptually contentful states is such that its having this capacity depends on the fact that any other system *lacking the capacity to be in conceptually contentful states* has the capacity to be in non-conceptually contentful states.

As far as I can see, this yields an intuitively plausible sense in which conceptual intentionality probably *is* original, and in which its being original wouldn't conflict with the fact that it is yet *dependent* on non-conceptual intentionality.

To round up this discussion, let's now ask on what grounds one could be tempted to claim that non-conceptual intentionality is derivative with respect to conceptual intentionality. It seems to me that (in Brandom's and Davidson's cases at least) the motivation for this view comes mainly from the interpretationist (or phenomenalist) principle according to which any state or performance can be contentful only insofar as (and in virtue of the fact that) it is or can be treated as such (i.e., insofar as it has been conferred some content by a set of practices), *together with* the feeling that (i) intentional systems lacking the capacity to be in conceptually contentful states cannot have the capacity to engage in any «content-conferring» practice (i.e., to practically attribute any relevant normative status), and that (ii) such systems (e.g., non-linguistic animals) therefore count as intentional (if at all) only in virtue of the fact that other systems, which *are* capable to be in conceptually contentful states, treat them as capable to be in (non-conceptually) contentful states and attribute such states to them. In other words, this attitude seems to rest on the conviction that only systems which have the capacity to be in conceptually contentful states can have the *practical* attitude of treating anything as contentful. Thus the intuition behind this talk of «original» and «derivative» intentionality (in the context of an interpretationist perspective) may well be that an intentional *system* (as opposed to a kind of intentionality, such as conceptual or non-conceptual intentionality) counts as *derivatively* intentional when it lacks the capacity to treat anything as contentful or intentional (i.e., to attribute content or intentionality to anything), and as *originally* intentional when it has this capacity. At least this seems to fit nicely with the way in which I have proposed to understand these notions, in (3) and (7) above. According to what has just been suggested, if non-conceptual intentionality were to satisfy (3), and to count as derivative with respect to conceptual intentionality, it would have to be in virtue of the fact that all systems having the capacity to be in non-conceptually contentful states (whether or not they are also capable to be in conceptually contentful states) are such that they have this capacity only in virtue of the fact that some systems which have the capacity to be in conceptually contentful states *treat them as having this capacity*⁶. And as for conceptual intentionality, it does satisfy

⁶ Which is *not* to say that all systems capable of non-conceptual intentionality would therefore count as derivatively intentional *systems*, for among these, some would be capable conceptual intentionality as well. Hence, the claim that some system is originally or derivatively intentional should not be

(7), and this is in virtue of the fact that no system having the capacity to be in conceptually contentful states is such that it has this capacity only in virtue of the fact that some system lacking the capacity to be in conceptually contentful states (but having the capacity to be in non-conceptually contentful states) treat it as having this capacity.

Now, the idea that no system is capable of treating anything, in practice, as contentful unless it has the capacity to be itself in some *conceptually* contentful states (on which the view that non-conceptual intentionality is merely derivative rests, in part) strikes me both as questionable (or at least in need of further argument) and as little more than an empirical conjecture. It should at least be pointed out that Brandom's program itself requires that it be possible to practically attribute some (non-discursive) normative statuses without having the capacity to be in any conceptually contentful state. And if this is right, then it is hard to see why only systems capable of conceptual thought could have the capacity to practically attribute (non-conceptually) contentful states.

Thus, the claim that non-conceptual intentionality is merely derivative (with respect to conceptual intentionality) seems insufficiently supported, even from within Brandom's perspective. But what is more important, in the present context, is (i) that rejecting it would not involve renouncing all forms of interpretationism, and (ii) that it is in any case perfectly compatible with the claim that non-conceptual content does not depend on conceptual content. For the latter is the claim that a system could have the capacity to be in non-conceptually contentful states without having the capacity to be in any conceptually contentful state, while the claim that non-conceptual content is derivative amounts to the claim that no system can have the capacity to be in non-conceptually contentful states unless some *other* system has the capacity to be in conceptually contentful states. It follows that one could explain conceptual content in terms of non-conceptual content even if the latter were derivative with respect to the former. It would not be exaggerated to suggest that if interpretationism doesn't require that non-conceptual intentionality be derivative, then its opponents are thereby deprived of one of their strongest arguments.

One potential problem with the suggestion that Brandom's practical deontic attitudes may have non-conceptual content may however come from the fact that, in case non-conceptual content could be accounted for in non-normative terms, one would run the risk of being forced to conclude that it is, after all, in non-normative, rather than in normative terms that content has been accounted for. In other words, Brandom's normativism would then threaten to change into its opposite (but wouldn't that be very *hegelian*?).

There is a sense in which any explanation presupposes conceptual intentionality, just in virtue of the fact that to explain anything involves a piece of discursive thinking. But this obviously isn't enough to make it objectionably circular to appeal to non-conceptual intentionality in accounting for conceptual intentionality (and since only systems who enjoy conceptual intentionality could set themselves the task of accounting for it, such circularity is in any case unavoidable). The main thing is that to attribute non-conceptual intentionality is not yet to *attribute* conceptual intentionality, and that it is conceivable that some systems be able to have non-conceptually contentful attitudes without being able to have conceptually contentful attitudes, or perhaps even that some systems might be able to have practical deontic

confused with the claim that conceptual (non-conceptual) intentionality is original (derivative) with respect to non-conceptual (conceptual) intentionality.

attitudes (conceived as a special kind of non-conceptually contentful attitudes) without their capacity to have such attitudes being sufficient to make them capable of being in conceptually contentful states (e.g. because their use of this capacity doesn't exhibit the right kind of structure).

§3.— Original Intentionality in Practice and in Theory

My aim in this section is to see how and to what extent Brandom's brand of interpretationism departs from Davidson's view that one can be a thinker only if one has the concept of a concept.

On *one* natural reading of Davidson's views, he holds that a necessary condition of having the capacity to be in any intentional state is having the capacity to attribute such states to others, where to attribute an intentional state involves (not just having some practical attitude, as suggested by Brandom, but) mastering the concept of a contentful, intentional state. In other words, to attribute an intentional state, on this understanding, is to be in some (higher-order) conceptually contentful intentional state whose content can be specified only by using indirect discourse (and thus amounts to thinking or judging that someone *thinks that* so-and-so). This might be encapsulated in the following «principle of intentional ascent»:

- (8) If S has the capacity to think that p, then S has the capacity to think that S' thinks that p.

Insofar as Davidson's interpretationist perspective implies that a system is derivatively intentional only if it has the capacity to be in contentful states (i.e. to have «thoughts», in some generic sense) without itself being able to attribute such states to others, this further claim obviously doesn't leave any room for derivative intentionality, which could at best be construed as unreal and metaphorical. In this perspective, the capacity to have thoughts (or to be in contentful states) at all thus goes hand in hand with the capacity to have higher-order conceptually contentful thoughts.

But in view of the fact that it is certainly possible to think *that p* without thinking *that anyone thinks that p*, it seems that it should also be possible to attribute *the thought that p* without attributing *the thought that anyone thinks that p*. And if this is possible, then there doesn't seem to be any reason to deny that it is possible to attribute the thought that p without attributing the capacity to have the thought that anyone thinks that p. In other words, it is hard to believe that there could be systems with the capacity to attribute the capacity to attribute thoughts (in the present context, the capacity to have higher-order conceptual thoughts) without the capacity to attribute thoughts *simpliciter* (i.e., first-order thoughts) in the first place. But if these really are two distinct capacities, there is no reason why such systems should not also be able to attribute thoughts to systems to which they don't attribute the capacity to attribute thoughts themselves. This obviously doesn't show (8) to be false, but is enough to show that it's not necessary (and cannot be sustained as any kind of conceptual truth, as Davidson would have it), and strongly suggests that if some systems have original intentionality (which, in this context, means the capacity to attribute intentional states), then it must be *possible* that some systems have only derivative intentionality (which, in this context, means the capacity to be only in simple, first-order, intentional states). Hence, there seems to be a sense in which original intentionality depends on the *possibility* of derivative intentionality; that is to say, there would be no originally intentional systems if they didn't have the capacity to attribute derivative intentionality (i.e., if they couldn't treat some systems as having only «simple», derivative, intentionality).

The significance of this conclusion is not only that it *may* go some way towards doing justice to the common sense idea that non-linguistic animals might enjoy genuine intentionality even if it should turn out that they lack any capacity to attribute any intentional states, but that it may provide some ground for holding that the notion of the objectivity of conceptual norms must make sense, or at least help to make that notion intelligible. Since a derivatively intentional system is a system which can be in contentful states, but lacks the capacity to think that it (or anything) is in such states, the contentful states of such a system may turn out to be incorrect or inappropriate without its having the capacity to recognise that this is so. To admit derivative intentionality is thus to admit that some intentional states may be incorrect without the bearer of these states recognising that they are. But if it makes sense to admit this possibility, it seems it must also make sense to admit that one's originally intentional states may be incorrect without one's realising that they are, and that if this can happen, then it can also happen that no one takes (or will ever take) them to be incorrect.

I conclude that the principle of intentional ascent, as stated by (8), is untenable. But as will shortly become apparent, this is not to say that other, restricted forms of this principle may not be acceptable.

Recall that one main difference between Davidson's perspective (as just reconstructively modified) and Brandom's, is that the latter allows that to attribute an intentional state (most fundamentally, a doxastic commitment) may consist, not in being in any conceptually contentful state, but in having some practical deontic attitude. In other words, while Davidson takes all attributions of intentionality as being themselves conceptually contentful attitudes, Brandom introduces a distinction between attributions of intentionality which are somehow «implicit in practice» and consist in taking (certain kinds of) practical deontic attitudes, and attributions which are explicit in thought or discourse (and involve being in higher-order intentional states, in the form of having higher-order *discursive* commitments). This (as will soon appear more clearly) seems to be what permits him to draw a line separating two kinds, or grades, of original conceptual intentionality; that is to say, to allow that some systems may have only the practical ability to attribute (conceptually) contentful states, while others have both this practical ability and the capacity to make such attributions explicit (and thus not only have concepts but have specifically mastered the concept of a conceptually contentful state).

As already remarked, Brandom's project of giving an account of conceptual content (and discursive deontic attitudes) in terms of practical deontic attitudes precludes one to take the latter to be themselves conceptually contentful attitudes, and thus to assume that one who practically attributes some doxastic commitment is thereby doxastically committed to anything. I suggested above that taking these practical attitudes to be non-conceptually contentful (instead of non-contentful at all) would neither conflict with the view that non-conceptual intentionality is derivative, nor be objectionably circular, if it were allowed that a system might have such attitudes without yet being able to be in or to attribute conceptually

contentful states at all⁷. But this suggestion still raises many other questions, some of which concern the status of various forms of intentional ascent which are now made available.

As can now be clearly seen, the principle of intentional ascent, as originally given by (8), can be read in *at least* eight different ways, depending on (i) whether various occurrences are taken to refer to conceptual or non-conceptual thoughts, and (ii) whether attributions of intentional states are taken to involve being in some conceptually contentful state, or being in some non-conceptually contentful state, i.e., depending on whether we're talking about «theoretical» (and explicit) or «practical» (and implicit) attributions. This accordingly yields *four* principles of *theoretical* intentional ascent :

- (9) If S can conceptually think that p, then S can conceptually think that S' conceptually thinks that p (i.e., S can theoretically attribute the conceptual thought that p),
- (10) If S can conceptually think that p, then S can conceptually think that S' non-conceptually thinks that p (i.e., S can theoretically attribute the non-conceptual thought that p),
- (11) If S can non-conceptually think that p, then S can conceptually think that S' conceptually thinks that p (i.e., S can theoretically attribute the conceptual thought that p),
- (12) If S can non-conceptually think that p, then S can conceptually think that S' non-conceptually thinks that p (i.e., S can theoretically attribute the non-conceptual thought that p),

and *four* corresponding principles of *practical* intentional ascent :

- (13) If S can conceptually think that p, then S can non-conceptually think that S' conceptually thinks that p (i.e., S can practically attribute the conceptual thought that p),
- (14) If S can conceptually think that p, then S can non-conceptually think that S' non-conceptually thinks that p (i.e., S can practically attribute the non-conceptual thought that p),
- (15) If S can non-conceptually think that p, then S can non-conceptually think that S' conceptually thinks that p (i.e., S can practically attribute the conceptual thought that p),
- (16) If S can non-conceptually think that p, then S can non-conceptually think that S' non-conceptually thinks that p (i.e., S can practically attribute the non-conceptual thought that p).

⁷ But it must be admitted that Brandom doesn't give any clear indication that he would embrace this position, and gives instead the impression (see, e.g. 1994 : xiii) that the practical deontic attitudes that are supposed to institute deontic statuses are to be taken as non-intentional at all, which is the source of some puzzlement. On either view, however, it could be granted that not any practical deontic attitude is to count as a practical attribution of a conceptually contentful state, or even of a contentful state in general, for not all practical deontic attitudes are part of practices exhibiting the right kind of structure and complexity. Brandom's account would therefore be immune to at least one kind of circularity.

Just as the original principle (8) amounts to the claim that all intentional systems are originally intentional, these various principles claim either that conceptually intentional systems or non-conceptually intentional systems are originally intentional. But since an intentional system may count as originally intentional either in virtue of the fact that it can make practical attributions of intentionality, or in virtue of the fact that it can make theoretical attributions of intentionality, and since a distinction must be made between attributions of conceptual intentionality and attributions of non-conceptual intentionality, we now have four different ways in which an intentional system may turn out to be originally intentional.

Now, taking account of the fact that (i) to conceptually (non-conceptually) think *that S' conceptually (non-conceptually) thinks that p* is an instance of conceptually (non-conceptually) thinking *that p*, and that (ii) to conceptually (non-conceptually) think that S' (conceptually or non-conceptually) thinks that p is to theoretically (practically) attribute the (conceptual or non-conceptual) thought that p, it will be noted that each of (9)-(16) yields a corresponding higher-order principle as a special case :

- (9*) If S can theoretically attribute the (conceptual or non-conceptual) thought that p, then S can theoretically attribute the attitude of theoretically attributing the (conceptual or non-conceptual) thought that p,
- (10*) If S can theoretically attribute the (conceptual or non-conceptual) thought that p, then S can theoretically attribute the attitude of practically attributing the (conceptual or non-conceptual) thought that p,
- (11*) If S can practically attribute the (conceptual or non-conceptual) thought that p, then S can theoretically attribute the attitude of theoretically attributing the (conceptual or non-conceptual) thought that p,
- (12*) If S can practically attribute the (conceptual or non-conceptual) thought that p, then S can theoretically attribute the attitude of practically attributing the (conceptual or non-conceptual) thought that p,
- (13*) If S can theoretically attribute the (conceptual or non-conceptual) thought that p, then S can practically attribute the attitude of theoretically attributing the (conceptual or non-conceptual) thought that p,
- (14*) If S can theoretically attribute the (conceptual or non-conceptual) thought that p, then S can practically attribute the attitude of practically attributing the (conceptual or non-conceptual) thought that p,
- (15*) If S can practically attribute the (conceptual or non-conceptual) thought that p, then S can practically attribute the attitude of theoretically attributing the (conceptual or non-conceptual) thought that p,
- (16*) If S can practically attribute the (conceptual or non-conceptual) thought that p, then S can practically attribute the attitude of practically attributing the (conceptual or non-conceptual) thought that p.

It goes without saying that any of these higher-order principles could be true even if the corresponding first-order principle had to be rejected.

These various «combinatorial» possibilities are of course not on a par, and the purpose of making them explicit is to help to clarify Brandom's view and how it is related to

Davidson's. Hence, the task before us is to determine, for each of (9)-(16), how it (and its higher-order corollary) fares with respect to Davidson's and Brandom's perspectives (on the working assumption that practical deontic attitudes are non-conceptually contentful). This fortunately turns out to be less painstaking than it might seem.

In the present larger setting, it can easily be seen that (9) is the only reading of the original principle of intentional ascent (the one given by (8)) which accords with Davidson's claim that all intentionality is conceptual (and original) intentionality, and which provides a plausible way to understand his well-known claim that no one can be interpretable unless one is an interpreter (i.e., unless one has the concept of a thought, and is a *theoretical* attributor of intentionality). By contrast, Brandom clearly rejects not only this principle⁸, but all of (10)-(12) as well (that is to say, all forms of first-order theoretical intentional ascent), since he holds that it is possible to have conceptual (or for that matter, non-conceptual) thoughts without having the capacity to make any explicit, «theoretical» attribution of intentionality⁹.

It should however be emphasized that even though he rejects (9), Brandom could (and probably would) accept its higher-order corollary (9*), thus endorsing the claim that no one can have the capacity to theoretically attribute thoughts unless one has the capacity to theoretically attribute this very capacity (which is *one* version of the Davidsonian claim that no one can be an interpreter unless one has the concept of an interpreter). As far as I can see, (10*) likewise seems compatible with everything Brandom says, though it remains unclear to what extent he would be prepared to endorse it (and how plausible it really is). And as for (11*)-(12*), they clearly are excluded by Brandom's claim that one may have the capacity to practically attribute intentional states without yet having the resources to make explicit, theoretical attributions (i.e., by Brandom's distinction between «merely rational» and «fully logical» conceptually intentional systems).

Let's now turn to the «practical» forms of intentional ascent, (13)-(16) and (13*)-(16*).

Clearly, since Brandom wishes to maintain that all conceptually intentional systems are originally intentional, while allowing that some of them may yet lack the capacity for theoretical attributions of intentional states, this leaves him no choice but to endorse (13) (and (13*)) with it). This actually provides his version of Davidson's claim that no one can be interpretable without being an interpreter, one in which being interpretable is restricted to being «conceptually» interpretable, and being an interpreter is restricted to being a «practical» interpreter.

Furthermore, the argument I gave above, to the conclusion that no one could have the capacity to attribute intentional states at all unless one has the capacity to attribute derivatively intentional states, strongly suggests that (on the assumption that derivatively intentional systems, if any, must be capable of non-conceptual intentionality) (14) should also be acceptable, for what it says, in effect, is that no one could have conceptual thoughts without having the capacity to practically attribute non-conceptual thoughts, and it has already been granted that no one could have conceptual thoughts without having the capacity to practically attribute at least conceptual thoughts. But as with (10*), it is hard to tell to what extent

⁸ Even though it is perhaps less clearly untenable than (8).

⁹ It should also be pointed out that (11)-(12) are incompatible with the assumption (made in section 6) that non-conceptual intentionality doesn't depend on conceptual intentionality.

Brandom would be prepared to endorse (14), even though it seems compatible with his views. It all depends on whether the capacity to practically attribute conceptual thoughts (which all conceptually intentional systems must have, according to Brandom) should be taken to imply the capacity to practically attribute non-conceptual thoughts.

Insofar as Brandom is committed to accept that intentional systems which have the capacity to be in non-conceptually contentful states need not have the capacity to attribute (even practically) any intentional state, he clearly must reject both (15) and (16). Moreover, it would seem that if (as Brandom claims) one may be able to practically attribute thoughts without being able to theoretically attribute them, then (*a fortiori*) one may be able to practically attribute thoughts without being able to practically attribute theoretical attributions of thoughts; and hence, that (15*) probably must be rejected as well. The same doesn't hold for (16*), however, which is compatible with Brandom's views.

The upshot of this discussion is that Brandom is clearly committed only to (13)-(13*), though nothing (so far) seems to preclude his acceptance of (9*), (10*), (14), (14*) and/or (16*). Note that only two of these (namely, (9*) and (10*)) are forms of *theoretical* intentional ascent, and then only *higher-order* ones. Since in the context of Brandom's attempt to explain conceptual intentionality, it clearly is *practical* intentional ascent that is of prime importance, they can safely be ignored here. But perhaps it is worth giving a closer look at (14)-(14*) and (16*), despite the fact that they don't seem to be strictly required by Brandom's explanation of conceptual intentionality (and that Brandom doesn't pay much attention to attributions of non-conceptual intentionality).

On the face of it, (16*) looks interestingly less compelling than its (Davidsonian) direct opposite (9*). Indeed, since practical attributions of thoughts (even conceptual ones) are not (and cannot be) *conceptual* thoughts, there is no obvious reason why the capacity to practically attribute (even conceptual) thoughts should entail the capacity to practically attribute *practical attributions* of thoughts.

Furthermore, given that in Brandom's terminology, the following special case of (16*):

(17) If S can practically attribute the conceptual thought that p, then S can practically attribute the attitude of practically attributing the conceptual thought that p,

would translate as:

(18) If S can have the practical deontic attitude of taking S' to be doxastically committed to p, then S can have the practical deontic attitude of taking S' to have the practical deontic attitude of taking S'' to be doxastically committed to p,

to endorse (16*) would imply that to be involved at all in discursive practice (i.e., to have conceptual thoughts) requires that the participants have the capacity, not only to keep deontic scores on each other, but also to keep scores on each other's scores, etc... But this raises a problem, insofar as to keep score on the score kept by S must be *something different* from keeping score on S's (higher-order) discursive deontic statuses (since by hypothesis, practically attributing a doxastic commitment is not having oneself any *doxastic* commitment). Perhaps this point could be dealt with by taking practical deontic attitudes to be commitments of a *non-discursive* kind (taking in practice a performance as correct might *commit*, and not only dispose, one to sanction it), a course which, as far as I can see, is not precluded by the non-

circularity condition (and accords with some of Brandom's remarks¹⁰). However, it would still be hard to see how such higher-order attitudes could nevertheless be purely *practical* attitudes, if this is meant to imply that they must be such that those who have them may yet lack the capacity to have corresponding doxastic commitments (as Brandom seems to require).

Another potential problem with (16*) stems from the fact that the following also is a special case of it:

- (19) If S can practically attribute the non-conceptual thought that p, then S can practically attribute the attitude of practically attributing the non-conceptual thought that p,

and that to endorse it would thus commit one to the view that no system could have original non-conceptual intentionality without having the capacity to treat, in practice, other systems as themselves enjoying *original* non-conceptual intentionality (i.e., as practical attributors of non-conceptual intentionality); a capacity which one would perhaps not want to grant to all systems which are capable of practically attributing non-conceptual thoughts without being capable of having conceptual thoughts (if there are such systems). Of course, this would not be much of a problem for Brandom, since he denies that there could be any such intentional systems.

But even if (16*) has to be rejected (which is by no means clear), it might still be the case that all intentional systems of some interestingly restricted class are required to have the (higher-order) capacity to practically attribute practical attributions of thoughts. Indeed, it is easily seen that, on the plausible assumption that if no one can have the capacity to practically attribute the attitude of theoretically attributing the thought that p unless one also has the capacity to practically attribute the attitude of practically attributing the thought that p, (13*) (and hence (13)) entails something between (13*) and (16*), namely:

- (14*) If S can theoretically attribute the (conceptual or non-conceptual) thought that p, then S can practically attribute the attitude of practically attributing the (conceptual or non-conceptual) thought that p,

(in Brandom's terminology: if S can be doxastically committed to someone's having the thought that p, then S can have the practical deontic attitude of taking someone to practically attribute the thought that p).

It would be highly interesting to be able to argue as well that *only* intentional systems of some interestingly restricted class can have this (higher-order) capacity to practically attribute practical attributions of thoughts, e.g., to establish the claim that:

- (20) If S can practically attribute the attitude of practically attributing the thought that p, then S can theoretically attribute the thought that p,

which would restrict higher-order practical attributions to what Brandom calls «fully logical» intentional systems; or perhaps only the weaker claim that:

- (21) If S can practically attribute the attitude of practically attributing the thought that p, then S can have the *conceptual* thought that p,

¹⁰ However, this would require committing oneself to the claim that even non-conceptual content is essentially normative; and this would need some justification.

which would restrict them to conceptually (but not necessarily «fully logical») intentional systems. In either case, it would have to be granted (as expected) that some practical attitudes (namely, higher-order ones) are such that only conceptually intentional systems can have them. It has however to be observed that the weaker claim (21) would seem to be of special significance only if it could also be shown that:

(22) If S can have the conceptual thought that p, then S can practically attribute the attitude of practically attributing the thought that p,

something which doesn't follow from any of the forms of intentional ascent which have so far been considered, but to which Brandom may well be committed.

Enough has been said, I think, to demonstrate that the hypothesis that Brandom's practical deontic attitudes are to be taken as non-conceptually contentful attitudes, though *prima facie* coherent and appealing, raises a number of intriguing and potentially fruitful questions. Exploring them further would however require something which I don't yet have, namely, a substantial theory of non-conceptual content.¹¹

References

- Brandl, Johannes and Wolfgang L. Gombocz eds. (1989) *The Mind of Davidson*, Amsterdam, Rodopi
- Brandom, Robert B. (1985) «Varieties of Understanding», Rescher ed. (1985) 27-51
- Brandom, Robert B. (1994) *Making It Explicit*, Cambridge (Mass.), Harvard U. Press
- Brandom, Robert B. (1997a) «Précis of *Making It Explicit*», *Phil. and Phenom. Research* 57, 153-156
- Brandom, Robert B. (1997b) «Replies», *Phil. and Phenom. Research* 57, 189-204
- Brandom, Robert B. (2000) *Articulating Reasons*, Cambridge (Mass.), Harvard U. Press
- Child, William (1994) *Causality, Interpretation and the Mind*, Oxford, Oxford U. Press
- Coates, Paul and Daniel Hutto eds. (1996) *Current Issues in Idealism*, Bristol, Thoemmes Press
- Davidson, Donald (1980) *Essays on Actions and Events*, Oxford, Oxford U. Press
- Davidson, Donald (1984) *Inquiries Into Truth and Interpretation*, Oxford, Oxford U. Press
- Davidson, Donald (1989a) «The Conditions of Thought», Brandl and Gombocz eds. (1989) 193-200
- Davidson, Donald (1989b) «What is Present to the Mind?», Brandl and Gombocz eds. (1989) 3-18
- Davidson, Donald (1991) «Three Varieties of Knowledge», Griffiths ed. (1991) 153-166
- Davidson, Donald (1992) «The Second Person», French *et al.* eds. (1992) 255-267

¹¹ Part of this paper has been read at the annual meeting of the CPA, at the Université Laval, on May 26th 2001. I am grateful to Don Ross for his stimulating comments.

- Davidson, Donald (1996) «Subjective, Intersubjective, Objective», Coates and Hutto eds. (1996) 155-177
- French, P. A., T. E. Uehling jr and H. K. Wettstein eds. (1992) *Midwest Studies in Philosophy 17: The Wittgenstein Legacy*, Notre Dame, U. of Notre Dame Press
- Gampel, E. H. (1997) «The Normativity of Meaning», *Phil. Studies* 86, 221-242
- Gibbard, Allan (1994) «Meaning and Normativity», Villanueva ed. (1994) 95-116
- Gibbard, Allan (1996) «Thought, Norms, and Discursive Practice: Commentary on Robert Brandom, *Making It Explicit*», *Phil. and Phenom. Research* 56,699-717
- Griffiths, A. P. ed. (1991) *A. J. Ayer: Memorial Essays*, Cambridge, Cambridge U. Press
- Kripke, Saul (1982) *Wittgenstein on Rules and Private Language*, Oxford, Blackwell
- McDowell, John (1997) «Brandom on Representation and Inference», *Phil. and Phenom. Research* 57, 157-162
- Peacocke, Christopher (1992) «Sense and Justification», *Mind* 101, 793-816
- Rescher, Nicholas ed. (1985) *Reason and Rationality in Natural Science*, Lanham, U. Press of America
- Rosen, Gideon (1997) «Who Makes the Rules Around Here?», *Phil. and Phenom. Research* 57, 163-171
- Rosenberg, Jay F. (1997) «Brandom's *Making It Explicit*: A First Encounter», *Phil. and Phenom. Research* 57, 179-187
- Villanueva, Enrique ed. (1994) *Philosophical Issues 5: Truth and Rationality*, Atascadero, Ridgeview

Daniel Laurier

Département de Philosophie. Université de Montréal

C.P. 6128, Succ. Centre-ville

Montréal, Québec H3C 3J7, Canada

<Daniel.Laurier@umontreal.ca>

SORITES (ΣΩΡΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 23-32

Metaphor and Meaning

Copyright © by SORITES and Alec Hyslop

METAPHOR AND MEANING

Alec Hyslop

How do we manage to understand metaphors? Do we understand them? What are they? What kind of meaning do they have? Do they have any? If they do, why are they not paraphrasable?

I begin with what I am sure about. Metaphors are, indeed, for practical purposes, elliptical similes: though not all metaphors can be given the surface form of a simile, they all involve comparisons. Metaphors, of course, are not reversible, but neither are similes. Our understanding of a metaphor only begins when we realise that the literal sense is not the point (or the only point) of the utterance. The same is true of similes. In each case we have to make what we can of the utterance. How we do this is the same for both figures. What we make is the same in each case, and subject (even if not subjected) to the same controversies. Both, as Robert Fogelin has put it, express figurative comparisons: similes explicitly, metaphors implicitly.

How do we unravel metaphors? If rumour is claimed to be (like) a disease then we compare the two. However, we proceed by looking at disease and seeing what salient features might be attachable to rumour. This order is crucial. We are interested in what features are believed to be attached to disease whether or not they are. Context might determine attachability, particularly in poetry. We also need to move from talk of objects to talk of words; not because rumour and disease are funny objects, but because we work with the words. As we go about (and about) construing the metaphor, everything that can be done in the object mode can be done in the verbal mode, but things can be done in the verbal mode that cannot be done in the object mode. Words refer so we retain our objects, odd or otherwise, but words have their own features, particularly their allusive power. The Church of England as the Tory Party at prayer would be needed to underpin our responses to the claim that Marxism-Leninism was the Communist Party of Australia at prayer.

But what do we emerge with from this exercise in comparison? What does a metaphor give us? Do we emerge with a metaphorical meaning, or two, or more? What kind of meaning would a metaphorical meaning be? If we come out with no more meaning than the literal meaning we went in with, do we come out with a new way of seeing, or experiencing? That metaphors are comparisons does not give us answers to such questions. Someone says «Rumour is a disease». Salient features of disease are contagion and damage and these features seem «attachable» to rumour. So has it been claimed that rumour is a disease and, more particularly, a social disease and harmful (even fatal, potentially)? Or have we been invited to try thinking of rumour as a disease; or to have the experience of contemplating rumour as contagious and harmful; or to entertain that thought; or, to have that thought entertain us?

Knowing how we manage to respond appropriately to this metaphor, insofar as we do know this, does not help us to decide between these options.

The fundamental disagreement about metaphor is between those who think a metaphorical utterance has a meaning other than, or in addition to, its literal meaning, and those who do not (most famously Donald Davidson). But there are important, though lesser disagreements. Are metaphors a case of speaker's meaning? Or do the words used metaphorically have, on that metaphorical occasion, a meaning other than their normal, literal sense? Does such a meaning attach to the metaphorical utterance as a whole or to a word or words in that sequence of words? Is the literal meaning retained or discarded?

I have not talked of special metaphoric meaning. That seems to offer no addition to the embarrassment of choice already on offer. However, elucidation of the relevant concepts is needed to make sure the options are clear.

'Inspissated' means 'complexly dense' (near enough) and started life as a term of haute cuisine, so that a traditional French sauce is inspissated. Naturally enough it expanded its reach into the haute cuisine of the intellectual kitchen, so that «This book is inspissated» means that the book in question is complexly dense. What if, unsurprisingly, some tyro, hearing 'inspissated' used, thinks it means 'unduly complicated'. They say: «This book is inspissated». They mean, by saying this, that it is unduly complicated. Here is a case of speaker's meaning diverging from what is normally (standardly) meant. Of course, speaker's meaning need not diverge.

What does the word 'inspissated' mean, as used on this occasion? It means 'unduly complicated'. What if our neophyte, fresh from a lecture which has been only too easily comprehensible, says: «That was inspissated». What has been said, strictly, is that the paper is complexly dense. What is meant by the speaker is that the paper was unduly simple. What 'inspissated' (the word) means on this occasion is 'unduly complicated'. So in this case the speaker's meaning diverges from the standard meaning, and from the occasion meaning, and the occasion meaning diverges from standard meaning. Of course, we need have no divergence at all between the three, and, of course, the terminology is negotiable. Notice that occasion meaning attaches to a word or words: what the word or words mean on that occasion. Speaker's meaning is a matter of what is meant by uttering those words on that occasion. Standard meaning covers both what the words normally mean and what would normally be meant by uttering them.

We are not done backgrounding. Take an ambiguous utterance. Standard meaning provides more than one meaning. Normally, one of those will be what the word means on that occasion, as used by the speaker. Now take the role that context can play in modifying standard meaning. In Mark Antony's speech in *Julius Caesar*, when he says, frequently, «Brutus is an honourable man», what 'honourable' means does not change as it is repeated. What he means though, by the utterance, does change. But the context makes clear how the speech is to be understood, independently of any other evidence, with in this case, the immediate linguistic context, the repetition of the word 'honourable', being crucial. The context might not make this clear, however. Literary texts in particular are likely to provide several possible contextual meanings as candidates for the author's meaning.

Let me make sure that contextual meaning is understood, as it occurs in ordinary (non-literary) contexts. Take the case of an ambiguous sentence, S, with two possible standard meanings, A and B. So if S is uttered the speaker might mean A or might mean B. The

speaker might also mean both (or neither, as in the case of irony) or, indeed, whatever the speaker happens to mean. So the occasion meaning could be A, or could be B, or could be A and B, or, whatever. But a choice has to be made between the various possibilities. The occasion meaning will be just one of these possibilities; viz. A, or B, or A and B, or, whatever. Speaker's meaning allows for various possibilities in terms of what might be meant by uttering S, but, as in the case of occasion meaning, must choose between them. Since context, linguistic and non-linguistic, both adds and removes possibilities, contextual meaning offers comparable possibilities for consideration. However, contextual meaning does not involve a choice being made between the possibilities. All are equally instances of contextual meaning, of what might be meant by uttering that (ambiguous) sentence, given the context. Not all, of course, need be equally likely to have been what was in fact meant; and what was in fact meant might not be predictable from the context, being totally idiosyncratic, and so not a case of contextual meaning.

In the case of a literary text, some would opt for speaker's meaning (provided it fitted the text) as giving the correct contextual meaning (interpretation); some the contextual meaning thought to be the most aesthetically satisfying; some would think it wrong to make any choice. But all the interpretations would have equal standing as contextual meanings of that text, provided each fits that text. Rejected interpretations remain contextual meanings of that text, of what might have been meant by that text, given the context, linguistic and non-linguistic.

So we now have the possibility of a four way divergence, though it remains the case that we need not have any divergence. It is important to see that occasion meaning and contextual meaning can diverge, both from one another, and from speaker's meaning (and, of course, from standard meaning). A specific, non-linguistic context might limit, or add to, what might be meant by and within a text, given what the words mean or might mean; and, of course, speaker's meaning might be gloriously or ingloriously private, indeed, idiosyncratic. Notice that it is only occasion meaning that can attach a divergent sense to a word or words. Contextual meaning is at one with speaker's meaning in generating, not word meaning, but what might be meant by an «utterance», conceived as encompassing a text, or stretch of text. Though contextual meaning is distinct from speaker's meaning, context includes the speaker, and one (or more) of the contextual meanings might be what the speaker meant in particular cases.

We need these four concepts of meaning for the ordinary, non-metaphorical range of linguistic facts. They offer enough options for metaphorical meaning, though contextual meaning does not seem to be among the options on offer in the writings on metaphor.

Sufficiently backgrounded, we can now return to the various disagreements. There are those who believe metaphorical utterances have non-literal meaning ranged against those who believe that metaphors have (at most) literal meaning. This is the big disagreement. The former group divides between those who are for speaker's meaning and those who are for occasion meaning (word meaning). There seems to be virtually universal agreement, however, that metaphors are not paraphrasable (at any rate, interesting metaphors are not).

But why are they not paraphrasable? Davidson¹ is triumphantly clear: given the absence of any meaning other than the literal there is nothing to paraphrase! With one bound Jack is free. Whereas those espousing some form of metaphorical meaning seem to struggle to have any answer, he has the knock-down answer to why metaphorical utterances are not paraphrasable.

However, Davidson would seem to have what many would see as a knock-down problem. If we lose meaning, do we not lose truth? If we lose truth, do we not lose metaphor? Surely metaphors can be true, even interesting metaphors. Where there is meaning there can be truth. Devoid of meaning, Davidson is devoid of truth.

To decide for or against Davidson I need first to consider the choice confronting those attracted to metaphorical meaning. This has heretofore been that between speaker's meaning and occasion meaning, between Searle most famously on the one hand, and Beardsley and Black on the other. Various objections have been advanced against metaphorical meaning as speaker's meaning, but they can be exemplified by Beardsley's objections.

He claims that speaker's meaning (as utterer's meaning 'cannot account for our ability to interpret metaphorical expressions even when we know that the relevant properties were not meant by any utterer.'² In addition, he claims that occasion meaning (my terminology) fits better with the move from live metaphor to dead metaphor; and with the fact that there is a continuity between metaphorical «senses» and literal senses; and with the fact that metaphorical «senses «behave in many of the same ways as literal senses. For example ... we can devise metaphorical equivocations» (p.11). He also insists on the «rule-guided character of literary interpretation» (p.11).

What these objections all trade on is surely the essential freedom of speaker's meaning, indeed its possible total arbitrariness. This freedom is, however, incompatible with metaphorical utterance. You cannot decree that your utterance be a metaphor by fiat. Your utterance can, indeed, be totally idiosyncratic, and risk, or court, being incomprehensible, but it cannot then be a metaphor. No more than a raised eyebrow can a metaphor exist by itself. It needs a linguistic context that will sustain a metaphorical interpretation. In this it is different from ironic utterance, where my insistence that I am speaking ironically may be treated with incredulity by all and sundry, but no matter; I can insist, and it is possible that I am.

So speaker's meaning is ruled out for metaphors. It is too wide. Of course, speaker's meaning may be the same as metaphorical «meaning» in a particular case, but a metaphorical «meaning» may force itself on us in another particular case, whether or not such a «meaning» was, or could have been, the speaker's. The words themselves, sometimes perhaps with a little help from the non-linguistic context, invite a metaphorical interpretation. So speaker's meaning is also too narrow. But if metaphorical «meaning» is seen as contextual meaning, and generated by the context, linguistic and non-linguistic, then it is not speaker's meaning, it is not loose, and it is not vulnerable to Beardsley's objections.

¹ Donald Davidson, 'What Metaphors Mean', *Critical Inquiry* 5 (1978): 31-47; reprinted in Sheldon Sacks (ed.) *On Metaphor* (Chicago: University of Chicago Press, 1979). 29-45, especially 44-45. Page references are to Sacks.

² Monroe Beardsley, 'Metaphorical senses', *Noûs*, 12 (1978), 3-16.

It is true that metaphorical utterance has the potential to be highly specific, depending on the features of that particular occasion, its circumstances, and participants. Such a moment for metaphor can, indeed, pass. But the dependence on the words remains. Even so, the outsider may well be at a loss, not knowing the particular circumstances. Contextual meaning still covers such a metaphorical utterance because of its dependence on the words, on the linguistic context.

If metaphorical meaning is not speaker's meaning, is it a case of occasion meaning; a matter of what the word, or words mean as used on that (metaphorical) occasion?

Surely attaching such new senses would be *ex post facto*? Faced with a text we make what sense of it we can, as a text, as a whole. What might be meant by this text, by this or that passage, by particular words within the text? What might be meant by particular words, given all the other words, and whatever else is germane? We work with, and within, the context. We need answers at this level before there could be any chance of assigning metaphorical meanings to individual words, and assigning such meanings is what the alternative to contextual meaning demands.

Also, we are not always able to assign meaning to individual words. Often metaphor works with phrases as the relevant unit and it seems strained to attach a new meaning to the phrase where we are unable to do so for the individual words making up the phrase. Such cases are handled straightforwardly within contextual meaning: as what is meant, or might be meant, by uttering the phrase.

A fortiori, if we think we should work on the level of sentences, then contextual meaning seems the natural way to go, and I think we should accept that we work no lower than on that level. Take George Herbert's two lines from his poem, *Virtue*: «Only a sweet and vertuous soul, like seasoned timbre, never gives». Try attaching new senses to some of these words, as opposed to the whole quotation. The sounds of words, the rhythms, allusions, (both internal and external) all go to require at least the sentence as the minimal unit to which a new meaning or meanings could be attached.

If that is accepted then what would be wrong with attaching a new sense to George Herbert's lines? What would be wrong is that this case of sentence meaning is unlike standard sentence meaning, and unlike divergent sentence meaning as it otherwise occurs. Here the sentence as a whole acquires a new sense or senses but not the individual words. Whatever understanding we have of this kind of sentence meaning seems parasitic on our understanding that what is meant by uttering a sentence may be different from what is normally meant. Insofar as it is thought to be a distinct notion, it is not surprising that this has led to notions of a special kind of meaning, and then to a special kind of (metaphorical) truth. Contextual meaning carries no such danger, being a particular case of something's being what is meant, or might be meant, by uttering (in this case) a sentence, where speaker's meaning is another such particular case. There is, therefore, nothing mysterious about either of these cases.

So contextual meaning should be the choice for metaphorical meaning. However, Searle's choice, speaker's meaning, is in fact proffered as possible speaker's meaning.³ So perhaps contextual meaning is nothing more than that: possible speaker's meaning? Not so: possible

³ John R. Searle, 'Metaphor', in Andrew Ortony (ed.) *Metaphor and Thought* (Cambridge: Cambridge University Press, 1979), 92-123. See 93.

speaker's meanings cast far too wide a net to catch metaphorical meaning, which is found only in contexts which include a linguistic context. Also, possible speaker's meanings which are candidates for metaphorical meaning are so only because they are contextual meanings. Because they are something that could be meant by the relevant utterance, given the context, then they could possibly be meant by a speaker. Such speaker's meanings are parasitic on contextual meaning. In addition, there is no guarantee that a contextual meaning will be such that it might, in fact, be uttered and meant by a speaker. There might be some taboo, or mental barrier or whatever. Again, contextual meaning is the prior notion, and a possible speaker's meaning is so only because it is a contextual meaning. So: a possible speaker's meaning can qualify as metaphorical meaning only insofar as it is a contextual meaning, and cannot qualify unless it is a contextual meaning.

So metaphorical meaning had better be regarded as contextual meaning. However, according to Davidson the only meaning a metaphorical utterance has is its literal meaning. There is more. Metaphors have no cognitive content, no cognitive content whatever, although, given that it would seem they share this feature with pictures, perhaps they might yet be allowed to be useful. Davidson's banishment of metaphorical meaning seems intuitively implausible, after all we think metaphors can be true or false, and we run them in arguments and inferences.

Davidson's extended broadside against metaphors' involving any meaning other than literal meaning is conducted, in fact, as a broadside against any notion that *words* have meaning other than their literal meaning. Speaker's meaning is not mentioned, never mind contextual meaning. However, the positive Davidson picture that emerges in the course of the broadside, makes it unlikely that either speaker's meaning or contextual meaning would shift him.

According to Davidson we are to respond to the metaphor, to give ourselves over to it. The effects the metaphor produces in us are what matters. Notably, 'there is no limit to what a metaphor calls to our attention and much of what we are called to notice is not propositionain character' (p.44). What matters is the experience. That is why there is no meaning other than the literal; that is why metaphors are not paraphrasable. 'Metaphor makes us see one thing as another... Since in most cases what the metaphor prompts or inspires is not entirely, or even at all, recognition of some truth or fact, the attempt to give literal expression to the content of the metaphor is simply misguided' (p. 45).

But metaphors seem to be capable of being true or false, and they cannot be so and lack meaning. Now Davidson is happy to accept that metaphor 'does lead us to notice what might otherwise not be noticed, and there is no reason I suppose, not to say these visions, thoughts, and feelings inspired by the metaphor, are true or false' (p. 39). However, he insists that 'the sentences in which metaphors occur are true or false in a normal, literal way, for if the words in them don't have special meaning, sentences don't have special truth. This is not to deny that there is such a thing as metaphorical truth, only to deny it of sentences' (p.39).

Why should it be thought that meaning which is other than standard will produce non-standard truth? If it is a different concept of meaning then the consequence might follow, just as if metaphorical meaning is mysterious then metaphorical truth might be mysterious. But speaker's meaning is not mysterious, nor is it a different concept of meaning. Standard meaning is, of course, different from speaker's meaning but that is not because there are two different concepts of meaning. Why should speaker's meaning produce only speaker's truth?

What might be meant by such a claim? Speaker's meaning gives us no clues to what might be meant by speaker's «truth».

Perhaps we should regard Davidson as having done nothing other than reject the idea that metaphorical sentences acquire new meanings. So interpreted, he would be happy with metaphorical meaning otherwise understood. However, that would bring the problem of paraphrase back again, and also seems incompatible with his positive characterisation of metaphor. Perhaps his view is that metaphor may cause us, for example, to see that some thought is true. But that would be one only of many possible effects, and would be an effect of the metaphor, not intrinsic to it. The inspired thought could be called true, not the metaphor, while the inspired vision could only metaphorically be called true.

Is he right, so interpreted? Those who attach metaphorical meaning to words are claiming that the words on that occasion of their use have a meaning other than their standard meaning. Just as the word 'insipissated', used by someone who has the wrong notion of what 'insipissated' means, will have a sense other than its standard sense. They will also (generally) mean by the utterance, in which no doubt 'insipissated' will be the star turn, something other than what is standardly meant. If they have expressed, in so doing, the belief (say) that my categorisation of meaning is unduly complicated then their claim is, of course, false, while what they have (strictly) said (that it is complexly dense) might by some be thought to be true. Surely, here, true is true and false is false, and the twain are no more likely to meet here, than they are anywhere else.

Another example of nonstandard meaning's allowing for truth, is irony. Someone says, of an embarrassingly elementary paper, that it was insipissated, and opines, by saying that, something like that it was embarrassingly elementary. Someone could disagree and think the opinion false, or agree and think it spot-on true. This is surely plain old, boring old truth, spot-on or otherwise.

Truth has been attached in these two examples in the one case to sentences (as uttered) and in the other to what has been meant by uttering the sentence. In the examples given, the point against Davidson could have been driven home this way: what has been (strictly) said (that the paper was insipissated) is not what the sentence means, in Unduly Complicated's case, nor what is meant by the sentence, in the case of Complexly Dense's ironic utterance. But what is, in fact, meant could as well, though no better, have been said directly. Why then, given this equivalence, should direct opining be true or false but its deviant sibling be denied this excitement?

It is true that metaphors differ from the examples above in that they are (generally) not paraphrasable. But being paraphrasable is neither necessary nor sufficient for truth. 'This little rooster is red' is not paraphrasable. I pluck this example fairly randomly from many on offer. 'Rhett Butler was a gentleman'. 'Dalziell is a drongo'. 'He's a ratbag but he's me mate'. But it might be thought that metaphors are in principle, if they are so, not paraphrasable. Why, though, should this difference, if indeed it exists, have as a consequence that such metaphors have only literal meaning? Trying to explain what a ratbag is, in practice is very like trying to explain what a metaphor means. We grope, and continue to grope, never doubting that we know.

Those clinging to this difference, should consider some literal comparisons. 'She looks like Bette Davis', where it is known how Bette Davis looks, will be loaded with information, which cannot, in principle, be paraphrased. I might well say, finally having found this Bette

Davis look-alike, all hope suddenly gone: ‘She doesn’t at all!’ More was meant by saying what was said, than was said.

This talk of meaning should not be resisted. Someone is claimed to be like a colleague, where the context makes clear that what is being claimed is that they are kind. There should be no unease in treating this as a case of speaker’s meaning, where what is meant by saying ‘He is like Alec’ is that the person is kind. Now this is paraphrasable; but what if the claimed likeness is in respect of my unusual walk. Then I think they mean more by what they have said than they have said, but that more is not paraphrasable.

There is, then, no problem in principle in metaphors meaning other than their literal meaning; nor is there a problem in principle in what they mean or might mean not being paraphrasable; nor in their being true or false. However, Davidson seems to have the advantage that he has a clear explanation for metaphors not being paraphrasable: nothing other than the literal is meant, so there is nothing to paraphrase.

Searle believes that metaphors are intrinsically not paraphrasable, ‘because without using the metaphorical expression, we will not reproduce the semantic content which occurred in the reader’s comprehension of the utterance’ (p.123). But this would seem to hold equally of ironic utterances and they do not seem to be intrinsically not paraphrasable.

Black refers to Toynbee’s ‘No annihilation without representation’, and says that this ‘could no doubt be spelled out to render his allusion to the familiar slogan boringly explicit...[but]...something of the force and point of the original remark would then have been lost.’⁴ However, it would not be boring to someone who did not know of the allusion, and would surely be egregious rather than boring to those who did. Once again, though, ironic utterance would turn out not to be paraphrasable on this test.

What these examples demonstrate is that we should not confuse a metaphor’s being paraphrasable with its being replaceable without loss. A poem is not replaceable by a literal paraphrase. If the poem has meaning, what is meant by the poem would not exhaust the features that make it a poem, make it valuable. No more is the claim that someone is AC/DC, replaceable; yet it is no trouble to say what it means. So a metaphor can have meaning even though there is more to a metaphor than its meaning.

So why are metaphors generally not paraphrasable? Consider a feature of non-metaphorical language. ‘Red’ enables us to refer to an element of our colour experience. ‘Bette Davis eyes’ allows us to refer to such eyes. However, ‘It’s red’ isn’t paraphrasable, and intrinsically so, because it depends on the link with experience; as, for another example, does our understanding of what a burning pain is like. The experience can be picked out, but cannot be put into words. The way someone looks can be picked out, but cannot be put into words. Metaphors enable us to pick out more complex features of reality. It has been claimed that the tango is a metaphor for the way men and women relate. If it is indeed so then that view of an aspect of reality, and how it is experienced, has been presented by the metaphor, but cannot otherwise be presented.

Metaphors can make very specific links with features of reality. Leave the metaphorical utterance the same but change its context, then it has changed. The metaphor only exists in

⁴ Max Black, ‘How metaphors work: a reply to Donald Davidson’, in Sacks (op.cit.), 181-192. See 192.

context, so the metaphorical utterance can be the same and the metaphor different, if the context is different. What metaphors convey is a singularity.

Is there still a problem for the devotees of meaning? Simon Blackburn thinks so, stressing that metaphors are open-ended, indefinite. Referring to Romeo's calling Juliet the sun, he claims that our response to the metaphor 'is quite open-ended. Shakespeare need have had no definite range of comparisons which he intended, and it is quite wrong to substitute some definite list and suppose the exploration is complete. The metaphor is in effect an invitation to explore comparisons.

But it is not associated with any belief or intention, let alone any set of rules, determining when the exploration is finished.'

The first point to make on indefiniteness is that literal meanings can be indefinite. Any literary text can receive more than one interpretation. It is not thought that there is a limit to the number of possible, even sensible, even eminently satisfying interpretations. Nor is it thought that this means that complex texts lack meaning. By parity of reasoning, the fact that there need not be just one «meaning» to a metaphor, is no reason for denying that metaphors have meaning. They can have «meanings», so they can have meaning. Presumably the controversies that surround the interpretation of texts will gather round the interpretation of metaphors: is there a correct interpretation?; is it the author's?; and so on and so on.

If it is insisted that the indefiniteness that attaches to metaphor attaches to the individual indefiniteness of the list of comparisons rather than there being an indefinite number of lists; to the individual «meaning» rather than a range of «meanings», the response is that this feature is also found in literal language. Open-texture and family resemblance are relevant theoretical notions that immediately spring to mind.

Another response to the indefiniteness problem is this, using the «comparison» approach for illustration. That A is like B need not tell us much. In context, that A is like B might tell us a good deal, to the extent of shocking us, by way of telling us, via a literal comparison, that A is evil. But the information might be less specific. In context, that A is like B tells us that A has some of the salient features of B, warmhearted to a fault etc., but does not pick out just which of those features A shares with B. Context is generally important here, as much to rule out as to rule in possible features. Figurative comparisons exhibit the same range within the informative band. The relevance of these facts to the indefiniteness issue does not depend on accepting the «comparison» approach. All that is needed for their relevance is the indefiniteness possible in that approach.

Susan Haack has stressed the cognitive usefulness of metaphor, and she has emphasised that 'it is precisely because *metaphorical* statements are unspecific or open-textured that they are apt for representing novel conjectures in their initial and undeveloped stages, and for prompting investigation of what might be specific respects of resemblance.'⁵ She has outlined elsewhere in some detail the exploratory usefulness of metaphor by way of her recounting how she worked her way towards her «foundherentism»; and the indispensable role played in that by the notion that 'the way a person's beliefs about the world support one another is rather

⁵ Susan Haack, 'Surprising Noises: Rorty and Hesse on Metaphor', *Aristotelian Society Proceedings*, New Series 88 (1987/1988), 293-301. See 299.

like the intersecting entries in a crossword.’⁶ The responses above accommodate these important truths.

The attention just given to indefiniteness should not be seen as countenancing the degree of open-endedness to a metaphor’s invocations that Robert Fogelin has characterised as allowing a ‘drift into the Davidsonian void’ (p.112) of possible comparisons (or experiences or whatever). Any open-endedness in metaphor is always subject to constraints, sometimes powerful constraints. Fogelin has done a splendid job of elaborating the role of context in shaping the interpretation of metaphor, most markedly in the way a poem sets limits to our understanding of its metaphors (see pp.108-112).

Which brings me back to contextual meaning. I have argued that if meaning other than literal meaning inhabits metaphor, it has to be contextual meaning, not speaker’s meaning, not word or sentence meaning (occasion meaning). I am now able to claim more boldly that metaphors do have meaning other than literal meaning and that this is contextual meaning. If metaphors are to be allowed the possibility of being true, then metaphors had better have meaning.

Alec Hyslop

<A.Hyslop@latrobe.edu.au>

⁶ Susan Haack, ‘Dry Truth and Real Knowledge’, in Hintikka (op.cit.), 1-22. See 18.

SORITES (ΣΩΡΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 33-47

The Justification of Deduction

Copyright © by SORITES and Silvio Pinto

THE JUSTIFICATION OF DEDUCTION

Silvio Pinto

§1. The Problem

It has been a long-standing claim of Michael Dummett that deduction poses a philosophical problem of its own justification. According to him, when philosophers set out to look for a justification for our practice of inferring, they want an explanation of the validity and fruitfulness of the rules of inference that we accept as valid and fruitful. The fruitfulness of valid deductive argument simply means that new knowledge is gained in the transition from the argument's premises to its conclusion. Another way of expressing this is to say that knowledge of the conclusion of the argument is not entailed by knowledge of its premises.

Probably the first philosopher to recognize the essential fruitfulness of deduction was Frege.¹ For him, understanding a proof always requires a creative act of forming new concepts via the process of carving the thoughts expressed by its various steps in ever novel ways. Frege's insight was meant to accommodate the undeniable fertility of mathematical proof but it was never developed into a systematic explanation of fruitfulness. This may be due to the fact that his recognition of both the validity and the fruitfulness of logic created a tension in his philosophy. For crucial to the Fregean explanation of validity is the idea that articulated thoughts exist independently of us so that their logical relations (for instance, entailment) obtain or not no matter whether we grasp them or not. But if a pair of thoughts is to stand in the relation of logical consequence then not all of their possible partitionings into constituent senses will be allowed. Validity must therefore impose certain ways of carving each of the thoughts in a proof into its respective constituents. That this is so in Frege's view, however, seems to go against his own explanation of fruitfulness since this latter forces upon us the opposite, apparently very anti-Fregean picture of concept formation, namely, that of proofs as producing new concepts—or senses, as Frege would say.

Following Frege, Dummett also claims that deduction is essentially fruitful. As I understand it, his argument for this claim is the following. If we do not assume that each case of deduction is—even if sometimes infinitesimally—fertile then there would be no way of explaining the numerous obviously fruitful proofs in mathematics.² One of Dummett's preferred examples is that of Euler's famous proof of the impossibility of an uninterrupted route that crosses over all the 7 bridges in eighteenth century Königsberg without crossing any

¹ See, for example, Frege 1950, par. 88.

² This can be found in Dummett 1973b, p. 297 and also in Dummett 1991, pp. 175-6.

one of them twice.³ Before the proof was found, there was already a way of checking the conclusion of Euler's argument, which proceeded by enumerating all the finite many possible routes and showing that each one of them required that at least one bridge be crossed twice. What Euler discovered was a new and more elegant method for proving this. He gave us a new way of representing an arbitrary seven-step route. Similarly, Cantor invented a new method for constructing a real number which is different from all the denumerably many rational numbers between 0 and 1. But could we not equally well explain all these cases of undeniable fertility of deduction in terms of the non-preservation of the triviality of the basic deductive steps when they appear linked together in an innovative proof? Or maybe by claiming that what makes them really surprising is the introduction and later elimination of unexpected premises?

I take it that Dummett would deny that such proposals were really offering any explanation of fruitfulness. For if the notion of a trivial deductive step is to mean that the step is such that there is no epistemic gap between the step's premises and its conclusion, then, to use the analogy between epistemic and spatial distance, we are left with no clue as to how a proof can be fruitful—that is, how there can be an (in many cases, considerable) epistemic distance between its assumptions and conclusion—given that it is assumed to consist of a sequence of deductive transitions whose premises-conclusion epistemic length measures zero. On the other hand, if the explanatorily relevant notion for fruitfulness is that of unexpected premises, then it is not at all clear how the presence of these premises could account for the epistemic gap between the premises and the conclusion of many deductive arguments. What needs clarification is the epistemic distance between two locations in the deductive chain. A fruitful assumption or definition is just a location on our geometrical picture of fruitfulness. But we want to know what makes the distance between two locations in the chain possible.

The problem of the justification of deduction, as Dummett sees it, is therefore that of providing a philosophical explanation for the validity and the fruitfulness of those forms of argument which are valid and fruitful. It is incumbent on the theory of meaning for the language, in his view, to supply such an explanatory argument. The difficulty is that the requirements of either one of them seem to conflict with the requirements of the other: validity seems to demand that the conclusion brings in nothing new with respect to its premises whereas fruitfulness seems to require exactly the opposite.⁴ This is how Dummett expresses the puzzle:

For it [deduction] to be legitimate, the process of recognising the premises as true must already have accomplished whatever is needed for the recognition of the truth of the conclusion; for it to be useful, a recognition of its truth need not actually have been accorded to the conclusion when it was accorded to

³ See Terquem & Geronio 1851, pp. 106-119 for the French translation of Euler's paper (1736).

⁴ One attempt to dismiss Dummett's problem was made by Susan Haack (1982, pp. 225-227). According to her, the tension between the requirements of validity and fruitfulness can be relieved if we distinguish between two senses of deduction. The first is that of deductive implication, which she characterizes as the relation of logical consequence between propositions. The second sense of deduction is that which corresponds to the intentional act of making an inference, which she calls 'deductive inference'. Haack's claim is that validity is a property of deductive implications whereas fruitfulness applies only to deductive inferences. This strikes me as incorrect, since deductive inferences can also be evaluated as valid or not; neither can I see why the pair of predicates 'fruitful/non-fruitful' could not be attributed to deductive implications.

the premises. Of course, no definite contradiction stands in the way of satisfying these two requirements: recognising the premises as true involves a possibility of recognising the conclusion as true, a possibility which will not in all cases be actualised. Yet it is a delicate matter so to describe the connection between premises and conclusion as to display clearly the way in which both requirements are fulfilled. (Dummett 1973b, p. 297)

Dummett maintains that only a molecularist theory of meaning can deliver such a harmonious explanation of validity and fruitfulness. Generally speaking, a meaning molecularist believes that the meaning of a word or a sentence is fixed by a linguistic unit which is well short of all the language. His opponent in this debate is the meaning holist, someone whom Dummett represents as denying that for each word or sentence of a language there is a group of sentences the knowledge of whose meanings is sufficient to determine the meaning of the mentioned word or sentence. In other words: the meaning holist does not believe that each expression of the language splits the set of sentences of which it is a constituent into a constitutive and a collateral group. Clarification of the constitutive/collateral terminology as applied to words and sentences as well as of Dummett's reason to maintain that the holist cannot reconcile the validity with the fruitfulness of deduction now requires that we get into the detail of the controversy between molecularism and holism.

§2. Molecularism versus Holism

The terms 'molecularism' and 'holism' will be used here to designate two mutually exclusive positions in the philosophy of language concerning what should be taken as a sufficient basis for fixing the meaning of each expression of a language. But how can the notion of the basis for fixing meaning or its equivalents—for instance, unit of meaning—be cashed out? Quine,⁵ in his well-known criticism of the logical empiricists' criterion of empirical significance, made a very strong case for the thesis that linguistic content cannot be attributed to sentences in isolation but only to a larger linguistic unit, which he sometimes identifies with our language and some other times with our present total scientific theory. This is so because in Quine's own terminology sentences do not face the tribunal of experience—which in agreement with the logical positivists he takes to be the source of linguistic content—piecemeal, i.e. one by one or even in small groups. Hence, the unit of meaning is, according to Quine, that minimal linguistic whole to which empirical content can be attributed independently of any content attribution to any larger linguistic unit.

Another way of spelling out the notion of meaning unit and one in which the constitutive/collateral dichotomy comes out quite naturally was suggested by Dummett himself. In the opening pages of his first book on Frege,⁶ Dummett mentions a circularity (certainly known to Wittgenstein in the *Tractatus* and to Russell) which appears to threaten any view that takes sentence meaning to be explanatorily prior to word meaning. If the latter is explained in terms of sentence meaning by saying that it consists of the contribution the word makes to the meaning of the sentences in which it occurs,⁷ then sentence meaning would be circularly characterized were it also proposed that the meaning of a sentence is to

⁵ See Quine 1951, sections 5 and 6.

⁶ Dummett 1973a, pp. 4-5.

⁷ As someone who accepts Frege's context principle (which makes its first appearance in Frege's writings in Frege 1950, introduction) must admit.

be explained in terms of the meanings of its constituent words. A possible way out of the circle is the one which, according to Dummett, Frege uses, namely: to elucidate sentence meaning in terms of truth-conditions. In this early book, Dummett did not think there was any danger of circularity at the level of understanding; knowing the meaning of all the constituent words in a sentence is obviously necessary, and therefore in this sense prior, to knowing the meaning of the sentence.

In his later work,⁸ however, Dummett suggests that the circularity problem discussed above can also arise at the level of understanding for the anti-atomist conceptions of linguistic content recognition (i.e. for those who regard as absurd the idea that grasping the meaning of words is prior to understanding whole sentences in which these words appear). The circle seems very present, for example, in the famous Tractarian passages where it is said that understanding sentences requires understanding their respective names and that knowing the meaning of each name presupposes understanding the sentences containing it.⁹ The problem is then how to spell out a competent speaker's knowledge of the meaning of a word in such a way as, on the one hand, to express it in terms of his primary understanding of sentences (as should be expected from an anti-atomist about content recognition) and, on the other hand, not to fall into the sentence/word meaning circularity. Dummett proposes to tackle this problem by distinguishing two groups of sentences for each expression of the English language. Let us call them, respectively, the expression's constitutive and collateral group.¹⁰ The first consists of all the sentences containing the expression which together provide the basis of our understanding it. The second group comprises all the other sentences containing the expression; it is called collateral group because understanding any sentence belonging to this second group requires a prior understanding of the expression in question. Here is what Dummett says about the constitutive/collateral distinction:

the priority of sentence-meaning over word-meaning requires the understanding of a word to consist in the ability to understand *certain* sentences, or more exactly, at least some of the sentences of a certain range, in which it occurs. (...) The compositional principle demands that, for any given expression, we should distinguish between two kinds of sentence containing it. An understanding of the expression will consist in the ability to understand representative sentences of the first kind and does not, therefore, *precede* the understanding of sentences of that kind. By contrast, an antecedent understanding of the expression will combine with an understanding of the other constituent expressions to *yield* an understanding of a sentence of the second kind, which demands an understanding of the expression but is not demanded by it. (Dummett 1991, p. 224)

Later in the same text, Dummett adds that the constitutive group of a word must contain the simplest sentences containing it, whereas there is no limit to the complexity of the sentences that can figure in the word's collateral group. He illustrates this with the word 'fragile': while the sentence «this plate is fragile» should be part of its constitutive group, «I'm afraid I forgot that it was fragile» should belong to 'fragile's collateral group. Dummett equates understanding a word with the ability to understand the simplest sentences containing the word. According to him, this is the picture one gets if he takes the molecularist standpoint

⁸ For instance, in Dummett 1991.

⁹ Wittgenstein 1922: 3.263, 3.3.

¹⁰ The terminology is not Dummett's. I took it from a paper on meaning holism by Eric Lormand (1996).

on the issue of recognition of linguistic content. That is, if he insists that knowledge of the meaning of a word does not require more than understanding a sufficiently small fragment of the language. All the sentences containing the word which are situated beyond the boundaries of this fragment do not contribute to fix its meaning; on the contrary: a previous knowledge of the word's meaning is necessary in order to understand them.

That there is for each word such a fragment well short of the totality of language is what the meaning holist wants to deny. Although he may accept that sentence meaning is—at least partially—determined by the meaning of its constituent words plus the sentence's syntactical structure, the holist will take issue with the thesis that word meaning is always determined by an appropriate fragment of the language consisting of logically non-complex sentences. His rejection of the thesis may take various forms. A more radical holist might, for example, hold that the meaning of any sub-sentential expression is fixed only by the whole of the linguistic network. Echoing a familiar sort of holism about belief, such a holist would say that the content of any word or sentence can only be individuated in the context of the totality of its linguistic web; any smaller unit would leave its meaning substantially undetermined. Quine's holism, for instance, has been interpreted in this way.¹¹ A moderate holist might, however, reject the molecularist thesis merely on the ground of his belief that the meaning of a new sentence is not completely determined by the constitutive groups of its respective constituent words and adequate hypotheses about the sentence's syntax. This kind of holist will be willing to allow for the possibility that the new use of the sentence be tracked down to a new use of one or more of its constituent words. Davidson's linguistic holism seems to me to be of this latter moderate variety.¹²

Dummett's purely constitutive way of characterizing the constitutive/collateral distinction and therefore the molecularist/holist dichotomy does not seem to me to be entirely satisfactory. By 'purely constitutive way' I mean to point out that his rendering of the distinction is in terms of what constitutes a speaker's grasp of the meaning of a word. His answer to this constitutive question appeals to an ability to understand certain simple sentences containing the word. My uneasiness with Dummett's route to establish the constitutive/collateral distinction relates to what I take to be a gap in his explanation of the priority of sentence meaning over word meaning for the members of a word's constitutive group. If for the members of this group sentence meaning is prior to word meaning, how can a speaker ever get from the former to the latter? In other words, how can his grasp of the meaning of a finite number of sentences ever be sufficient for the speaker to derive the meaning of a word they all have in common? After all, one might say, the kind of understanding of sentences that is at stake for the members of the constitutive group cannot be further analyzed in terms of the understanding of these sentences' constituent parts. But, he might insist, if compositionality and structure do not play a role in the grasp of the content of any member of the constitutive group, then the non-analyzed content of any finite group of sentences, however numerous, cannot on their own ever yield the content of any of these sentences' constituent parts.¹³

¹¹ See, for example, his celebrated *Two Dogmas of Empiricism* (1951).

¹² I was convinced of this by his paper on malapropisms (Davidson 1986).

¹³ Crispin Wright reminded me of this weakness of Dummett's distinction.

Within the framework of Dummett's explication, the constitutive/collateral distinction threatens to collapse.

I reckon the gap in the account of the constitutive/collateral dichotomy offered by Dummett can be bridged if we switch from his constitutive perspective to a radically interpretive one. The suggestion is that we consider the possibility of drawing such a distinction from the perspective of someone who is trying to learn the language for the first time. A very illuminating description of the interpretive framework has been provided by Neil Tennant in the following passage:

When we discern meanings as we learn the language, it is primarily whole sentences whose meanings we work out in context. (...) Sufficiently many sentences tentatively grasped allow meaning to coagulate upon the words they have in common. This coagulation of meaning is constrained by conjecture as to the structure of each sentence grasped thus far. The structure of a sentence is a matter of how the words and phrases and clauses have been composed within it. Once global sentence meanings plus structural hypotheses have conferred meanings upon individual words, the process is then reversed. New sentences are understood on the basis of one's assignment of meanings to words, and one's analysis of how the words are put together to make the sentence. The possibility in principle of the eventual success of the compositional method is a methodological principle guiding the language learner. (Tennant 1987, p. 31)

In this account, the language learner (or radical interpreter) starts with hypotheses about the content of the speech acts of his subjects which involve whole sentences (his assertions, for example) and about the syntactical structure and composition of these initial sentences. From these, he derives hypotheses about the meanings of the sentences' component words. Once the latter hypotheses have been firmly corroborated by the data, then the initial hypotheses about sentence meaning are no longer needed. For, at this stage (when the interpreter has learned the language), the meaning of a new sentence can already be obtained from word meaning plus the hypotheses about the sentence's structure. In Tennant's picture, this is the stage at which the initial procedure of inferring claims about word meaning from hypotheses about sentence meaning and about syntactical structure is reversed.

To see how Dummett's distinction is presupposed by Tennant's account, consider again the sentences with which his compositional method starts. These are the sentences without which, according to Tennant, the hypotheses about the meaning of the words such sentences have in common would find no support. The content of these sentences constrained by suitable further hypotheses about their structure will eventually converge into the content of the words in question. Hence the set of initial sentences make up their shared words' respective constitutive group. Once the sentences of a word's constitutive group have fully determined its meaning, which is guaranteed by the success of the compositional method, then any new sentence containing the word will have its meaning determined by the meaning of the word and the meanings of the sentence's other constituent expressions, as these latter meanings are already given by the expressions' respective constitutive groups; all such new sentences will belong to the word's collateral group. Tennant's description of the compositional method therefore corresponds to a molecularist view of language. For only a molecularist would maintain that finitely many hypotheses about sentence meaning could ever come to suffice for the determination of word meaning.

Tennant's approach also suggests a way of escaping the difficulty mentioned five paragraphs back about the transition from sentence meaning to word meaning. The problem, let us recall, was that if understanding the sentences of a word's constitutive group does not involve discerning in these sentences' contents any structure then it appears that no finite

amount of constitutive sentences understood by an interpreter will suffice for him to derive the content of the word common to them. According to the compositional method, what enables him to infer word meaning from hypotheses about the word's constitutive sentences' meanings are the cited additional hypotheses concerning the syntax of these sentences. Once word meaning is firmly determined, then this kind of understanding of sentences (i.e. unstructured grasp of their content) is no more necessary as, with the reversal of the process, sentences can now be understood on the basis of their constituent words' meanings and of their composition (i.e. grasp of the sentence's content as structured in a certain way). The first kind of sentence understanding (unstructured grasp of the sentence's meaning) is only prior to word understanding until the interpreter finds his way into the alien language, at which point it ceases to have any application and the direction of priority runs from word understanding to the second kind of sentence understanding (structured grasp of the sentence's meaning). This is how a molecularist account can avoid the sentence/word meaning circularity.

By contrast, this discussion of Tennant's compositional method also helps us see what a holistic picture of meaning would be like. For all the holist needs to reject is the molecularist assumption that the compositional method will eventually succeed. According to such a holist, the interpreter will never reach a stage where sufficient contact with sentences containing a word uttered by the alien speakers will enable him to settle on a meaning for that word which will determine the meaning of any new sentence containing it exclusively from the content of its constituent words and the way they are put together. The reason for this, claims this kind of holist, is that no matter how many sentences containing the word an interpreter considers, there will always be new sentences whose understanding cannot be reduced to the old meaning of their parts and the way they are combined together. In order to understand these new sentences, the interpreter will have to revise his previous hypotheses about word meaning and structure in the light of this new datum: a new use of a word. For this sort of holist a new use of a sentence is not always reducible to the previous uses of the words in it. This means that no fixed group of sentences is fully constitutive of the meaning of any sub-sentential expression, so that the rest is collateral. This holist can make no sense of the constitutive/collateral distinction. As seen above, one need not be more than a moderate holist in order to adopt the view of Tennant's opponent.

§3. Dummett on the Justification of the Deduction

The holist cannot, according to Dummett, offer a harmonious explanation of the validity and fruitfulness of deduction since he fails to account for validity. And this latter cannot be explained within a holistic approach to language, Dummett insists, because in the context of this approach there is no room for the constitutive/collateral distinction.¹⁴ The connection between the availability of this distinction and the possibility of a philosophical elucidation of validity is nonetheless not straightforward. Below, I will attempt to make it explicit by considering Dummett's account of validity.

§3.1. Dummett on Validity

As our interest here is in deduction, let us start with the question of how, for the molecular verificationist, the constitutive/collateral distinction applies to the logical constants. His claim is that the constitutive group of a logical constant comprises just its introduction

¹⁴ See, for instance, Dummett 1973b, pp. 300-305.

rules.¹⁵ In a system of natural deduction, each logical connective is characterized by a set of introduction and elimination rules¹⁶—a characterization which can easily be shown to be equivalent to one in terms of the connective’s truth-table.

But why does the verificationist privilege I-rules? The reason is this. It is a basic tenet of molecular verificationism that the meaning of a sentence consists in the canonical method of verifying it. This method—the sentence’s direct means of verification—will surely involve deductive inference if the sentence is not atomic. But in this case we need only apply the I-rules for the sentence’s respective logical constants. For instance, a sentence like «6 is even and is perfect» can be verified directly via the separate computations of the two conjuncts and posterior application of the I-rule for conjunction. There may be other methods of verifying this sentence which use logically more complex sentences—for example, the sentence «some perfect numbers are not odd»—or other auxiliary sentences. Think, for instance, of the verification of Euler’s proof’s conclusion (the sentence «any uninterrupted minimal route through Königsberg’s 7 bridges contains more than 7 bridge-crossings»). The proof gives us a new method of establishing the truth of this sentence which proceeds from the auxiliary sentence «consider any route through Königsberg’s bridges which is uninterrupted, minimal and crosses all the 7 bridges». By contrast, the canonical method of verifying the conclusion of Euler’s proof exploits only its internal logical concatenation of smaller sentence units; it makes exclusive use of the introduction rule of ‘any’, which in this case is equivalent to a finite conjunction. Any non-canonical method for establishing the truth-value of sentences—i.e. those which appeal to elimination rules as well—Dummett labels an indirect means of verification.

The constitutive/collateral distinction with respect to sentences corresponds in Dummett’s view to the dichotomy between direct and indirect means.¹⁷ According to him, to know the meaning of a sentence is to know how to verify it directly. All the other means of sentence verification do not contribute to constitute its meaning and therefore belong to the sentence’s collateral group. Thus, Euler’s proof offers an indirect means for verifying its conclusion; the canonical method of establishing its truth consists of enumerating and demonstrating all its conjuncts and then using the I-rule of conjunction.

Now, what does the characterization of the constitutive/collateral distinction for sentences have to do with the homonym distinction for the logical constants? The following is the verificationist story. First of all, as Frege had already realized, in order to give the meaning of a logical constant it is enough to present the truth table of an arbitrary sentence for which it is the main connective. This is equivalent to taking all the sentences for which the constant is the main connective as members of the latter’s constitutive group. Secondly, according to the verificationist, the meaning of all these sentences is identified with a direct means of verifying them and their main constant contributes to this means in so far as its respective I-rules warrant the transition from the direct methods of verification for both of the sentence’s immediate constituent sentences to the direct method of verification for the sentence itself. Thus, if we know the meaning of «6 is perfect» and «6 is even», then this together with the

¹⁵ Dummett 1991, p. 252.

¹⁶ I- and E-rules, for short.

¹⁷ Dummett 1991, p. 229.

I-rule for conjunction will tell us how to verify «6 is perfect and even» directly. Hence, within the molecular verificationist framework, the meaning of the logical constants is fixed by its respective I-rules.

Having established this, our next task is to clarify the connection between the verificationist characterization of the constitutive group for the logical constants and his explanation of validity. Dummett claims that this latter explanation must appeal to the notion of a harmony between the two main aspects of our practice of uttering sentences to make statements. Deduction is valid, according to the verificationist, if the practice of offering grounds for an asserted sentence—that aspect which is verificationally constitutive of its meaning—is in harmony with that of drawing consequences from it—the verificationally collateral aspect of the meaning of a sentence. Thus, Dummett says:

For utterances considered quite generally, the bifurcation between the two aspects of their use lies in the distinction between the conventions governing the occasions on which the utterance is appropriately made and those governing both the responses of the hearer and what the speaker commits himself to by making the utterance: schematically, between the *conditions for* the utterance and the *consequences of* it. (...) Plainly, the requirement of harmony between these in respect of some type of statement is the requirement that the addition of statements of that type to the language produces a conservative extension of the language. (Dummett 1973c, p. 221)

The quotation is explicit of how Dummett thinks a verificationist should cash out the vaguer notion of harmony, namely: in terms of the more precise notion of conservative extension. The latter has a precise sense, however, only when applied to formalized languages. Thus, of two formalized mathematical systems A and B expressed respectively in languages L_A and L_B we can say that A is a conservative extension of B if and only if L_A extends L_B and no sentence belonging to L_B could be deduced from A that was not already deducible from B. Nothing like the concept of deducibility exists for natural language though; our reasons for asserting the truth of non-mathematical sentences are normally defeasible to a lesser or greater extent. But if we cannot speak for natural language of conservatism with respect to provability with the resources of the non-extended language, how else could conservatism in this domain be spelled out?

Dummett claims that the verificationist has a plausible natural language surrogate for deducibility, namely, the notion of verifiability via whatever means is available for truth-value attribution.¹⁸ He suggests that an extension (NL_{f+e}) of a certain fragment of natural language (NL_f) is a conservative extension of NL_f if the incorporation of new methods of verifying sentences propitiated by NL_{f+e} does not change the truth-value of any sentence already verifiable with the resources of NL_f . In other words, the new language which results from the enrichment of a fragment of it by new vocabulary is conservative relative to the latter fragment if any sentence which is verifiable in the scanner language, and hence has already a determinate truth-value assigned to it, does not have that truth-value altered as a result of the application of any new means of verification made available by the richer language. Still another way of putting the verificationist account of conservative extension is this: NL_{f+e} extends NL_f conservatively if the new vocabulary introduced by NL_{f+e} does not conflict with that which determines the meaning of the expressions and sentences of NL_f . Now, we know that for the verificationist sentence meaning is fixed by its direct means. Hence, the conflict we alluded to must be between an eventual indirect means of verification of a sentence of NL_f

¹⁸ For example, in Dummett 1991, pp. 218-9.

introduced by NL_{f+e} and its direct means supposedly available in NL_f : these two means will be in conflict if they attribute different truth-values to the sentence.

Let us illustrate the Dummettian concept of conservative extension with a counter-example. Think of NL_f as any fragment of natural language and of NL_{f+e} as this fragment together with Arthur Prior's famous logical constant 'tonk' which he defined in the following way:

$$\text{I-tonk: } \frac{p}{p\text{-tonk-}q} \quad \parallel \quad \text{E-tonk: } \frac{p\text{-tonk-}q}{q}$$

where 'p' and 'q' are sentences of NL_f or NL_{f+e} .¹⁹ Tonk does obviously not extend NL_f conservatively since if we are justified (by a direct verification, say) in attributing the truth-values *true* and *false* to two sentences 'r' and 's' of NL_f respectively, then with the introduction of 'tonk' the truth of 's' can be (indirectly) established as well. The presence of 'tonk' would provoke a disharmony between the practice of grounding asserted sentences—for instance, '¬s' through its direct means—and that of drawing consequences from asserted sentences—for example, 's' as consequence of 'r'. The lesson to be learnt from 'tonk' is therefore the following: the I- and E-rules of logical vocabulary which does not extend its respective language conservatively in the verificationist sense are to be deemed invalid. As the logical constants are also an essential part of the indirect means of verifying sentences (via their respective I- and E-rules) and as for the verificationist sentence meaning consists in the direct means of establishing the sentence's truth-value, non-conservative logical constants—that is, those whose I- and E-rules are invalid—might easily upset the meaning of a sentence. Such a modification in sentence meaning would consist in the alteration of the truth-value conferred to the sentence before the logical constant's incorporation into the language. Alteration in sentence meaning would in turn be translated into an alteration in the meanings of their constituent words. So, if we do not want meaning to change in the transition from suitably rich fragments of a language to a larger one—i.e. if we side with the meaning molecularist—this must be our reason for banning non-conservative logical constants from our linguistic practice.

This is the molecular verificationist account of validity. According to Dummett, if one denies the constitutive/collateral distinction with respect to sentences on the grounds that their various indirect means also contribute to fix the sentences' respective meanings—i.e. if one adopts a verificational holistic position—then he would be left with no constraints to impose on a certain linguistic practice. He would be in possession of no criterion to exclude disharmonious practices of inference as unacceptable and could not explain why some other deductive practices are in order. I have argued elsewhere against Dummett's claim that a holist cannot provide a satisfactory alternative explanation of validity.²⁰ Here I will review some of the difficulties the molecular verificationist account of validity faces.

The first has to do with the applicability of the modified notion of conservative extension to natural language. One might complain that, as the truth-value of natural language sentences can normally not be established definitively, the verificationist version of conservative

¹⁹ This is in Prior 1960, p. 130.

²⁰ In Pinto 1998, chapter 7.

extension inevitably loses the classificatory power of the initial one. Thus, it is highly doubtful whether we could classify ‘tonk’ as verificationally non-conservative if an application of the direct means of verifying ‘q’ could only guarantee that we are fallibly justified in denying ‘q’, while the use of a certain indirect means—via the defeasible direct verification of ‘p’—established that we are justified in asserting ‘q’; the specific use of the two methods might be blamed instead for the apparent quasi-conflict. The point is that a defeasible ground for assertion does not make room for a sufficiently precise concept of lack of conflict between, on the one hand, reasons to assert and, on the other, reasons to deny. One application of a method of verification may give us a reason to assert which a later application of the same method (in a situation where our set of background information has changed) might withdraw.

A second difficulty with the verificationist notion of conservative extension concerns the issue of whether the meaning of a word can be completely determined by an appropriately specified fragment of its respective language. As discussed above, a moderate holist denies this and one could manifest his agreement with him by citing the constant modification of the meaning of natural language terms by new uses of sentences containing it or by the discovery and confirmation of new scientific theories. An example of the latter would be, say, the word ‘dog’ before and after the atomic theory became common currency. Or think of the word ‘funny’ just before and after it started being used to signify an odd person. Examples like these can be multiplied almost indefinitely. In my opinion, they highlight the point that the thesis according to which the meaning of a word is totally fixed by a certain group of logically non-complex sentences containing it is, at the very least, highly controversial. I do not want to proffer a final word on this matter here. But if the holist is right about the irreducibility of the new use of a sentence to the old meanings of its constituent words plus the sentence’s compositional structure, then there will be no room for a constitutive/collateral distinction neither for sentences nor for words. And without such a distinction no sense could be made of the verificationist concept of conservative extension. We must conclude hence that a big question mark still hangs above Dummett’s explanation of validity.

§3.2. Dummett on Fruitfulness

As far as the explanation of the fruitfulness of deduction goes, it must incorporate some sort of epistemic gap between premises and conclusion. Dummett claims that this epistemic gap can only be accounted for by theories of meaning which allow for a distinction between truth and the recognition of truth. According to him, radical verificationism offers the perfect example of an approach which is unable to explain fruitfulness because of its identification of the truth of a statement with its actual verification by whatever means (direct or indirect). The problem is that the explanation of validity will require the preservation of some property of statements which this verificationist claims to be that of actual verification. But if verification of the conclusion reduces to the verification of the premises of an arbitrary valid deduction—as the radical verificationist would have it—then there can be no epistemic gap between them.

Dummett acknowledges that verificationism runs the risk of being unable to explain fruitfulness. He suggests that, in order to avoid the difficulty discussed in the last paragraph, the reasonable verificationist should move some way towards realism, that is, he must replace

in the characterization of truth the implausible notion of actual verification by that of verifiability.²¹ Consider, for example, one of Dummett's partial characterizations of truth:

True statements must comprise, though they are not necessarily confined to, all those which would have been established as true had the relevant observations been made; 'observation' is, as before, not to be taken as mere passive exposure to sense experience but to include physical and mental operations and the discernment of structure (of patterns). In particular, we are able to say that a statement is true, in this sense, whenever we can show how observations that were made *could have* been transformed into ones that would have established it. (Dummett 1991, p. 181)

Truth, in Dummett's view, is to be elucidated counterfactually: a sentence *S* is true if were we in a position to verify other sentences (by observation or by some mental operation like computation or proof), we could thereby transform these verifications into a verification of *S*. In other passages, Dummett identifies this possibility of transformation with a method of verifying *S*.²² Hence, truth for Dummett must be predicated of those sentences concerning which at least a method is known to us for transforming the direct verification of other statements into a direct verification of *S*.²³ The necessary gap between truth and knowledge of truth is preserved, he insists, because a competent speaker who understands *S*—i.e. knows a canonical method of verifying it—may still never have, or could never have, verified *S* directly; besides, if he knows of any indirect means of establishing the truth of *S*, he may have equally never applied it.

Once truth has been described in the above manner, Dummett can then explain fruitfulness in the following way. He will say that knowledge of the truth of a sentence (*S*) always goes beyond knowledge of the truth of the premises *S* is inferred from even when *S*'s truth is established by a direct means. Of course, there are cases of the use of the direct verification procedure where the gap between knowledge of the sentence's truth and knowledge of the truth of the procedure's premise's may be infinitesimal as, say, for the sentence «6 is even and a multiple of 3». Other cases of application of the direct verification procedure are not so trivial and it would be plausible to say that there the mentioned epistemic gap is a bit wider. For instance, in the case of the sentence « $10^{1000} + 1$ is prime». But the really interesting cases, the ones where the gap is much wider according to Dummett, are those where deduction gives us a totally unexpected indirect way of verifying *S*. Examples of these cases abound in mathematics. One could mention Cantor's proof about the cardinality

²¹ We discussed this property of statements in the last section in connection with the verificationist notion of conservative extension.

²² For example, in Dummett 1973b, pp. 314-16.

²³ See, for instance, Dummett 1973b, p. 314. There are, according to Dummett, two reasons why the moderate verificationist notion of truth must have such a convoluted expression. On the one hand, truth cannot be equated with verification by a direct means because this, as our discussion of the radical verificationist shows, would make it impossible to explain the usefulness of deduction. On the other hand, truth must be related to the direct method of verifying statements otherwise one would fall into some version of verificationist holism. This last contention is far from straightforward unless one adds the further assumption that there must be a close connection between meaning and truth. Such a connection was first rendered explicit by Frege and Dummett is prepared to endorse it provided that truth is understood according to the moderate verificationist canon (see, for example, Dummett 1991, chapter 6).

of real numbers or Andrew Wiles' recent proof of Fermat's last theorem. For Dummett, there is a continuum of possible cases covering the whole scale measuring epistemic distance between premises and conclusion; in the most trivial cases such a distance will be extremely small, whereas the most ingenious proofs will be located at the other extreme of the scale.²⁴ The important point is that all cases of the application of deduction involve some epistemic distance, however small, between premises and conclusion. This distance is to be explained in terms of how novel the deductive transition under consideration is; that is, how much it diverges from the trivial cases where the canonical method of verifying the conclusion is what connects deductively premises and conclusion and from the even more trivial cases where such a canonical deduction is quite easily applicable. The distinction between truth and knowledge of truth guarantees that, although truth is automatically transferred from premises to conclusion in a correct deductive transition, there may be a quite pronounced epistemic gap between them.

Euler's proof nicely illustrates Dummett's explanation of fruitfulness. Recall that before the proof was found, there was already a canonical means of determining the truth-value of its conclusion—the sentence «any uninterrupted minimal route through the 7 bridges in Königsberg contains more than 7 bridge-crossings». Euler discovered a new, indirect means for transforming the canonical method of verifying the premise—the sentence «consider any route over the Königsberg's bridges which is uninterrupted, minimal and crosses all the 7 bridges»—into a method of verifying the conclusion. The proof—Euler's non-canonical method—is our warranty that the transition from premise to conclusion is truth-preserving. Moreover, knowledge of the proof together with the verification method for the premise provide us with a new route to the recognition of the truth of the conclusion. Imagine, for example, that all the bridges were numbered and that a device at each bridge detected only that this bridge was crossed but did not keep track of how many times it was crossed. Suppose also a central device which controlled all the bridge cross-detectors and signaled when all the bridges were crossed, and also whether the route was uninterrupted as well as minimal. This seems a good method for verifying the premise (MVP). Now, by Euler's proof, each step in MVP becomes a step in a new method of verifying the conclusion (NMVC). Application of MVP and of Euler's elegant procedure to MVP allows one to establish the truth of the conclusion, eliminating thereby the need to verify the latter by means of its cumbersome direct method. It is precisely this newness that makes Euler's proof fruitful.

So, the fertility of a deductive argument is to be explained, according to Dummett, in terms of how much it, when applied to the direct methods of verifying the premises, departs from the direct method of verifying the conclusion. In the *Logical Basis of Metaphysics*,²⁵ he claims that this explanation agrees with Frege's insight that the novelty of a fruitful deduction lies in the fact that a new pattern—a new procedure for connecting methods of verification, Dummett would add—has just been discerned; a pattern that was already there to be discerned. Now, we have seen that this supposed insight creates a major tension in Frege's philosophy: a tension between that which according to him validity requires—that is, that the conceptual framework we employ be prior to our deductive practice—and that which he rightly saw as essential to the fertility of deduction—that our deductive practice be partly responsible

²⁴ This is corroborated, for example, by Dummett 1973b, p. 297 and Dummett 1991, p. 176.

²⁵ Dummett 1991, pp. 197-199.

for the creation of new concepts. The question is: would Dummett's agreement with the Fregean picture of the fertility of deduction not create a similar tension in his philosophy as well?

§4. Concluding Remarks

In one of his most recent papers on the nature of deduction,²⁶ Dummett explicitly defends an externalist account of deductive consequence. He stresses his agreement with the Fregean picture of the practice of inferring as that of discerning changes in the patterns we impose on reality, where the patterns consist in our judgements about this very reality. The mentioned externalism translates into the idea that the possibility of transformation of one judgement with a certain discerned pattern into another—i.e. the possibility of a given deductive transition—is intrinsic to the judgements in question and not created by our inferring activities (like, for instance, a new proof). The patterns of our judgements are not imposed by us; their multiplicity is already contained in the judgements themselves. Dummett believes such a dose of externalism is necessary if validity is to be accounted for. Otherwise, as he claims, deductive transitions would not be meaning preserving; the judgements on which the patterns must be discerned would themselves change in the course of the transition from premises to conclusion.

Dummett's externalism about deductive inference matches very well with his explanation of validity in terms of conservative extension. Remember that, according to him, the requirement that our logical vocabulary be conservative with respect to verifiability is a necessary condition for valid transitions to be meaning-preserving. But this requirement, as we already know, rules out the possibility that the meaning of sentences and words might be changed by correct deductions. That is, if the explanation of validity is to appeal to conservative extension then the discovery of a new deductive argument cannot alter our conceptual framework. Here, I think, is where Dummett's view diverges from Frege's. For, in order to avoid inconsistency, Dummett must deny that the process of devising and understanding a new proof requires the deployment of new concepts. The conceptual apparatus does not change, he would probably insist; what varies is our command of such a conceptual network.

The problem with this way of accommodating validity and fruitfulness is that the explanation of the former is achieved almost at the expense of rendering the account of the latter very unsatisfactory. For to say that the fertility of deductive inference resides in its often contributing with new methods for verifying sentences whose possibility is nonetheless provided for by our language amounts to admitting only a very mitigated notion of newness: one that rests solely on our cognitive limitations with regard to seeing pre-existing conceptual connections. On Dummett's view, a being not subject to human epistemic constraints who could nonetheless make sense of our deductive practice would not find it useful at all. Put like this, the molecular verificationist account of deduction's usefulness does not look any more promising than the one in terms of the non-triviality of concatenations of trivial inferences. Perhaps Dummett's constraint on a plausible explanation of validity—i.e. that deductions should be meaning-preserving—was too strict. However, if that constraint was relaxed, the molecularist framework would have to be abandoned in favor of a holistic justification of deduction. Dummett, as noted above, is pessimistic about the prospects of such a justification. Many

²⁶ Dummett 1994.

people, including myself, do not share his pessimism. Expounding my reasons would nevertheless deserve another paper.

§5. Bibliography

- Davidson, D. 1986: «A Nice Derangement of Epitaphs» in Lepore, E. 1986: *Truth and Interpretation* (Oxford: Basil Blackwell).
- Dummett, M. 1959: «Wittgenstein's Philosophy of Mathematics» in Dummett, M. 1978.
- Dummett, M. 1973a: *Frege Philosophy of Language*. (London: Duckworth).
- Dummett, M. 1973b: «The Justification of Deduction» in Dummett, M. 1978.
- Dummett, M. 1973c: «The Philosophical Basis of Intuitionistic Logic» in Dummett, M. 1978.
- Dummett, M. 1978: *Truth and Other Enigmas* (London: Duckworth).
- Dummett, M. 1991: *The Logical Basis of Metaphysics* (London: Duckworth).
- Dummett, M. 1994: «Wittgenstein on Necessity: Some Reflections» in Clark, P. & Hale, B. 1994: *Reading Putnam* (Oxford: Blackwell).
- Euler, L. 1736: «Solution d'un Probleme appartenant a la Geometrie de Situation» (French transl.) in Terquem & Gerono 1851: *Nouvelles Annales de Mathematiques* Première Serie (10): 106-19.
- Frege, G. 1950: *The Foundations of Arithmetic*. transl. by J. L. Austin (Oxford: Blackwell).
- Gaifman, H. 1996: «Is the «Bottom-Up» Approach from the Theory of Meaning to Metaphysics Possible?» in *The Journal of Philosophy* (93): 373-407.
- Haack, S. 1976: «The Justification of Deduction» in *Mind* (85): 112-19.
- Haack, S. 1982: «Dummett's Justification of Deduction» in *Mind* (91): 216-39.
- Lormand, E. 1996: «How to be a Meaning Holist» in *The Journal of Philosophy* (93): 51-73.
- Pinto, S. 1998: *Wittgenstein, Meaning and Mathematics*. PhD dissertation (University of London).
- Prior, A. 1960: «The Runabout Inference-Ticket» in *Analysis* (21): 38-9 Also in Strawson P. F. 1967.
- Quine, W. 1951: «Two Dogmas of Empiricism» Quine, W. 1953: *From a Logical Point of View* (Cambridge: Harvard University Press).
- Strawson, P. F. 1967: *Philosophical Logic* (Oxford: Oxford University Press).
- Tennant, N. 1987: *Anti-Realism and Logic* (Oxford: Clarendon Press).
- Wittgenstein, L. 1922: *Tractatus Logico-Philosophicus* (London: Routledge).

Silvio Pinto

Universidad Autónoma del Estado de Morelos

<silvio@filosoficas.unam.mx>

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 48-51

New Systemic Hypothesis of Ageing

Copyright © by SORITES and Alexey Kolomiytsev

NEW SYSTEMIC HYPOTHESIS OF AGEING

Alexey Kolomiytsev

Great advances in biological research in recent years have resulted in a clearer understanding of biological processes at subcellular and molecular levels in physiological and pathological conditions. But today we are still far from resolving the problem of life duration. The increase in life duration — partly as a result of developments in medicine — is reaching a plateau. Further life lengthening may be possible by altering cellular processes. Recent research into apoptosis and other problems of the biology of ageing have revealed some molecular mechanisms for cell life control. However, these studies were done by cell culture techniques, and there is a great difference between genetic program realization in vitro and what happens in real life.

Some functional features of complex multicellular organisms may not be explained by subcellular processes alone. These organisms function as complex systems, and the life of cells is determined both by molecular processes and system organization. This approach may represent the general idea about any living organism functioning and processes that influence the duration of life at all and persistence of pathologic conditions.

Our new hypothesis does not contradict either evolutionary or molecular theories of ageing. This is just an attempt to understand how ageing developed in complex organisms.

Multicellular organisms function as complex systems, and the life of cells is determined both by molecular processes and system organization. The developmental process is determined by kinetic curve of population growth (Khokhlov AN., 1987, Varfolomeev SD., 1990), which is typical for every cell association (cell culture, colony of microorganisms, etc.) (Erbe W., et al., 1977, Foa P., et al., 1982, Bremer H., et al., 1983). It includes phases of induction, exponential growth, linear growth, deceleration, a stationary and atrophy phase.

The new hypothesis considers every multicellular organism as a system that consists of various cellular associations in symbiotic interaction. One of these associations dominates and determines the developmental kinetics of the whole organism.

The nervous system is the dominating cellular association in animals and human beings. The duration of its development is restricted by the capability of neurons to regenerate in adult organisms. This may determine the duration of life of the organism as a whole. So the organism is hypothetically presented as a «neuronal cell culture» that develops on the «medium» (the other tissues).

Although there is no nervous system in plants, this rule can also be observed in their organization. For instance, apical meristems in plants play the role of organogenesis units.

Because organs are produced continuously throughout the life cycle, the apical meristems maintain a pluripotent stem cell population (Bowman JL., 2000). Thus, apical meristems may play the dominating role in plant maintenance, development and life cycle.

Animal organisms (more exactly this applies to mobile animals and humans) may be hypothetically presented as a «neuronal cell culture» that develops on the «medium» — the other functional units of organism.

Although there are certain reserves of neural stem cells in the human CNS, they are unable to generate new nerve cells in any useful amounts. This is an evolutionarily conditioned feature (Aubert I., et al., 1995, Olson L., 1997).

New stereological techniques (Cotter D., et al, 1999, Kubinova L., et al, 1999, West MJ., 1999) have failed to confirm earlier findings regarding age-associated neural loss (Regeur L., 2000) but there is the evidence of focal neural death and synaptic or receptor loss.

Recent researches in neurobiology revealed the ability of human neural stem cells «(1) to differentiate into cells of all neural lineages (i.e., neurons — ideally to multiple subtypes, oligodendroglia, astroglia) in multiple regional and developmental contexts; (2) to self-renew (to give rise to new NSCs with similar potential); and (3) to populate developing and/or degenerating CNS regions» (Flax JD., et al. 1998, Brustle O., et al., 1998). The same features are essential in animal neural stem cells (Temple S., et al, 1999).

These features could be useful not only in treating neurodegenerative diseases (Brevig T., et al., 2000, Svendsen CN., et al., 1999) but also in a wider range of pathologic processes.

Constant alterations in neural tissue may lead to persistence of a number of pathologic processes and result in a decrease of life duration. The restoration of neural cells and their connections may be useful to modify the features of various chronic diseases and even increase life span.

New strategies including stem cells transplantation could be used for the stimulation of cell renewal processes in central and peripheral nervous system as a whole. This may result in continuous self-renovation of the whole organism as a complex system.

Moreover, epidermal growth factor (EGF) and transforming growth factor (TGF) — a member of the epidermal growth factor — are essential components for a stable growth of neural cells (Junier MP., 2000). This may mean that these growth factors are necessary for both epithelial and neural tissues as ectoderm derivatives, and in one's turn the growth factors application could be one of possible ways to increase stability of neural cells of human organism.

Our hypothesis does not violate any of the existing theories of ageing. The theories of life span gene regulation may also be explained by means of our theory. Even cell division limit (Hayflick L., 1998, 2000) may be considered a compensatory feature that arose as a result of cellular development in complex systems where nutritional competition occurs as well as the destruction product accumulation.

REFERNCES

01. Aubert, I, Ridet, JL, Gage FH.(1995) Regeneration in the adult mammalian CNS: guided by development. *Curr Opin Neurobiol.* 5: 625-635.
02. Bowman, JL, Eshed, Y. (2000) Formation and maintenance of the shoot apical meristem. *Trends in Plant Science*, 5:3: 110-115.
03. Bremer, H, Chuang L.(1983) Cell division in Escherichia coli after changes in the velocity of DNA replication. *J Theor Biol.* May 7 102:1 101-20.
04. Brevig, T, Holgersson, J, Widner, H. (2000) Xenotransplantation for CNS repair: immunological barriers and strategies to overcome them. *Trends in Neurosciences.* 23:8: 337-344.
05. Brustle, O, et al. (1998) Chimeric brains generated by intraventricular transplantation of fetal human brain cells into embryonic rats. *Nature Biohechnology.* 16: 1040-1044.
06. Cotter, D, Miszkiel, K, Al-Sarraj, S, Wilkinson, ID, Paley, M, Harrison, MJ, Hall-Craggs, MA, Everall, IP. (1999) *The assessment of postmortem brain volume: a comparison of stereological and planimetric methodologies.* *Neuroradiology.* Jul 41:7 493-6.
07. Erbe, W, Linden, WA, Reddy, SB, Zywiets, F., Baisch, H. (1977) Studies on the population kinetics of the Walker carcinoma by autoradiography and pulse cytophotometry. *Cell Tissue Kinet.* May 10:3 269-77.
08. Flax, JD., Aurora, S., Yang, C., et al. (1998) Engraftable human neural stem cells, respond to developmental cues, replace neurons, and express foreign genes. *Nature Biotechnology.* 16: 1033-1039.
09. Foa, P, Maiolo, AT, Lombardi, L, Toivonen, H., Rytomaa, T, Polli, EE. (1982) Growth pattern of the human promyelocytic leukaemia cell line HL60. *Cell Tissue Kinet.* Jul 15:4 399-404.
10. Hayflick, L. (1998) How and why we age. *Exp Gerontol.* Nov-Dec 33:7-8: 639-53.
11. Hayflick, L. (2000) The illusion of cell immortality. *Br J Cancer.* Oct 83:7: 841-6.
12. Junier, MP. (2000) What role(s) for TGF in the central nervous system? *Progress in Neurobiology.* 62:5: 443-473.
13. Khokhlov, AN, Golovina, ME, Chirkova, EIu, Nadzharian, TL. (1987) Analysis of the kinetic growth patterns of cultured cells. III. The effect of inoculation density, of a geroprotector-antioxidant and of stationary-phase aging. *Tsitologiiia.* Mar. 29:3.353-7.
14. Kubinova, L, Janacek, J, Guilak, F, Opatrny, Z. (1999) Comparison of several digital and stereological methods for estimating surface area and volume of cells studied by confocal microscopy. *Cytometry.* Jun 1 36:2.
15. Olson, L. (1997) Regeneration in the adult central nervous system: Experimental repair strategies. *Nature Medicine.* 3: 1329-1335.
16. Regeur, L. (2000) Increasing loss of brain tissue with increasing dementia: a stereological study of post-mortem brains from elderly females. *Eur J Neurol.* Jan 7:1 47-54.

17. Svendsen, CN, Smith, AG. (1999) New prospects for human stem-cell therapy in the nervous system. *Trends in Neurosciences*. 22:8:357-364.
18. Temple, S, Alvarez-Buylla, A. (1999) Stem cells in the adult mammalian central nervous system. *Current Opinion in Neurobiology*. 9:135-141.
19. Varfolomeev, SD, Kaluzni, SV. *Biotechnologia*. Moscow, 1990. 296p. (in Russian).
20. West, MJ. (1999) Stereological methods for estimating the total number of neurons and synapses: issues of precision and bias. *Trends Neurosci*. Feb 22:2: 51-61.

Alexey Kolomiitsev

Department of Pathology. Rostov-na-Donu Medical University

Address: Sverdlova 97-5

Taganrog, Rostovskaya obl.

347935 Russia

tel. 007-86344-66109, fax. 007-86344-44724.

E-mail: <betrg@altavista.com> & <betrg@pbox.ttn.ru>

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 52-71

Truth Conditions Without Interpretation

Copyright © by SORITES and John Collins

TRUTH CONDITIONS WITHOUT INTERPRETATION

John Collins

1: Introduction

Donald Davidson has long argued that the concept of meaning should be seen through the lens of a Tarski-style truth theory format. A theory answering to such a format will be successful as a theory of meaning for a language *L* if it offers an informative interpretation of *L*-speakers, where such an interpretation affords content attribution to both speech acts and mental states. My brief here is to separate these two components; the optic provided by the Tarski format will be commended but the stated criterion of success will be rejected. Thus, I shall argue that Tarski's format, once suitably re-jigged, is indeed apt to depict a speaker's core semantic competence (§2). This stance will be seen to run counter to a prevailing deflationism about truth, which sees the concept as (in some sense) explanatorily antecedent to that of meaning. Davidson tends to challenge deflationism on the grounds that it gets truth wrong precisely because, under its construal, the concept is no longer serviceable in a theory of meaning. This, I think, simply initiates a stand-off. My tack will be more circumlocutionary. I shall suggest that the Tarski format offers an *optimal* account of compositionality, while it is genuinely moot whether the kind of non-truth involving accounts of meaning the deflationist is obliged to deal with can so much as accommodate compositionality. Such an approach, of course, is non-demonstrative (and none the worse for that); still, on the assumption that a theory of meaning must account for compositionality, we have a *prima facie* case for truth being the core concept of semantic competence (§3). This argument, for sure, is Davidsonian in spirit, where I more fundamentally depart from Davidson is on his understanding of the aim of a theory of meaning.

It will be argued that there is no a priori link between truth and translation or interpretation of the kind Davidson attempts to derive from Tarski's work. Instead, it will be claimed that a truth theory's theorems do their work precisely because they are trivial. This stance will be seen to be the natural consequence of a cognitive conception of a theory of meaning which stands against Davidson's interpretive social view of language (§4). This cognitive approach views a theory of meaning as a component of a naturalistic framework for the investigation of human cognitive capacity. Now it might seem that we have extirpated ourselves from a stand-off about truth — deflationism vs. the truth theoretic conception of meaning — only to sink ourselves into another about meaning — interpretation vs. naturalism. It will be argued, though, that a theory of meaning cannot be screened off from empirical constraints deriving from psychology and linguistics, for the very characteristics of a truth theory which make it initially attractive to Davidson and others (its treatment of compositionality, unboundedness, etc.) are themselves such constraints. A recognition of this parity decisively militates for the cognitive conception over the interpretive social one (§5).

2: From Tarski to Truth

Why are truth conditions fit to articulate meanings? Because the disquotability of ‘is true’ shows that the conditions under which a sentence is true are the very conditions which express the content of the sentence. Schematically:

(TM) X means that $p \rightarrow X$ is true $\leftrightarrow p$,

where the substitutions of ‘ X ’ are quotation-names of sentences and those for ‘ p ’ are the sentences stripped of their quotation marks. If we are in a position to determine each instance of the antecedent in the schema, then, trivially, truth conditions are expressive of meaning.¹ Such a procedure, however, hardly counts as the explication of meaning at the hands of truth. Meaning here is a constant relative to truth. Davidson’s (1984) seminal idea is that we can reverse this relation by placing the right constraints on a theory which issues in the instances of the consequent of (TM):

(T) X is true iff p ,

such that the substitutions of ‘ p ’ interpret the named sentences which substitute ‘ X ’ without the trivialising benefit of an appeal to notions of *meaning*, *translation*, etc. In other words, a truth theory may serve as a theory of meaning because the procedure by which we would determine the truth conditions of a speaker’s utterances is one which would take us from a description of utterances to what the utterances *say*, to a position where we can attribute content to the speaker and so understand him as a rational creature with the same basic beliefs and desires as ourselves. I shall argue that this conception of a theory of meaning as essentially a theory of interpretation is mistaken. Showing this, however, does not involve dispensing with the idea that a speaker’s semantic competence amounts to knowledge of truth conditions. I shall turn to Davidson’s thoughts on interpretation in §4. Before that I shall discharge my first burden by offering what I think is a strong argument for the truth conditional conception independent of Davidson’s understanding of interpretation.

Davidson’s (1984) central idea is that Tarski’s (1956) format for the explicit introduction of truth predicates may double as a format for truth theories that are expressive of meaning. Tarski’s achievement was to show that, for formal languages that do not contain their own truth predicates, a predicate *Tr* may be defined that is co-extensional with our intuitive notion of truth over the given objectlanguage. That co-extensionality holds is not part of the definition, nor, of course, is it proved elsewhere; rather, it is enshrined in a *material adequacy condition* on the definition, *Convention T*, which constrains the metatheory, of which the definition is a part, to entail an equivalence answering to (T), for each sentence of the objectlanguage, with *Tr* in place of ‘is true’. The instances of (T) are theorems of the metatheory: their left-hand sides are metalinguistic names (structural descriptions) of objectlanguage sentences, and their right-hand sides are metalinguistic translations of the sentences described to their left.

The real interest Tarski has for a theory of meaning is his method of recursively defining a relation — *satisfaction* — from which a truth predicate may be explicitly introduced as a special case. We can think of satisfaction as a generalised reference function SAT that maps infinite sequences of objects from the universe of discourse onto each wff of the object-

¹ Throughout this paper I shall intend *first* or *literal* meaning by ‘meaning’, i.e., that notion which is constrained by (TM).

language L such that $\text{SAT}(\phi)$, for any wff $\phi \in L$, is computable from the values of SAT taking the primitive wffs that feature in ϕ as arguments. SAT defined for the primitive expressions of L constitutes the *base* of the definition. In this way the definition of SAT mirrors the recursive definition of ‘wff of L ’ which is also part of the metatheory. SAT, of course, can be *explicitly* defined, given some restrictions on the objectlanguage, either set theoretically or via second-order quantification. Being so defined, $Tr +$ metatheory is a *conservative extension* of the metatheory, i.e., the deductive strength of the metatheory is unchanged. The metatheory in general contains conceptual resources no greater than those expressed in L , save for the set theory, logic, and theory of syntax of L required for the definition of SAT. The eliminability of Tr at the hands of such resources shows that the metatheory is free from paradox, or at least the definition introduces no paradox.

Equally, however, eschewing explicit definitions, and so departing from Tarski’s aims, we can give the clauses of the definition an axiomatic status. The recursion is still effected just as it is when we let ‘ $0 + a = a$ ’ and ‘ $a + \text{succ}(b) = \text{succ}(a + b)$ ’ tell us the values for ‘+’ for any argument pair $\langle a, b \rangle$ without bothering to define explicitly the set of pairs $\langle \langle a, b \rangle, c \rangle$ thereof determined. Even if we lacked the techniques to carry the definition through, the postulates would still obviously give us the right results. Further, we can replace the notion of SAT in the axioms with that of *semantic value* expressed by the primitive 2-place predicate ‘ $\mathbf{v}(x, y)$ ’.

Unlike SAT, \mathbf{v} assigns objects to expressions that are specifically appropriate to them. Singular terms are assigned individual objects and n -placed predicates are assigned n -tuples. These axioms have the same form as SAT clauses; in particular, the conditions for x to be a semantic value of y are expressed by conceptual resources no greater than those of L as translated into the metalanguage. (As with Tarski’s explicit definitions, the rest of the metalanguage is simply to expedite the construction.) Thus, where ‘NN’ and ‘ $F(x_1, \dots, x_n)$ ’ are proxy for singular terms and n -place predicates respectively, the axioms will have the following form:

$$(A1) \mathbf{v}(\sigma, \text{‘NN’}) \leftrightarrow \sigma = \text{NN}$$

$$(A2) \mathbf{v}(\langle \sigma_1, \dots, \sigma_n \rangle, \text{‘}F(x_1, \dots, x_n)\text{’}) \leftrightarrow F(\sigma_1, \dots, \sigma_n)$$

Since our relation \mathbf{v} pairs the appropriate object with an expression, when it comes to sentences (as defined by our recursive definition of wff) we go the whole hog and introduce a truth predicate as a primitive symbol. Again, following the syntax of L , **true**, being the value of a sentence S , can be recursively computed from the axioms such that, with details elided, (T1) holds:

$$(T1) \mathbf{v}(\text{true}, \text{‘NN is F’}) \leftrightarrow (\exists \sigma)(\sigma = \text{NN} \ \& \ \sigma \text{ is F}) \leftrightarrow (\text{NN} = \text{NN} \ \& \ \text{NN is F}) \leftrightarrow \text{NN is F.}$$

The derivations are as cumbersome as they simple. Assuming axioms for each expression of L (including quantifiers and connectives), we may make a similar derivation for each of the infinite sentences of L . Our theory, therefore, satisfies a modified form of Convention T, as is patent from the above example.

3: Truth and Compositionality

Davidson (1990a, 1997a) holds that that the kind of re-jigging exhibited above is possible shows that the Tarskian definitional format reveals truth to be constitutively linked to the notions of reference and satisfaction. Such is what Davidson means by the *structure of truth*:

the truth relevant combination of primitive expressions that enter into the determination of truth conditions. For Davidson, such revelation of compositionality is non-negotiable for a theory of meaning; that a truth theory, therefore, turns the trick of specifying meaning through derivations that reveal compositional structure is not to be easily relinquished. Now Davidson believes that this spells insuperable difficulties for the common deflationary conception of truth.

Deflationism is a mixed bag; for our purposes, let the notion refer to the thesis that there is no more to an adequate, complete account of truth than the instances of (T), or some generalisation thereof. If this thesis is right, then it seems that truth cannot play the role Davidson envisages, for the instances of (T) can be viewed as constitutive of an agent's understanding of truth only if the agent understands the sentences which go into the schema. Otherwise, while an agent might well recognise the instances (*if* homophonic) as true, he would not understand them to the extent that he could not 'disquote'. In short, for deflationism, TM is basic, not (T) itself, and so some well behaved notion of meaning or content is assumed antecedent to truth conditions (Horwich (1990, 1998), Field (1994), and Soames (1999)).

As Horwich (1999) rightly complains against Davidson (1997a), it is pointless to challenge the deflationary reading of truth on the grounds that it makes the truth unfit to serve as the core concept of a theory of meaning. It is a tenant of deflationism that this is indeed a sound consequence of the thesis (e.g., Field (1994) and Horwich (1998)). The burden thereby acquired, though, is to account for meaning in a non-truth-involving way. Davidson thinks the attempt forlorn. One should not think that this so, however, because Tarski has shown us that truth is set up to play an interpretive role within a theory of meaning. This Tarski did not do (see §4). Equally, one should not think that Tarski showed that truth essentially goes beyond disquotation due to its conceptual entwinement with compositional structure in the shape of satisfaction.² Davidson, I think, tends to peddle these two thoughts.³

² See, e.g., Etchemendy (1988). Davidson's (1990a, §1) response to Etchemendy is, in essence, that we shall not necessarily fall into inconsistency if we forgo the explicit definition and read a definition's clauses as substantive axioms. But, if Tarski is to be read as showing us that truth is inherently compositional, then it can only be via satisfaction, yet truth and satisfaction are related in Tarski's method only via explicit introductions of truth predicates in terms of satisfaction definitia. *Tarski did not recursively define truth; he recursively defined satisfaction, from which a truth predicate may be explicitly introduced.* If we introduce a primitive truth predicate to occur in axioms (as presented in §2), then, clearly, one cannot say that Tarski has shown us how truth is essentially tied to the interpretation of sub-sentential structure, for, by such a method, truth is *not* defined in terms of satisfaction. As for explicit definitions, as Davidson well recognises, they are not expressive of compositionality: they amount to conditions on membership of a set TRUE. Suffice it to say, none of this is inimical to the method advertised in §2.

³ There is an interesting contrast here between early and later Davidson. The early papers look upon Tarski's work as offering *one* way to get at compositionality (see especially Davidson (1984, p. 61)). That is, there is no direct *conceptual* relation between Convention T and compositionality. This thought is clearly correct: Convention T is a criterion of adequacy that does not imply *how* a theory or definition should satisfy it. Likewise, a theory of truth, on Davidson's construal, should issue in interpretive instances of (T), but this demand does not in itself tell us just *how* this should be achieved. In his more recent writings, however, Davidson appears to argue that the *structure of truth* is

He has at his disposal, however, a much stronger argument against deflationism that, to my knowledge, he has never made explicit. The value of this argument is that it does not depend upon anything Tarski is supposed to have demonstrated, nor upon any judgement about what is essential to the concept truth. It does, however, trade upon the ingenuity of Tarski's notion of satisfaction. Explicitly, though, the argument depends merely on the idea that a theory of meaning must account for the compositionality of complex 'meanings'.

Let us assume that any theory of meaning must show how the meaning of a complex expression is composed from the meaning of its parts and their structural arrangement. This desideratum may be questioned (e.g., Schiffer (1987) and Jackendoff (1997)). Let me be clear: a theory of meaning is not, in the first instance, an account or analysis of the property of *having meaning*. No serious investigation is interested in properties as such; rather, the aim is always to explain the phenomena at hand, and properties are identified in terms of their ineliminably featuring in such explanations. A theory of meaning, therefore, is not beholden to preserve the integrity of our intuitive concept. Compositionality, though, is a datum of most of the constructions with which we are competent. Whether or not, then, a given language does, in general, have a compositional semantics, the extent to which it is compositional is the extent to which an account of meaning for the language is constrained to specify what speaker/hearers know such that the exhibited structured competence is explained. I say this prescriptively, but its value will emerge as we go along.

We know that a truth theory of the kind sketched in §2 specifies knowledge which, if possessed, would explain compositionality. We do not know, as indicated, how to generate truth conditions for all constructions; e.g., perhaps subjunctives, non-intersective modifications, modal adverbs, etc. Still, a truth theory cleaves to compositionality because *if* we know how to assign semantic values to lexical primitives and we know how to specify the semantic value of a construction type *C*, then we can show that the truth conditions of a token of *C* are compositionally determined through a derivation of the kind given in §2. This is far from a trivial achievement, as we shall see. *Ceteris paribus*, therefore, it is a sound hypothesis that truth is indeed the core concept of an adequate account of meaning, i.e., speakers know a truth theory for their languages. Now deflationism explicates truth on the basis of an antecedently determinate notion of meaning; patently, truth conditions cannot enter into this notion. Still, whatever notion is provided, it must be such as to account for compositionality; this is non-negotiable. In effect, then, deflationism is betting on a non-truth theoretic explanation of compositionality, which would show that *ceteris* is not *paribus* in favour of the truth theory. But is there such a thing as a non-truth theoretic explanation of compositionality?

There are no shortage of theories that seek to account for meaning apart from truth conditions (semantic values more generally): *use* theories, *conceptual role* theories, *paradigm* theories, *verificationist-recognitional* theories, etc. Jerry Fodor and others have for a number of years been arguing that none of these alternatives, notwithstanding any other virtues they

conceptually necessary to thought or meaning (especially see Davidson (1990a, p. 296; 2000, p. 72)). I do think (see below) that we haven't got a good idea how to account for compositionality apart from a satisfaction-like relation; even so, if we were offered a non-truth theoretic explanation of compositionality that was not obviously mistaken, it would be dogmatic to say that there *must* be something wrong with it because it failed to deal in reference and truth.

might possess, come close to accounting for compositionality.⁴ The core problem is easily explained. Consider $[_{NP} \text{ ADJ } N']$ constructions where the modification is intersective, such as in *pet fish*, *male nurse*, *red ball*, etc. Compositionality tells us that the meaning of, say, *pet fish*, must be a function of the meaning of *pet* and *fish*, and that, moreover, this meaning must determine that pet fish are those things which are both pets and fish. Such, after all, is what it means to be an intersective modification; schematically:

$$(1) \mathbf{v}(x, [_{NP} \text{ ADJ } N']) \leftrightarrow x \in \{\text{ADJ} \cap N'\}.$$

The problem with the alternatives is that they do not appear to have the required consequences. The properties they cite as constitutive of primitive meanings (lexical or conceptual) are often variant over changes in context. For example, a paradigm pet is not a fish, it is, say, a dog, but the concept of dog does not enter into the meaning of *pet fish* (similarly for recognitional theories: I can recognise pets and fish when I see them, but it does not follow that I can recognise a pet fish). Likewise, I might be *canonically committed* to assert that fish are good to eat, but my taste surely does not stretch to pet fish. Suffice it to say that the debate continues. The important moral for my purposes is that it is genuinely moot whether anything other than a truth theory can explain even the most elementary constructions. Such is why I said that its achievement in this regard is far from trivial. Compare, for instance, the problems raised above with the limpidity of the derivation a truth theory allows for:

$$(2) \mathbf{v}(x, \text{'pet fish'}) \leftrightarrow (\mathbf{v}(x, \text{'pet'}) \ \& \ \mathbf{v}(x, \text{'fish'})) \leftrightarrow (x \text{ is a pet} \ \& \ x \text{ is a fish}) \leftrightarrow \text{is a pet fish}.$$

The important feature of the concept of semantic value is its *uniformity*, i.e., the value of any expression e is the value which e contributes to the value of any complex in which it occurs. In this sense we may say that the concept of semantic value (or satisfaction) is inherently structured: the condition under which a complex expression has the value V just is the condition that holds for the composition of the values of its parts. Thus, to give the semantic value of an expression e is to specify that property of e which contributes to the truth conditions of the sentences in which it occurs. In the other direction, the truth conditions for any sentence may be depicted as systematically arising from the contributions of its parts. This is why a truth theory offers an optimal explanation of compositionality. As far as we can tell, *only* truth theoretic notions (or satisfaction) have this inherent structure; *use*, *conceptual role*, and the rest appear to lack it.

It is tempting to conclude that the concept of truth is essential to a theory of meaning, with, consequently, deflationism indirectly refuted. This is too strong; there is no need to gainsay the future ingenuity of philosophers and linguists. What is clear, however, is that truth has an optimality *vis-à-vis* compositionality; that is, a truth theory offers us conditions that *do* compose. Horwich (1998) has recently challenged this conclusion; he claims that compositionality is trivial because it is consistent with whatever we take primitive meanings to be; a truth theory does indeed cleave to compositionality, but so does (more or less) every other theory of which one can conceive. No value, therefore, is accrued by the truth theory's success.

⁴ The main texts are Fodor and Lepore (1991; 1992) and Fodor (1998a; 1998b, Chps. 4 and 5). Fodor, it should be said, is principally concerned with mental content rather than linguistic meaning, although, clearly, if we cannot account for the compositionality of *thought* with theory T , then we should not expect T to work with *meaning*.

Horwich's idea is this. Consider a sequence of words $\langle w_1, w_2, \dots, w_n \rangle$ that comprise a sentence S . Let $\langle p_1, p_2, \dots, p_n \rangle$ be the corresponding sequence of the meaning properties of the words of S , where these are whatever we like. Assume that S has the structure $[_S [_{NP} N] [_{VP} [V] [_{NP} ADJ N']]]$ (e.g., *Mary likes male nurses*), where this expresses some combinatorial principle P with the meaning property p_{n+1} . Now Horwich contends that there is nothing to compositionality other than the idea that to understand a complex expression is just to understand its elements and the way they are combined, an idea which is fully captured by a *construction property*. In the case of S , this amounts to (3):

(3) S means $\mathbf{S} = S$ results from the application of P , with meaning constituting property p_{n+1} , to $\langle w_1, w_2, \dots, w_n \rangle$, with the meanings $\langle p_1, p_2, \dots, p_n \rangle$.

Effectively, a construction property (as given on the right flank) is simply a statement of the value of the function expressed by the structure of a given sentence taking the sequence of primitive meanings of the sentence's words as argument:

(4) $p_{n+1}(\langle p_1, p_2, \dots, p_n \rangle) = \mathbf{S}$

The important thing to note here, is that we have let the meanings be anything we please, and we have still managed to specify a construction property. Trivially, we can do this for any complex. Since the construction property constitutes what the complex means, we have thus showed how compositionality can be satisfied without appeal to any *particular* theory as to the nature of meanings. A truth theory, therefore, is no more privileged than a use theory, say, in satisfying compositionality. Thus, deflationism can, *ceteris paribus*, appeal to a use theory, say, to account for meaning and then give a deflationary characterisation of truth upon this basis, which is just what Horwich proposes.

Horwich takes it to be a virtue of his approach that it shuns any uniform relation, such as satisfaction, that may relate parts and whole via their realising the same property or being a relatum of the same relation. By Horwich's lights, the meaning constituting property of *male nurse*, say, is not the same kind of property that constitutes the meanings of *male* and *nurse*; for the first meaning here is given by a construction property, which is available independently of whatever is understood to constitute the primitive meanings. This rejection of uniformity certainly allows Horwich to claim that, while the meaning of *nurse* might well be some set of paradigm nurse features (including, presumably, *being female*), the meaning of *male nurse* need not be a paradigm at all. This goes no way, however, to show that uniformity is not required to capture compositionality.⁵ It does not even answer the complaints raised above, for how the primitive meanings are expressed in the complex meanings remains obscure. What, on Horwich's account, do paradigm nurses, say, have to do with the meaning of *male nurse*? It seems that Horwich is saying that one could understand *male nurse* without having a clue about paradigm nurses, for what constitutes the meaning of the complex does not

⁵ Horwich (1998, pp.22-7) does identify what he terms the «constitution fallacy»: the idea that any analysis of a complex property must preserve its componential structure in its analysis. Thus, Horwich's claim is that the uniform notion of semantic value simply issues from the dogma that there must be some property to realise the uniform relation in each fact of the form ' x means y '. I suppose this is a fallacy, but it has nothing to do with the issue at hand, for (i) a truth theory does not analyse meaning properties and (ii) the uniformity of semantic value is not based upon any general principle of constitution, fallacious or otherwise, but on what appears to be demanded by the phenomenon at hand.

express what constitutes the meaning of its parts, the parts just need to have some meaning or other. The problem here is, I think, a species of a problem Russell faced.

The *unity problem* bedevilled Russell from the time of *Principles* (1903) up to the period of logical atomism (1918). Simply put, Russell's problem was how to understand relations (transitive verbs, say) as simultaneously constituents of propositions and relations between the other elements of the proposition, or, in Russell's (1903, §54) terms, how to distinguish between «a relation in itself and a relation actually relating».⁶ For example, enumerating the meanings of *John*, *loves*, and *Mary* will not issue in a structure that will articulate *that* John loves Mary; that is, the enumeration will not constitute what a speaker means who utters the sentence, for the enumeration does not relate John to Mary in *any* way. It will not do to add a further relation — a 'relating relation' — to the enumeration; we *already* have that relation: if we need a further relation, then we would need yet another one to express *its* 'relatingness', *ad infinitum*. What appears to be required is a constituting structure outside of the proposition, but this threatens to destroy the very notion of knowledge of meaning, i.e., to conflate the act of *saying* something with what is *said*. Russell did not resolve this dilemma. If, however, we eschew enumerations in the first place and depict a speaker's understanding as being constituted by knowledge of truth conditions, then the problem disappears, for to know truth conditions is *ipso facto* to know the semantic contributions the parts make to the whole. We do not have to specify this contribution as a further element to the knowledge since the operative notion of semantic value is uniform across constituents and host complexes: to know the conditions under which the parts have their semantic values just is to know the conditions under which the whole is true.⁷ Russell's problem appears to arise only when one attempts to specify lexical meanings independently of their structural options or, perhaps better, argument structure or logical form. As Russell himself put it: «when analysis has destroyed the unity, no enumeration of constituents will restore the proposition» (Russell (1903, §54)).

Now Horwich's theory precisely destroys unity in that individual words are depicted as having meaning constituting properties independently of the kind of arrangements they can enter into with other words. Horwich's appeal to construction properties and schemata arrives too late, as Russell would have understood. A construction property essentially tells us that if we put meanings together like *this*, then we get this other meaning. Well, sure; but that is a statement *about* meanings in themselves, as it were, it is not a statement that depicts

⁶ The problem was especially vivid for Russell because his propositions (in 1903, at least) were made up out of particulars and universals, the very stuff our words are about, not mental representations of that stuff. Thus, realism vs. idealism turned on whether we can have an understanding of a thing independent of grasping the set of relations into which it enters. If not, then relations are *internal*, and we are landed with the One, as it were.

⁷ Davidson takes this kind of constituency to entail holism. This thought is correct on an interpretive understanding of the aim of a theory of meaning, for what is to be interpreted are sentences not words, words get interpreted simply in terms of their sentential role. I do not, however, share Davidson's assumption. If a theory of meaning is something a speaker genuinely knows, then sentential structure may be understood as projection from the lexical entries as represented in the speaker's mind. See §5.

meanings ‘relatingly’ in their joint articulation.⁸ The logical form of the sentence — how the constituents are related one to another to engender the complex meaning — remains outside of the content a speaker articulates when using the sentence. But the problem of compositionality is precisely to account for the form of meaning: we want to know *what* a sentence means in terms of its parts; we do not need to be told that it means whatever it does because it is made up out of *these* parts rather than *those*. That is utterly trivial.

It bears emphasis that compositionality on the truth theoretic understanding is not explained because SAT (or its cognates) maps n -tuples onto lexical items; rather, it is because a truth theory states the conditions under which the assignment holds in complexes, where the content of such a statement may be attributed to a speaker in such a way as to explain the speaker’s performance. The inherent structure revealingness of satisfaction would be lost if we thought of a truth theory as simply a mapping of objects onto words, for no such mapping can constitute what a speaker knows who knows a language. The equivalence relations in (2), for example, do not state that *pet* contributes the set of pets to the meaning of *pet fish*. This would be of no help at all. Rather, (2) states the conditions under which *pet* would be satisfied in such a way that those conditions are re-articulated in the conditions for the satisfaction of *pet fish*. On the other hand, it just makes no sense whatsoever to attribute construction properties to a speaker. To repeat the moral from above: as far as we know, only truth theoretic notions admit to such an explanation, and only such an explanation appears fit to tell us *what* the meaning of a complex is on the basis of the meaning of its parts.

If these thoughts are anywhere near correct, the deflationist will have the most difficult task to find a non-truth-involving theory of meaning that explains compositionality. And since compositionality is a non-negotiable constraint on meaning, the deflationist will not be in a position to explain meaning *per se*. If we want a theory of meaning, therefore, we have to acknowledge, it seems, that truth will be a core concept of it, which is just to deny that truth can be explained by an antecedent notion of meaning, which in turn is one way of saying, albeit indirectly, that deflationism about truth is false. As I remarked earlier, I cannot find this argument stated explicitly in Davidson’s work. Nevertheless, it is clearly a ‘Davidsonian’ argument and should be where the truth theorist makes his stand against deflationism. This result has a direct bearing, I think, on how we should understand the relation between truth and translation. To this I now turn.

4: Truth and Translation

From his initial statement of the idea that a truth theory might serve as a theory of meaning to his latest writings on truth (e.g., Davidson (2000)), Davidson has claimed that truth is an apt concept to capture meaning because Tarski showed that it essentially involves a notion of *translation*. The initial idea was that where Tarski keeps translation constant (in homophonic definitions such constancy is simply morphological identity) to get at truth, one may reverse the relation by assuming that a stable attitude towards truth can be identified in

⁸ This kind of complaint was made by Davidson (1984, pp.17-8) against Frege’s notion of unsaturatedness. Higginbotham (1999) deploys the complaint against Horwich, who sees his schemata solution as being essentially Fregean. I was, however, unaware of the latter paper at the time of formulating my current ideas. Besides which, neither Davidson nor Higginbotham trace the issue back to Russell, where it belongs. The points I am making were influenced by Sainsbury (1997a), who relates Russell to Davidson.

speakers which may pivot a translation/interpretation of *that* which they hold true. This initial idea has undergone many changes in the intervening years, changes I shall not discuss. What I want to focus on is the core idea that truth and translation are conceptually entwined, such being a chief reason why a truth theory is the correct approach to meaning. I think this is a superfluous idea; what is important, I submit, is that a truth theory states knowledge which may constitute a speaker's linguistic competence. Whether a speaker possesses such knowledge is, ultimately, to do with the computational structure of his brain rather than how well we can get on with him; the form of interpretation Davidson takes to be constitutive is, I shall suggest, at best marginal data as to such structure.

Davidson (1990a) puts his point by saying that what is missing from Tarski is the *general* conditions for the application of truth predicates; if we could specify such conditions, then we would perforce have a translation of the languages to which the predicates apply. Here is how Davidson (2000, p. 70) puts the point:

[Truth is «interesting and important» because] of its connection with meaning. This is the connection of which Tarski makes use, for translation succeeds only if it preserves truth, and the traditional aim of translation is to preserve meaning.

The idea is this: if a theory cleaves to Tarski's Convention T, then the truth conditions for each sentence of can be furnished by respective translations of each *L*-sentence; translation thus provides the conditions for the application of a truth predicate to *L*. Translation, however, also preserves meaning. If, then, we can give the conditions for the application of a truth predicate for *L*, we will have concomitantly interpreted *L*. An argumentative motif of those who follow Davidson can be best understood, I think, in light of this methodology.

Consider:

- (5) 'Snow is white' is true(in-E) iff snow is white.
 (6) 'Der Schnee ist weiss' is true(in-G) iff snow is white.

(5) appears trivial to *us* because we know the objectlanguage (English). In other words, the conditions for the application of a truth predicate to our own language are antecedently known to us *qua* speakers of the language. Yet (6) reveals that the general case is not trivial: laying down the conditions for the application of a truth predicate for a language we do not know provides us with an informative interpretation. This means that the general notion of translation specified in Convention T does indeed equip truth to play a role in the empirical interpretation of languages. Davidson (1999, p. 85) takes this possibility to refute the idea that truth just is disquotation, for disquotation depends upon a particular instance of translation (homophonic), and so can say nothing about the general case. (Davidson correctly points out that the issue relates specifically to idiolects: if one didn't understand 'snow is white', then being an English speaker would not make (5) trivial for you.). McDowell (2000) gives the name *extended disquotation* to the kind of role the truth predicate plays in (6). This is unfortunate. Field (1994) uses the same phrase for a notion that is designed to show that truth just *is* disquotation after all. A happier nomenclature I have used elsewhere is *cognitive opacity* (Author 1). The notion is best explained in terms of heterophonicity, although it truly applies to idiolects, not languages as such (what Chomsky (1986) would call *E-languages*). Consider a monolingual English speaker, Smith, who is given the sentence

- (7) 'Der Schnee ist weiss' is true(in-G).

Smith can use (7) to express a judgement, say, on Gottlob's assertion (Smith agrees with everything Gottlob says). Smith, however, cannot disquote, for he does not know the truth conditions of Gottlob's assertion. There is more to truth, therefore, than disquotation. This final step, of course, is far too quick, although Davidson is happy to make it. The reason, I think, is that he views truth, retaining my jargon, as *cognitively transparent*; more precisely, to characterise truth for a language is to make the speakers cognitively transparent to the interpreter, it is to reveal the details of their intentional profile. This is achieved through the essential connection of truth with interpretation (translation) which provides for (6). Disquotationism, on the other hand, is bound to particular languages; it cannot tell us what truth predicates have in common without appealing to a substantial notion of translation, but just such a move is to give up on deflationism.⁹

It was suggested above that Davidson, at best, initiates a stand-off with deflationism if he simply claims that truth is not trivial because of its role in a theory of meaning: the deflationist may simply reject such a role and seek to explicate meaning in a non-truth involving way. What does challenge the deflationist is the apparent of optimality of truth in an explanation of compositionality, as elaborated in §3. Do the considerations just rehearsed show otherwise? I think not.

Davidson contends that Tarski established a coupling of truth with translation; in one sense he did, but not so as to confute deflationism. A Tarskian explicit definition makes use of the syntax of the objectlanguage (what, in general, Tarski referred to as the *morphology* of the language) and the ontology therein expressed. Translation is required for the latter factor: the metalanguage needs to be able to express what the objectlanguage expresses. As Tarski (1956, p. 253) notes, this «has no advantage for the pursuit of the 'pure' morphology of the language». The qualification of 'pure' is important: Tarski still saw semantic concepts defined for the objectlanguage as being reducible to its morphology; witness Thesis A' (*Ibid*, p. 273). The reason for this is that the metatheory does not tell us what the objectlanguage sentences mean, the definition of a truth predicate is not informative about meaning or translation; *a fortiori*, a definition of a truth predicate couched in a metalanguage that did not contain a homophonic translation of the objectlanguage would not be informative either. For example, (6) is logically equivalent to (6') (for convenience, the mass noun *snow* is treated as a singular term):

(6') $[(\exists \mathbf{T}_G) [\text{'Der Schnee ist weiss'} \in \mathbf{T}_G \ \& \ \text{'Der Schnee is weiss'} \in \mathbf{T}_G \leftrightarrow \text{'Der Schnee ist weiss'} = \text{'Der Schnee ist weiss'} \ \& \ \text{'Der Schnee'} = \text{'Der Schnee'} \ \text{and} \ \text{'ist weiss'} = \text{'ist weiss'} \ \& \ (\exists \sigma)(\sigma = \text{snow} \ \& \ \sigma \text{ is white})]] \leftrightarrow \text{snow is white.}$

(6') makes patent that a Tarskian truth predicate is employable to tell us what is true given the way the world is, but not what is true given the way meanings are attached to sentences. Convention T can give the opposite impression. An instance of (T) appears to tell

⁹ A tempting line for the disquotationist is the attempt to generalise (T) in some manner so that it is no longer tied to particular languages. Such an attempt, however, appears forlorn (see Gupta (1993), also Davidson (1990a, p. 295-6)). A more promising line is one, very roughly, where *truth* is taken to be primitive, with assent to instances of (T) (or a propositional analogue) being taken to be necessary and sufficient for possession of the concept of truth (see Horwich (1998) and Soames (1999)). This is promising, however, only in the sense of avoiding the immediate generalisation problem; serious difficulties remain concerning compositionality and other factors (see Collins, J. (forthcoming). 'Truth or Meaning? A Question of priority', *Philosophy and Phenomenological Research*).

us about the meaning of the objectlanguage sentence, but when we substitute *definiendum* for *definiens* the appearance melts away. This reveals that, while the definition exploits a translation, the translation is not stated *through* the definition, it is exogenous. Put differently, Convention T is based on the (TM) platitude:

(TM) X means that $p \rightarrow X$ is true $\leftrightarrow p$

(TM) is common currency for all. The deflationist, for example, can take (TM) as indicating that, given a well-behaved translation (a meaning preserving relation), the content of a truth predicate is exhaustively specifiable in terms of that to which it applies (e.g., Soames (1999)). Such may be taken to be the moral of Convention T.

For all Tarski showed, the aptness of a truth theory to serve as a theory of meaning does *not* issue from a reflection on truth itself, for there is no conceptual entwining of truth with an *informative* translation relation. This conclusion does not preclude the attempt to forge such a relation, but there is no *a priori* reason to expect success. What we do know for free is that truth conditions can specify content, so much is enshrined in (TM); but this is not what Davidson is after, for (TM) is not informative in the right way, i.e., it is not the application of a truth predicate that makes content transparent where otherwise it is not. It is here, I think, where the project of a theory of meaning, as conceived by Davidson and followed by many others, goes wrong. In simple terms, a theory of meaning need not be informative in the sense of offering a ‘translation’ of a speaker’s words.

The idea that a theory of meaning should be informative in the way translations are is not best understood as deriving from Tarski. It is better, rather, to understand the idea as derivative of the view that language is, in some sense, intrinsically social: language is behaviour. If this is so, then a translation constraint makes perfect sense, for understanding a language would amount to being able to interpret and be interpreted. This I think, is Davidson’s reasoning.

Davidson’s background assumptions are largely inherited from Quine (1960). In particular: what there is to meaning is that which translation makes available. Davidson (1990b, p. 78) takes it to be ‘obvious... that there can be no more to meaning than an adequately equipped person can learn and observe; the interpreter’s point of view is therefore the revealing one to bring to the subject.’

A theory of meaning is not thereby a *mere* description of behaviour; it does offer a sense of speaker *knowledge*. Even so, the knowledge condition is counterfactual: *if* one were to have explicit knowledge of a theory of meaning for L , then one would be in a position to understand L -speakers (Davidson (1984, p. 125; 1990a, p. 312)). We do not, then, advance beyond (or below, perhaps) behaviour. The theory describes behaviour in an informationally rich, structured, normative way, but «it does not add anything to... say that if the theory does correctly describe the competence of an interpreter, some mechanism in the interpreter must correspond to the theory» (Davidson (1986, p. 438)).

It is difficult to find an argument in Davidson’s work for why the interpreter’s point of view should prevail; why, that is, a theory of meaning should be constitutively insensitive to the ‘mechanism’ that generates coherent language use. Davidson (1990a, p. 314) simply states that «what is to be *explained* is a social phenomenon... language is intrinsically social» [my emphasis]. But how is counterfactual knowledge supposed to be on explanatory duty? My worry here is not so much that of Schiffer (1987), who cannot see how a speaker might know

what Davidson attributes to him; rather, it is that a theory of meaning is *explanatory* of a speaker's linguistic behaviour only if the speaker *actually* possesses the information recorded in the theory. Davidson is certainly correct if he means that the validity of a theory of meaning does not presuppose a particular instantiating mechanism; after all, information is one thing, its instantiation is another. But if that information is going to enter into an explanation of the speaker's performance, then the information does need to be instantiated (somehow) such that the speaker can employ it. Look at it this way: either the speaker (subconsciously) possesses the information our theory of him records or he doesn't. If he does, then its cognitive availability to the speaker will enter into the explanation of his linguistic behaviour and judgements, otherwise the information would be, absurdly, an idle cog. But if he does not possess the information our theory records, then he must have some other information, in which case we want a theory of *that* information.

A natural remedy to this quandary, once all species of behaviourism are foregone, is to follow Chomsky (e.g., 1980, 1986, 2000) and claim that the appropriate *object* for a theory of any aspect of linguistic competence is the internally represented information a subject exploits *qua* language user. In this light, a theory of meaning is a theory of a system of mental representations, the possession of which empirically accounts for the facts of semantic competence.¹⁰ The general support for such a move is very familiar, and I shall not dwell on it. My aim, rather, independent of whether or not Davidson has a coherent conception of speaker knowledge, is to signal what happens to the putative central role of translation if we make this turn towards *internalism* and then in the next section offer a suggestion that so turning saves us from an insuperable tension.

On the prevailing view in linguistics, a truth theory is a description of the information represented by a semantically competent subject. It is, in other words, a theory of *knowledge of meaning*. This knowledge is not constituted by our being able to do anything, still less being disposed to do anything, whatever that might mean. It counts as knowledge just in the sense that it is information we possess and employ in our production and understanding of language. This means, of course, that anyone with the knowledge will, *ceteris paribus*, be able to do many things, such as translate another speaker. But, again, no such abilities amount to what it is to know a language. *Meaning is not use*. This being so, how are we to understand the apparent translation a truth theory makes available? Is it accidental?

Rather than say that a truth theory provides a *translation* or *interpretation*, we may say that it should issue in a statement of the conditions for sentences' evaluation as true or false (not necessarily disquotationally). A truth theory should do this precisely because it amounts to a test against the intuitive data that the theory records the information relevant to the understanding of the sentence in a way that explains how a sentence's meaning is composed from the meaning of its parts.¹¹ This was the big moral of §3. So understood, the articulation

¹⁰ This kind of approach to the theory of meaning (as truth theory) is set out in Higginbotham (1985, 1986) and Larson and Segal (1995).

¹¹ It might seem here that a traditional worry resurfaces, *viz.*, that, for many constructions (famously, context variable ones), the right flanks of instances of (T) will be ampliative in a way which either obviates the 'absolute' reading of *truth* or vitiates the instances' intuitive acceptability. The worry is misplaced. There is no demand for homophonicity. The requirement, rather, is that the instances of (T) record intuitively acceptable paraphrases on their right flanks of the sentences described to their left.

of content the theory affords will indeed, at least for many constructions, provide translations or interpretations, but these will not be informative in the sense asked for by Davidson's conception of success. Indeed, if they were so informative, the point of the exercise would be lost, for instances of (T) count as secure data *because* they are intuitively obvious; if they were in doubt, then our theories would not be corroborated by entailing them.

In effect, my suggestion here is the converse of Davidson's initial idea in 'Truth and Meaning', where homophonic instances of (T) are taken to be non-corroborative of the theory that entails them because they are uninformative; heterophonic instances, on the other hand, do corroborate their entailing theories because they show that the theories issue in genuinely informative interpretations. Once, however, we drop the notion that a theory of meaning should constitutively double as a theory of interpretation, then the idea that we should be able to frame an English truth theory for German, say, is wholly otiose. Our ability to do such a thing does not enter into the justification of a truth theory as theory of meaning. If truth theories are theories of internally represented information, then neither an English speaker nor a German speaker, *qua* such speakers, represent an English truth theory for German, or vice versa; and so a truth theory for one language couched in the other will not be a reflection of any underlying competence. Maybe we can construct and test such theories as Davidson envisages, but nothing important about our semantic competence would be obviously revealed by the exercise. For instance, getting to know the truth conditions of utterances we do not antecedently understand will certainly give us a systematic grip on what the speakers are saying (e.g., Field (1978) and Lepore and Loewer (1987)); but this would be due to the co-operation of our *theory of mind* and our own *antecedent* semantic competence: one's gaining such a grip does not constitute either one's own understanding or that of one's interpretee.

The justification of a truth theory as a description of a speaker's semantic competence ultimately rests on evidence, not conceptual reflection. Does the information recorded enter into explanations of our judgements about meaning and our typical performance? Does the structure of the theory explain the apparent structure of semantic competence, such as systematicity, compositionality, unboundedness, etc.? Is the information such as to be computationally tractable? Does the semantics offer a smooth interpretation of syntactic structure at the level which interfaces with human conceptual resources in general (LF, say)? E.g., is the theory consistent with the existence of 'empty categories'? Answers to these kinds of questions and many others will tell us if we are on the right track.

5: *Language in Mind*

From Davidson's remarks quoted above, it is clear that he considers the psychology or cognitive science of language to be one thing, and a theory of meaning to be quite another. This separation of the conceptual/normative from the empirical is now very common in a variety of forms. It might thus be countered that the above recommendations are a change of subject rather than a challenge to the interpretive conception of the job of a theory of meaning. This issue, suffice it to say, is far too wide ranging to be properly engaged with

The paraphrases will be as ampliative as the lexical axioms from which the theorems are derived. That a truth theory has such consequences while cleaving to compositionality is what, in part, corroborates the hypothesis that a speaker's semantic competence is constituted by knowledge of the entailing theory. I think that there is good evidence that this demand can be met without a retreat to model theory. The issue, though, I should say, is wholly empirical.

here. Still, on the assumption that empirical inquiry into the nature of cognition is not in any sense illegitimate, I think a general problem obtains for the imagined separation, and its instance with regard to linguistic meaning is especially vivid.

Davidson himself was one of the first to appreciate that a theory of meaning must satisfy constraints that are essentially empirical, such as explaining unboundedness, novelty and compositionality. These kinds of constraints now go largely without comment, but they are empirical in that they reflect real contingent features of our linguistic competence. Once such constraints are accepted, then, it is difficult to screen off one's theory from empirical findings. It has proved convenient for those who *a priori* insist on the intrinsic sociality of language to characterise a naturalistic attitude to the mind — what McDowell (1994) disparagingly refers to as «bald naturalism» — as one which hankers after a reduction of meaning and related concepts to mechanical/computational states: what is not causally accountable for just ain't real. This characterisation presents naturalism as crude scientism, an ontological thesis that straightforwardly conflates *is* with *ought*, aetiology with constitution. It is certainly true that many philosophers have reductionist ambitions, but they are not *required* by naturalism; indeed, a lot of such positions, I should say, are more informed by long standing philosophical intuition than any scientific research. Naturalism is a *methodological* thesis that has no *a priori* commitment to what meaning, or anything else, for that matter, really *is*; rather, it is committed to the construction of empirical theories that attempt to explain the myriad of factors that enter into linguistic competence. There is very little mileage to be gained, therefore, from an attack on naturalism as if it were a thesis about what there is, for such benighted assaults go no way to secure the legitimacy of the initial move to screen-off one's philosophy from empirical findings. It is the *very idea* that one can determine substantive claims about the sociality of language from conceptual consideration alone that is in tension with naturalism, not the putative irreducibility of the mental.

I am not here begging the question in favour of a language organ (faculty) or any other naturalist project, the point is a general one: whatever the shape of our concepts about natural phenomena *appears* to demand may turn out to be wrong. In the present case, I have little doubt that the best description of our *Lebenswelt* involves linguistic inter-subjectivity as constitutive. I also think that mutual interpretation in a certain mode is something language enables rather than what it is; the ability to communicate is neither necessary nor sufficient for linguistic competence. In this regard, our *Lebenswelt* is misleading and the *a priori* constitutive theses are just false. One may abjure interest in anything other than the conceptual layout of *what it is like* to be human, but if so, then it ill behoves one to make substantive claims. The crucial point is that once empirical constraints are acknowledged (e.g., unboundedness), then it is illicit to commit to just those features of linguistic understanding that are consistent with one's pre-conceived *a priori* views to constrain one's theory. Unfortunately, to acknowledge empirical constraints is one thing, to treat them seriously is another. Let us first look at Brandom as an example.

Brandom (2000, pp. 124-9) wants to show that the novelty and unboundedness of linguistic understanding is flush with his inferential-social model of thought. The explanation offered, however, is that the child is *trained* on a finite number of sentences and is then able to project to the «correct uses» of boundlessly many sentences through a segmentation of the members of the training set into substitution classes of singular terms and predicates. This Bloomfieldianism is patently motivated by Brandom's social view of linguistic understanding rather than a concern with novelty as a real phenomenon that is independent of any favoured

theory. Unsurprisingly, it faces serious problems. First, the model is false. Any developmental text book will tell one that children are not trained on a finite number of sentences. Children begin with single words, go through a two word phrasal stage, and then onto full grammatical complexity (function words, inflections, etc.). The whole process is characterised by novelty and creativity. Second, there is no known projection function from a corpus of sentences to a grammar proper. Brandom's substitutional method fails even with such a crude partition as singular term and predicate. It is, on the one hand, too broad, e.g., '...is dangerous' admits names (paradigm singular terms) as well as general nouns, gerunds, infinitives, quantifier phrases, *wh*-phrases, clauses, etc.; on the other hand, it is too narrow; e.g., '...washed herself' excludes all substitutends apart from third person, singular, feminine. The fact is that outside of artificial languages, substitution does not determine grammatical categoricalness; and there are, of course, 'scrambled' languages, such as Warlpiri and Latin, with free order, where phrasal constituents may even be discontinuous. Third, what explains novelty is structure, e.g., that $[_{NP} NP PP]$ is recursive means that if one understands $[X [PP [Y]]]$ (e.g., *the boy behind the girl*), then one will understand $[X [PP [Y [PP [X]]]]]$, etc. If one does not have the grammatical concept of a *preposition*, under which a given preposition may freely take NP arguments, then being able to use a given PP will not *ipso facto* enable you to use any other PP iterated thereof. The best explanation we have of this is that grammatical selection features are given in the lexicon. Perhaps this is wrong, but if words are not distinguished by grammatical features, then novelty and boundlessness are not explained; and if words are so distinguished, then the projection Brandom has in mind is wrong: it is projection from the lexicon that is explanatory, not from a corpus of sentences. Similar complaints can be made against Davidson.

Davidson (1997b, p. 20) says that that linguistic capacities are largely innately endowed is not «philosophically significant», although he does think that it is probably true (so much the worse, perhaps, for such 'significance'.) Davidson's point appears to be that such 'facts' shed no light on the *content* of our thoughts and how we come by such content, the facts concern only syntax. The motivation here, however, derives from his unabashed «social engineering» view of language rather than any considered view of the empirical data, which, I should say, does not support his position.¹² Even so, does Davidson here offer a principled screen behind which an interpretive conception of meaning may remain unmolested by data or empirical theories?

It is fair to say that in modern discussion the view that language is characterised by novelty, systematicity, and unboundedness comes directly from Chomsky's early pioneering work (1957), although the characterisation goes back at least as far Descartes. The mere

¹² Curiously, Davidson (1997b, pp. 20-1) claims to have no particular problem with language acquisition being largely governed by innate principles, and he cites Pinker's (1995) survey of the evidence in support of this nativism. He then, however, goes on to offer what is essentially an *ostensive learning by negative evidence* model (*triangulation*) of conceptuality as such, including lexical acquisition (also see Davidson (1997a, 2000)). But, famously, negative data appears to play *no* role at all in language acquisition; certainly the ostensive model does not square with the lexical 'explosion' that takes place around 24-30 months. Elsewhere, Davidson (1989, p. 164) is more conciliatory: «Of course very many words and sentences are not learned [by ostension and correction], but it is those that are that anchor language to the world». I know of no data, however, to support the claim that a certain class of words is learnt by ostension. Besides which, the assumption that language *is* anchored to the world begs the question in favour of the social externalism which is presently in question.

existence of such features, of course, does not reveal the falsity of the social view of language.¹³ Indeed, for Chomsky in the late '50s psychological issues were very much in the background. It is also fair to say, however, that linguistics of the past thirty years or so has made it beyond informed doubt that linguistic understanding involves sensitivity to information that is not communicated or in any way encoded in the data (input) the hearer or learner is faced with. Nor is this information necessarily syntactic as opposed to semantic, as Davidson's dismissal of the significance of the internalism of syntax would suggest. In fact, the kind of distinction between syntax and semantics in which Davidson trades is problematic.

Syntax is no less conceptual than semantics. That a speaker can recognise that, say, it is the object of *meet* which is being questioned in *Whom does Bill want to meet?* (I help out here by using the unfortunately moribund accusative form), presumably requires the concept of transitive verb, and much else besides, just as, I should say, the recognition that Rover is a dog requires the concept of a dog. If there is a fundamental difference here, no-one has made clear what it is. Even if there is a principled separation, it just does not follow that the difference is that semantics is external (social) while syntax is internal (private). (See Chomsky (2000) and Segal (2000) for a wholly narrow view of content.) There is the tendency to think of syntax as meaningless symbols, as we think of '1's and '0's on a Turing machine tape that do not mean anything until we interpret the machine as, say, computing in base 2. Deep confusion arises when we think this way, for it encourages the thought that syntax and semantics are related as representation to represented, as if syntax is 'inside' and semantics is 'outside'. Patently, however, when we theorise about the syntactic competence of a subject, we do not think that syntax is a meaningless representational medium upon which the subject drapes meaning. If that were the case, a subject would not be in a position to make *judgements* about grammaticality; contrary to fact, competence would be an automatic mechanical response rather than a resource we can *freely* employ. Syntax, along with semantics, is represented in the mind, it is not the representational medium of the mind. Syntax and semantics are properly thought of as different levels of information that a competent subject has at his disposal *qua* competent.¹⁴ Unsurprisingly, where syntax ends and semantics begins is not clear.

One may, of course, think that there is something special, conceptually revealing, about compositionality in distinction to empty categories and rest of the abstract structure revealed by current linguistics, although I know of no argument that even suggests that this is true; and even if it were true, we would still be owed an explanation of the phenomenon. If, then, one constrains a theory of meaning to be compositional simply because linguistic understanding appears to be compositional, by parity of reasoning, one must constrain one's theory of meaning to accommodate empty categories, etc. But if one is so consistent, then the kind of «social engineering» model Davidson favours is in real trouble. It *might* be, as Davidson (1997b, pp. 25-6; 2000, pp. 71-2) suggests, that we learn concepts through ostensively learning

¹³ One might argue that there are syntactic constraints, what is social is 'meaning'. It is difficult to take such arguments seriously without being given a principled *cognitive* distinction between syntax and semantics (see below).

¹⁴ It does not follow, of course, that grammatical information is consciously accessible. Although sometimes it is, and when it is not, it clearly shapes our choices and judgements rather than simply causes what we say.

the reference of words, with ourselves forming a *triangle* with teacher and the world, where the former provides shared saliences and the latter correction. I think not. But it just makes no sense at all that we ostensibly learn empty categories, for there is no ‘word’ to learn. The point here is quite different from, say, the situation with inflectional properties. Such properties are phonetically marked, but they have no reference that may be ostended, they simply mark agreement features. Still, we can just imagine, I suppose, some extrapolation from a matrix of previously fixed words (no-one, of course, has a clue how this might be achieved). Empty categories, however, do refer, possessing pronominal and/or anaphoric features. This is crucial. Empty categories are not expletives such as *it* and *there* (in some constructions) in English that occur without making a semantic contribution. The presence of expletives in English is due, it seems, to a *structural* requirement for overt subjects. Empty categories, on the other hand essentially enter into the semantic evaluation of the constructions in which they appear. Empty categories, therefore, cannot be dismissed as mere «syntactic constraints». Let us briefly exhibit this difference with a grammatical chestnut.

In *Bill is eager to please*, the infinitive *to please* is transitive, but what are the arguments of the verb? No competent speaker has trouble telling us that *Bill* is not the object, for it is Bill who wants to be pleasing, even though, *qua* infinitive, the verb’s subject is not marked; in finite paraphrase, *Bill is eager that he (Bill) please someone or other*. But compare, *Bill is easy to please*; here *Bill* must be the object of the verb, with the subject being empty (this becomes apparent with the paraphrase, *It is easy to please Bill*.) The difference is obvious but subtle still, for it depends on the fact that *easy*, unlike *eager*, does not categorise for an external (subject) argument; hence, the subject of the infinitive complement of *easy* is an empty category PRO, where the subject of *eager* is the raised subject *Bill*. In sum, a competent subject must understand the semantic values of empty categories, but no such values can be ostended, there is nothing to *triangulate*. Again, as far as we can see, the only place for empty categories to be is as represented in the minds of speakers/hearers. To acknowledge this is to admit that language is not «intrinsically social».

We have come some distance from the idea that a truth theory ought to embody a translation constraint; I hope, however, that the dialectic is clear. We can state a truth theory for a speaker (idiolect) in ways that satisfy constraints of novelty, compositionality, etc., and also offer intuitive tests of its correctness in the form of the demand that it entails articulations of the content of the speaker’s sentences. None of this involves translation in the substantive sense of a truth theory being constrained to issue in informative translations of ‘foreign’ sentences. The change of tack is a move from a social perspective, one under which a translation constraint makes sense, to a cognitive or internalist perspective, under which translation is otiose. I have attempted to show that staying with the former perspective lands the truth theoretic approach to meaning in an empirical dilemma out of which there is no apparent escape.

6: Concluding Remarks

Davidson’s initial presentation of the truth theoretic approach showed the greatest sensitivity to empirical issues and could be profitably understood as continuous with Chomsky’s work, as Davidson (1984, pp. 22/30) himself suggested. Davidson’s clarity on the demand that a *theory* is required rather than consultation of dictionaries has justly proved to be seminal. Above all, Davidson’s contention that a truth theory is what is needed to account for compositionality and unboundedness is a lasting achievement. As my title intimates, I seek a revision of Davidson, not a refutation. Once we get clear on the relationship between Tarski,

a truth theory and deflationism, we can see Davidson as offering a profound insight into truth and meaning, one which is flush with an empirical approach to linguistic understanding.

References

- Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, Mass.: Harvard University Press.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Chomsky, N. (2000). *New Horizons in the Study of language and Mind*. Cambridge: Cambridge University Press.
- Collins, J. (forthcoming), 'Truth or Meaning? A Question of priority'. *Philosophy and Phenomenological Research*.
- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Davidson, D. (1986). 'A Nice Derangement of Epitaphs'. In E. Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (pp. 433-46). Oxford: Basil Blackwell.
- Davidson, D. (1989). 'The Myth of the Subjective'. In M. Krausz (ed.), *Relativism: Interpretation and Confrontation* (pp. 159-72). Notre Dame, Ind.: University of Notre Dame Press.
- Davidson, D. (1990a). 'The Structure and Content of Truth'. *Journal of Philosophy*, 87: 279-328.
- Davidson, D. (1990b). 'Truth, Meaning, and Evidence'. In R. Barrett and R. Gibson (eds.), *Perspectives on Quine* (pp. 68-79). Oxford: Blackwell.
- Davidson, D. (1997a). 'The Folly of Trying to Define Truth'. *Journal of Philosophy*, 93: 263-78.
- Davidson, D. (1997b). 'Seeing Through Language'. In J. Preston (ed.), *Thought and Language* (pp. 15-28). Cambridge: Cambridge University Press.
- Davidson, D. (1999). 'Reply to W. V. Quine'. In L. E. Hahn (ed.), *The Philosophy of Donald Davidson* (pp. 80-6). Illinois: Open Court.
- Davidson, D. (2000). 'Truth Rehabilitated'. In R. Brandom (ed.), *Rorty and his Critics* (pp. 65-74). Oxford: Blackwell.
- Field, H. (1978). 'Mental Representation'. *Erkenntnis*, 13: 9-61
- Field, H. (1994). 'Deflationist Views of Meaning and Content'. *Mind*, 103: 249-285.
- Fodor, J. (1998a). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.
- Fodor, J. (1998b). In *Critical Condition: Polemical Essays on Cognitive science and the Philosophy of Mind*. Cambridge, Mass.: MIT Press.

- Fodor, J. and Lepore, E. (1991). 'Why Meaning (Probably) Isn't Conceptual Role'. *Mind and Language*, 6: 328-76
- Fodor, J. and Lepore, E. (1992). *Holism: A Shopper's Guide*. Oxford: Blackwell.
- Gupta, A. (1993). 'A Critique of Deflationism'. *Philosophical Topics*, 21: 57-81.
- Higginbotham, J. (1985). 'On Semantics'. *Linguistic Inquiry*, 16: 547-93.
- Higginbotham, J. (1986). 'Linguistic Theory and Davidson's program in Semantics'. In E. Lepore, *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (pp. 29-48). Oxford: Basil Blackwell.
- Higginbotham, J. (1999). 'A Perspective on Truth and Meaning'. In L. E. Hahn (ed.), *The Philosophy of Donald Davidson* (pp. 671-86). Peru, Ill: Open Court
- Horwich, P. (1990). *Truth*. Oxford: Blackwell.
- Horwich, P. (1998). *Meaning*. Oxford: Oxford University Press.
- Horwich, P. (1999). 'The Minimalist Conception of Truth'. In S. Blackburn and K. Simmons (eds.), *Truth* (pp. 239-63). Oxford: Oxford University Press.
- Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, MA.: MIT Press.
- Larson, R. and Segal, G. (1995). *Knowledge of Meaning: An Introduction to Semantic Theory*. Cambridge, MA.: The MIT Press.
- Lepore, E. and Loewer, B. (1987). 'Dual Aspect Semantics'. In E. Lepore (ed.), *New Directions in Semantics* (pp. 83-112). New York: Academic Press.
- McDowell, J. (1994). *Mind and World*. Cambridge, Mass.: Harvard University Press.
- McDowell, J. (2000). 'Towards Rehabilitating Objectivity'. In R. Brandom (ed.), *Rorty and his Critics* (pp. 109-23). Oxford: Blackwell.
- Pinker, S. (1995). *The Language Instinct*. London: Penguin.
- Quine, W. V. O. (1960). *Word and Object*. Cambridge, Mass.: MIT Press.
- Russell, B. (1903). *Principles of Mathematics*. London: Allen and Unwin.
- Sainsbury, M. (1996). 'How Can We Say Something?' In R. Monk and A. Palmer (eds.), *Bertrand Russell and the Origins of Analytical Philosophy* (pp.137-153). Bristol: Thoemmes Press.
- Schiffer, S. (1987). *The Remnants of Meaning*. Cambridge, MA: MIT Press.
- Segal, G. (2000). *A Slim Book About Narrow Content*. Cambridge, Mass.: MIT press.
- Soames, S. (1999). *Understanding Truth*. Oxford: Oxford University Press.
- Tarski, A. (1956). 'The Concept of Truth in Formalized Languages'. In *Logic, Semantics, Metamathematics* (pp. 152-277). Oxford: Oxford University Press.

John Collins

<JCollins42@compuserve.com>

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 72-79

Identity, Analyticity and Epistemic Conservatism

Copyright © by SORITES and Paul Tomassi

IDENTITY, ANALYTICITY AND EPISTEMIC CONSERVATISM

Paul Tomassi

Introduction

Rigorous, i.e. non-metaphorical, formulations of the analytic/synthetic distinction are thin on the ground. Moreover, many will sympathise with Quine's explanation of that fact.¹ Famously, Quine distinguishes two purported types of analytic statement:

Those of the first class, which may be called *logically true*, are typified by:

(1) No unmarried man is married.

...But there is also a second class of analytic statements, typified by:

(2) No bachelor is married.

The characteristic of such a statement is that it can be turned into a logical truth by putting synonyms for synonyms; thus (2) can be turned into (1) by putting 'unmarried man' for its synonym 'bachelor'.²

It follows, in Quine's view, that analyticity is not properly defined in the absence of an adequate definition of cognitive synonymy. The attempt to construct such a definition invariably brings (vicious rather than virtuous) circularity in its wake. The correct explanation of that fact, Quine argues, is that the analytic/synthetic distinction is merely (one) empiricist dogma. Quine's critique of a second empiricist dogma (reductionism) reinforces his earlier conclusion: there is no such distinction to be drawn. Consequently, no statement, laws of logic included, is ultimately immune to revision in the light of experience.³ Here I suggest that even if much of Quine's reasoning is cogent his conclusion is not warranted. Complete repudiation of the analytic/synthetic distinction is not the only conclusion which can validly be drawn. While I accept that there is a genuine problem of demarcation here I hope to show that the particular demarcation considered by Quine is the result of epistemic aggrandisement. The account of analyticity proposed may also cast new light on questions of warrant as regards essentialist accounts of identity-statements; these consequences are considered in the final section of the paper.

¹ Quine, W.V.O., 'Two Dogmas of Empiricism' in *From a Logical Point of View*, Harper Torchbooks, Harper and Row, New York. Second Edition. 1963. Pages 20-46.

² *ibid.* Pages 22-23. The italics are Quine's.

³ *ibid.* Page 43.

I

With no intended sleight of hand, the account of analyticity considered by Quine might be crystallised as a distinction between two kinds of identity-statement. The first of these is generally taken to exemplify the law of identity, i.e. the thought that everything is identical with itself:

Type I $a = a$

In ‘Two Dogmas of Empiricism’, Quine’s attitude to statements of this form is ambiguous. Laws of logic are revisable in the light of experience. More specifically, it is suggested that this fate might already have befallen the law of excluded middle.⁴

While it is now clear that the latter claim is erroneous (at least for the reasons Quine gives), by the former claim, Quine would appear committed to the potential revocability of the law of identity on empirical grounds. Current physical theory may offer no basis for any such revision but science is progressive and, in the longer run, who knows. On the other hand, Quine’s case against the first dogma is not made *modulo* logical truths but rather on the basis of a second type of purportedly analytic statement:

Type II $a = b$ where and only where (a, b) is a synonym-pair.

As Quine puts it:

... the major difficulty lies not in the first class of analytic statements, the logical truths, but rather in the second class, which depends on the notion of synonymy.⁵

Given the recalcitrance of cognitive synonymy as regards (non-circular) definition, Quine is right to claim that no analytic/synthetic distinction has yet been cogently drawn. He may, however, be wrong to conclude that there is no such distinction. Quine’s basic concern is how scientifically-informed belief-sets warp in experience’s light. Given that belief-sets warp under guiding principles which are pragmatic, conservative and simplicity-seeking, it is difficult to see what could possibly motivate rejection of the law of identity, and more difficult still, to imagine how any belief-set could actually be formed (or adjusted) in ways which did not keep faith with the law of identity.

The matter will not be settled by consensus but, I take it, few if any will want to dispute the analyticity of statements of type I. I further assume that the reason underpinning that reaction is constituted by the form of the sentence itself. Type I statements are epistemically reassuring in the sense that we generally grasp the truth of any such sentence as soon as we grasp the meaning of the symbols which compose it. The source of that reassurance consists in our understanding of the identity sign and our recognition that the same term fills both gaps in the identity relation. Indeed, in this we have a very clear definition of type I identities, i.e. *identity-statements in which one and the same term plugs both gaps in the relation*. Thus, it is very natural to assert that every such sentence is true in virtue of the meaning of the symbols which compose it, i.e. at first blush, at least, every sentence of that form seems guaranteed to be true. Moreover, it would appear that the content of any such statement can be paired down very considerably while retaining (intuitively, at least) self-evidence. Even

⁴ *ibid.*

⁵ *ibid.* Page 24.

here, however, there are limits. For example, the poverty of current flubjub-theory notwithstanding, one might be tempted to accept the truth of the following statement:

Flubjubs are flubjubs.

Given the logical form of the statement and the validity of existential generalisation, one might then infer:

Flubjubs exist.

Given that there are no flubjubs, such reasoning is invalid. Hence, we may require extraneous reassurance in some form that the name in question is not vacuous; in other words, that the concept in question has an extension. Given such minimal extraneous information, type I statements are robustly epistemically reassuring, i.e. Type I statements as defined which satisfy the reference-condition wear their analyticity on their faces; as we might put it, such statements exhibit *orthographic identity*.

Whether or not Quine ultimately shares with his opponents recognition of the epistemically reassuring character of type I statements, it is clear that, in Quine's view, the real trouble kicks in with the attempt to extend analyticity to type II statements. The key assumption here is just that there is a class of such sentences. However, if there is any such class, no member of that class exhibits orthographic identity. Type II identity-statements do not wear their analyticity on their faces. Indeed, the logical form of a type II statement is generally definitive of syntheticity rather than analyticity. Further, if the claim is that a certain class of such statements has this form but is nonetheless properly understood as analytic rather than synthetic, the key question is exactly which class? In other words, which substitution-instances are properly included in the analytic 'box' and which excluded?

The answer seems obvious. Type II statements of the form $a = b$ are analytic where and only where (a, b) is a synonym-pair. But now Quine's points as regards the recalcitrance of cognitive synonymy all become pressing. For singular terms in the relation, for example, it is now not enough simply to know that both terms refer. Equally, for general terms, it is not enough to know that the relevant extensions are non-empty or even that such extensions are identical. Thus, there is not only a challenge to clearly pin down type II analytic truths but also a fundamental epistemic contrast between the two purported forms of analytic truth. The challenge here is not adequately met by pointing to one or two statements of the relevant form which, most would agree, assert true identities. Rather, what is required is clear demarcation of the relevant class.

Type II identity-statements strongly contrast with those of type I. The former, unlike the latter, clearly are not orthographic identities. We noted above that recognition of the truth of a type I statement requires only that we are aware that the statement has type I form and that the first term in the relation refers. In contrast, knowing that a statement is of type II form and that the first term in that relation refers is not a state of information adequate to establishing the truth of any type II identity. Moreover, if the first term in a type I identity refers it is certain that the second term also refers. Again, however, the referential success of the first term in a type II identity offers no guarantee of reference as regards the second term. Therefore, there are clear epistemic differences between the two kinds of identity-statement. Given the lack of a clear demarcation of the purported class of type II analytic truths together with the fact that there is no reassuring orthographic identity in the case of any type II statement, the question naturally arises: why should we accept that there is any such class?

Thus Quine's boundary problem may indeed be one of aggrandisement, i.e. the notion of analyticity has been stretched beyond the genuinely self-evident. Traditional attempts to draw the distinction fail with Quine's attempt precisely because these strategies try to draw the line between the analytic and the synthetic in the wrong place.

II

Given the foregoing, we may rationally accept Quine's critique of type II 'analytic' statements while rejecting his conclusion, i.e. we may choose to warp our belief-set in the light of our experience rather differently. Given the Duhem-Quine thesis, our hand is not forced to any particular local fix and thus the move may even be Quinean in spirit. To be quite clear, the alternate conclusion is a position which confines analyticity to orthographic identity; given the background to this debate, a position aptly entitled *epistemic conservatism*, i.e. conservatism *modulo* analyticity. In Quinean spirit still, we ought to keep an eye to the knock-on effects of such an adjustment to, and in, the bigger picture. What does philosophical cost-benefit analysis reveal here?

Before considering that question, it is worth forestalling one obvious objection. Thus far, the only restriction imposed upon orthographic identities over and above logical form is the reference-condition, i.e. the term in that relation must be non-vacuous. Those who sympathise with Kripkean accounts of the semantics of names and natural kind terms in terms of rigid designation, for example, may rest easy given that condition alone.⁶ While I cannot fully make the case out here, given the work of Gareth Evans and others, it seems likely that some descriptive content (however minimal, e.g. sortal) must attach to names.⁷ However, the move to a (weakly) descriptivist, anti-Kripkean, position does nothing to undermine the key epistemic differences between the two types of identity-statement highlighted here. Assuming that there is some such content, that content will attach to the terms in any identity-statement. As regards orthographic identities, we can be equally certain *ex hypothesi* that exactly the same content attaches to each term in every such case. By the same assumption, there is no guarantee whatsoever that the same descriptive content will attach to the terms of any non-orthographic identity. Thus, a descriptivist stance preserves key epistemic differences between the two types of identity-statement.⁸

⁶ See, for example, Kripke, S., 'Identity and Necessity', reprinted in Moore, A. (ed.) *Meaning and Reference*, Oxford Readings in Philosophy, Oxford. 1993. Pages 162-191.

⁷ See, for example, Evans, G. 'The Causal Theory of Names', reprinted in Moore, A. (ed.) *Meaning and Reference*, Oxford Readings in Philosophy, Oxford. 1993. Pages 208-227.

⁸ More robustly descriptivist analyses are possible here. But these may come at a price. On a Russellean analysis, for example, it would seem to follow that there might be worlds in which such terms failed to refer, i.e. worlds in which orthographic identities might be false. To adopt a descriptivist position does not entail accepting every detail of Russell's analysis of descriptions, however. We may, for example, sympathise with Strawson's critique of Russell's theory (see, for example, Strawson, P.F., 'On Referring' reprinted in Moore, A. (ed.) *Meaning and Reference*, Oxford Readings in Philosophy, Oxford. 1993. Pages 56-79). According to Strawson, reference is presupposed rather than entailed. Thus, where referent is lacking so too is truth-value, i.e. the relevant class of orthographic identities is neither true nor false. Orthographic identities containing referring terms are therefore true while orthographic identities containing vacuous terms are neither true nor false. It follows that no orthographic identity is ever false. In other words, there is no possible world in which any orthographic

To sum up: the account of analyticity proposed here confines the analytic to orthographic identity-statements satisfying the reference condition. Clearly, every such identity is necessarily true in a highly intuitive sense. Moreover, that intuition can be made more precise. In every possible world in which the first term in an orthographic identity-statement refers that statement is guaranteed true. Therefore, there is no world in which any such orthographic identity is false. Given the familiar interpretations of the notions of possibility and negation, it follows that true orthographic identities are true necessarily, i.e. no such orthographic identity is false in any possible world. Further, while it is clear that the epistemic character of type I identities is underwritten merely by the form of the statement and satisfaction of the reference-condition, it is equally clear that the same basis is inadequate as regards establishing the necessity of any type II identity-statement. Here, merely grasping the form of the statement and knowing that the first term refers is plainly insufficient to establish the truth, let alone the necessity, of any type II identity-statement. There remain, therefore, important epistemic differences between the two kinds of identity-statement precisely as regards warrant for analyticity and, by the same token, necessity.

III

The account of analyticity proposed here not only allows a clear (if modest) analytic/synthetic distinction to be drawn but also provides a useful framework within which questions of warrant as regards essentialist accounts of identity-statements can usefully be considered. As is very well known, certain authors (chief among them Saul Kripke and Hilary Putnam) have argued for an account of a certain class of type II identity-statements which, if true, are true necessarily.⁹ Famously, the terms in such identity-statements are the prime candidates for rigid designation, i.e. names and natural kind terms. Of course, such identity-statements are not canvassed as simply analytic in the sense proposed here. Rather, in the light of identities of this type, we are urged to distinguish epistemic necessity from metaphysical necessity. Identities of the relevant kind are metaphysical necessities known a posteriori.¹⁰ It follows that there is, indeed, a class of properly type II identities which, if true, are necessarily true. Given the foregoing discussion, knowledge of the logical form of any such statement and the fact that the first term in any such statement refers is, just as such, insufficient to warrant recognition of the necessity of that statement. In these terms, the key question is exactly how that state of information is supplemented to constitute warrant? To that end, I separate cases. Consider, for example, Kripke's claim to the necessity of the statement:

identity is false. Given the familiar interpretation of the notions of possibility and negation, orthographic identities are necessarily true, i.e. no orthographic identity is false in any possible world. Of course, this analysis assumes the correctness of Strawson's position; a precarious assumption given Stephen Neale's critique of Strawsonian analyses in his *Descriptions*, Cambridge MA, MIT Press, 1990.

⁹ Re Kripke, for example, see reference 6 above. Re Putnam, see, for example, Putnam, H., 'Meaning and Reference', reprinted in Moore, A. (ed.) *Meaning and Reference*, Oxford Readings in Philosophy, Oxford, 1993. Pages 150-161.

¹⁰ I am indebted to Timothy Kenyon for many of the points made in this section of the paper which, moreover, shape the paper as whole.

The lectern is not made of ice.

Were the truth of the statement established by fiat, there could be no sense in which our knowledge (if indeed we have such knowledge) is the result of empirical discovery. That this is not Kripke's position is quite clear:

What we know is that first, lecterns usually are not made of ice, they are usually made of wood. This looks like wood. It does not feel cold and it probably would if it were made of ice. Therefore, I conclude, probably this is not made of ice. Here my entire judgement is a posteriori.¹¹

Here, the key point is that to establish the necessity of the statement in question is to establish that it is not possible that the statement be false, i.e. that there is no possible world in which the statement is false. But what in the passage quoted assures us of that fact? The statement is highly probable. But any degree of probability less than 1 is plainly inadequate to Kripke's task. The statement's being .99 probable and yet being false is perfectly consistent. It follows that there is a possible world in which the negation of the statement holds; whence, then, its necessity? Kripke outlines an argument-form which, under substitution, is intended to establish the necessity of the statement as conclusion:

$P \supset \Box P$

P;

Therefore

$\Box P$

The conclusion — ' $\Box P$ ' — is that it is necessary that the table not be made of ice¹²

To be clear, the validity of the argument-form outlined is not in question here. However, to draw the conclusion soundly requires establishing the truth of all the premises. Thus, the key question is the nature of the warrant for P itself. Certainly, given P, and given that if P then necessarily P, necessarily P. But, again, the truth of P is by no means conclusively established by Kripke's argument. In fairness to the position in question, and lest anything significant should hang merely on one particular example, consider Putnam's candidate identity:

Water is H₂O

Putnam is clearly aware that establishing the necessity of the identity-statement requires ruling out the possible non-obtaining of this truth (if it is one) in some possible world:

... we can perfectly well imagine having experiences that would convince us (and that would make it rational to believe that) water *isn't* H₂O. In that sense, it is conceivable that water isn't H₂O. It is conceivable but it isn't possible! Conceivability is no proof of possibility.¹³

¹¹ Kripke, S., 'Identity and Necessity', reprinted in Moore, A. (ed.) *Meaning and Reference*, Oxford Readings in Philosophy, Oxford. 1993. Page 180.

¹² *ibid.*

¹³ Putnam, H., 'Meaning and Reference', reprinted in Moore, A. (ed.) *Meaning and Reference*, Oxford Readings in Philosophy, Oxford. 1993. Page 159.

Again, were the truth of this identity-statement established by fiat, there could be no sense in which such knowledge is the result of empirical discovery. But Putnam is equally clear that the truth of the identity-statement in question is an empirical matter: ‘Once we have discovered that water (in the actual world) is H_2O , *nothing counts as a possible world in which water isn’t H_2O .*’¹⁴ Given that the discovery in question is empirical in character it remains to ask what it is that ultimately assures us of the truth of the identity-statement in question to the extent that we may confidently assert the necessity of that statement?

Undoubtedly, that water is H_2O is currently an uncontroversial commonplace of a mature science. However, it does not follow that the truth of that statement is thereby guaranteed. Scientific statements are, generally, defeasible statements. Therefore, it remains possible that any such statement will turn out to be false. In the bigger picture, Larry Laudan’s ‘confutations’, may show that the appropriate general meta-induction is pessimistic rather than optimistic.¹⁵ Certainly, chemistry is no exception to the rule. The history of science does not bear out the thought that the percentage composition of familiar compounds within established chemical theories is immune to revision. Indeed, just that point was effectively exploited by Thomas Kuhn precisely in order to establish the revocability not only of scientific theory but of scientific data.¹⁶ Certain such changes could force us to revise our view of the claim that water is H_2O and any such revision would undermine the plausibility of Putnam’s candidate type II identity.

I do not claim here that water is likely to turn out not to be H_2O . Rather, my point is that any appeal to empirical knowledge as regards bridging the justification-gap between type I and type II necessities faces the apparently recalcitrant problem of the principled defeasibility of scientific claims. If the Kripke-Putnam position is underpinned by the thought that there are points at which scientific hypotheses pass from being hypothetical to being written in metaphysical stone then we must ask at precisely which points that transition is accomplished? In other words, we should recognise here a surrogate for Quine’s original challenge to demarcate that class of type II identities which are properly understood to be necessary rather than contingent. Moreover, in this arena, any positive answer to the question would also tell us just when and where science could stop.

Two further important points are worth emphasis. We may very well agree that if a true type II identity is to be had then there is a valid inference to the necessity of that identity-statement. But the question remains: what assures us of the truth of the antecedent and thereby licenses *modus ponens*? Most sentences of this form are synthetic and license no such inference. At root, the problem is epistemic: under exactly which kinds of circumstance are we warranted in allowing that we have grasped the truth of any such identity-statement? Further, to the claim that the critical discussion of Kripke and Putnam in this section simply conflates epistemic necessity with metaphysical necessity, the rejoinder is that the candidate type II identity-statements considered here are supposed by both authors to motivate just that

¹⁴ *ibid.*

¹⁵ Laudan, L., ‘A Confutation of Convergent Realism’, reprinted in Papineau, D. (ed.) *The Philosophy of Science*, Oxford Readings in Philosophy, Oxford. 1996. Pages 107-138.

¹⁶ Kuhn, T.S., *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago, 1970. Second Edition. Chapter X, see especially pages 129-135.

bifurcation of modalities. Unless we can recognise that those statements do enjoy the special status claimed on their behalf the motivation to draw the modal distinction would appear to be lost.¹⁷

Ultimately, the soundness of the particular arguments presented here is less important than the fact that kinds of doubt can be raised as regards warrant for type II identity-statements to which type I identity-statements are immune. Given the fundamental epistemic differences between the two, type II identity-statements will always require special pleading re necessity. I have argued here that, in just this respect, the Kripke-Putnam case remains not proven.

Paul Tomassi

<p.tomassi@abdn.ac.uk>

¹⁷ I am again indebted to Timothy Kenyon for these points.

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 80-89

Gorgias the Sophist on Not Being:
A Wittgensteinian Interpretation

Copyright © by SORITES and Michael Bakaoukas

GORGIAS THE SOPHIST ON NOT BEING: A WITTGENSTEINIAN INTERPRETATION

Michael Bakaoukas

Provocative aphorisms of the most notable fifth century Sophist, Gorgias, such as «Nothing actually exists» or his realist tenet that «it is not speech (logos) that serves to reveal the external object, but the external object that proves to be explanatory of speech» (DK 82B 85) have been subject to endless cycles of interpretation. I do not propose to offer here a full analysis of Gorgias' thought, but rather some new suggestions as to how to interpret Gorgias on the basis of how scholars have interpreted Gorgias.

Untersteiner (1954: 163-5) and Kerferd (1981a: 93-95) distinguish between three approaches to Gorgianic texts. According to the first approach, Gorgias' treatise *On What is Not* is just a rhetorical parody of philosophical doctrines (philological or rhetorical approach) [Bux, 1941: 403 ff]. Following the rhetorical approach, researchers in speech communication and rhetoric attempt to attribute to Gorgias an epistemology and a genuine philosophy of rhetoric (Gronbeck, 1972: 36 — Engnell, 1973 — Enos, 1976 — Cascardi, 1983 — Walters, 1994).

According to the second approach, Gorgias is just a nihilist (or a negative dogmatic or a forerunner of scepticism) attacking the doctrines of the Eleatics and the Presocratics (ontological approach). There are many interpreters who hold that Gorgias is attacking the ontological doctrines of the Presocratics: Grote (1869: VII 331 ff and 1875: 107-109), Gomperz Th. (1901: 480-496), Maier (1913: 223-226), Reinhardt (1916: 39 ff), Joel (1921: 726), Nestle (1922: 554), Lattanzi (1932), Calogero (1932: 157-222), Brocker (1958: 438), Mondolfo (1936: 177-182), Levi (1941: 32-34 and 1966: 204 ff), Zeller (1963: 1305-1310), Sicking (1964: 225 ff), Guthrie (1969: 199 and 1971: ch. 11), etc.

According to Bakaoukas' Ph.D. dissertation, both Gorgias and Aristotle refer to the contradicting views of some presocratic philosophers who argue with each other about one and the same thing, i.e. the «being» (*on*). For Aristotle, «we cannot be right in holding the contradicting views [sc. of Heracleitus and Anaxagoras]. If we could, it would follow that contraries are predicable of the same subject [sc. which is not the case]» (*Metaph.* K 1063b24-26). In the same way, Gorgias says in his rhetorical work *Palamedes* that we should not believe those people who contradict themselves (*Pal.* 25). Obviously, the «quarrelling» philosophers at issue (in Gorgias' time) are the Atomists and the Eleatics. As far as we can tell from Gorgias' treatise *On What is Not*, the Gorgianic arguments and counter arguments refer to the Eleatics who had engaged in a controversy with the Atomists about being and non-being (or *kenon*).

In the third approach (which attributes to Gorgias an interesting philosophic position), Gorgias is seriously interested in the problems of predication and meaning (philosophical

approach) [Kerferd, Mourelatos, etc.]. According to Kerferd, «there is nothing in the treatise (sc. of Gorgias) which might not have been expressed by Gorgias in the fifth century and there the matter is perhaps best left [5] (...) there have indeed been those who have treated the work seriously. But its interpretation undoubtedly presents quite extraordinary difficulties, and those who have treated it seriously have arrived at very different views as to what Gorgias is saying [3]» (Kerferd, 1955: 3, 5). So, «what is needed, I believe, at the present stage of Gorgianic scholarship is a *programme* of discussion and research (...) — this is to identify certain broad philosophic features in Gorgias' thought in order to provide a kind of philosophic sketch-map» (Kerferd, 1981[b]: 322-3).

For Mourelatos the presocratic fragmentary texts encourage the study of commentaries and interpretations, that is to say the study of «secondary literature». The original works are lost, so one should «seek to come to terms with alternative views already on record» (Mourelatos, 1993: 1). This goes for Gorgias' paraphrased, fragmentary texts as well, i.e. Sextus' sceptical paraphrase of Gorgias' treatise *On What is Not* (2nd cent. AD) [hereafter DK B30] and the pseudo-Aristotelian paraphrase of Gorgias' *On What is not De Melisso Xenophane Gorgia* (2nd cent. BC) [hereafter MXG]

Following the third approach, Grote (1869: VII 331 ff) and then Joel (1921: 726) were the first scholars who attempted to interpret Gorgias *from a modern philosophical point of view*. They put forward a Kantian interpretation according to which Gorgias distinguished between the phenomenal and the noumenal. According to them, the Gorgianic word 'being' refers to ultra-phenomenal or noumenal object of which Gorgias denied the existence (not being). In this regard, according to Hamberger (1914: 53, 55), Plato misunderstood the Gorgianic theory on the relationship between noumenal and phenomenal objects. Newiger (1973: 186) emphasises the same interpretative line. But, as Chiapelli (1890) points out, we should translate the Gorgianic «not being» into «unknowable material Being», for the distinction between «phenomenal» and «noumenal» is foreign to all ancient thought before Plato (cf. Untersteiner, 1954: 164, n. 2).

A modern philosophical interpretation is also adopted by Mansfeld, stated as follows: «The point of Gorgias' argument seems to be that the only knowledge (...) is absolute or unqualified knowledge, i.e., knowledge of things as they are in themselves. Personal knowledge, being relative and not of things as they really are in themselves, is not knowledge in the required sense ... is not absolute knowledge of things as they really are but personal knowledge of things as they are experienced. This knowledge cannot be communicated to someone else» (Mansfeld 1985: 252). In this regard, Mansfeld (1985: 258) holds that «some of Gorgias' points (...) are philosophically immensely interesting because they deal with the problem of private vs public knowledge».

Recent interpretations of Gorgias' texts treat Gorgianic arguments as serious and valid. For example Schiappa and Hoffman say that «we ought to treat the *On What is Not* as a work of careful argumentation and not of inconsiderable philosophical significance» (Schiappa and Hoffman, 1994: 160). According to them, Gorgias refutes successfully the Parmenidean premise «if (A) can mention (O) or can think of (O), then (O) exists». Along this line of reasoning, Gorgias refutes the claim that what is thought of is necessarily existent (DK B3 79); that is he argues «against the existence of thought-about-objects» using a *reductio ad absurdum*. Namely, the Parmenidean premise «if (A) can mention (O) or can think of (O), then (O) exists» is refuted, for we can think of non-existent things like chimera or chariots running over the sea. Hence, there is no «identity relationship between things-thought-about and things-

that-are» (Schiappa and Hoffman, 1994: 157-8). For Barnes as well (1993: 171) Parmenides' premise is fallacious, «for Scylla and Chimera, and many non-entities are, as the Sophist Gorgias says, thought upon». Also it is noteworthy that logicians like Bochenski (1951: 17) and Thom (1986) take Gorgias' arguments into serious consideration.

As regards the relation between Plato and Gorgias it has been argued by Newiger (1973: 177-188) and Hays (1990: 336-7) that there are some important parallels between Gorgias' *On What is not* and Plato's *Parmenides*, *Meno*, *Theaetetus* and *Sophist*. These parallels have not yet been investigated in detail. As Calogero and Mansfeld point out, «there is not a systematic comparison of concrete parallels between Gorgias and Plato» (Calogero, 269 ff., 311ff; Mansfeld, 1985: 258, n. 48). In this respect, the philosophical implications of Gorgias' views at issue are very important for future studies, for in order to compare Plato's and Gorgias' arguments we should first examine Gorgias' own views (M. Angelini: 2).

In Hay's words (1990: 336-7), «it would seem prudent for scholars of Plato to re-acquaint themselves with the treatise and to keep in mind that Plato had to respond to these Gorgianic arguments». Crivelli (1996) who holds that the target of Plato's *Sophist* is Gorgias has done himself this to a certain extent. Many parallels between Plato's *Sophist* and Gorgias' works corroborate this approach. That is, Gorgias' example of «thinking of a non-existent entity» is «a flying man» (DK B3 79) which is reminiscent of Plato's example of «flying men» (*Theaetetus* 158b3-4; *Sophist* 263a8). Also Gorgias' treatment of the contradictory and contrary properties (DK B3 67, 80) is reminiscent of the sophistic argument in the *Sophist* (240b5, 240d6-8, 257b3-4, 258e6). Furthermore, Gorgias' arguments «had posed formidable challenges to Eleatic philosophy, and (...) [Plato's] quest for forms was particularly vulnerable to the same arguments, because its ontological assumptions were similar to those of Eleatics» (Hays, 1990: 336).

Researchers in Gorgianic scholarship recognise the philosophical significance of Gorgias. In this respect, Mourelatos and Kerferd attribute to Gorgias a sophisticated theory of meaning and interpret Gorgias in terms of a theory of meaning. Mourelatos detects in Gorgias' *On What is Not* and *Helen* a behavioural account of the nature of meaning as opposed to a referential and an ideational one. For Mourelatos, «Gorgias attacks two captivating conceptions of the nature of linguistic meaning, viz., that meaning is reference, and that meaning is mental image or idea. The attack is in the form of a series of puzzles. These are by no means sophistic in the pejorative sense. Indeed, the puzzles have recurred in the history of philosophy and have specifically played a role in the development of twentieth-century philosophical analysis» (Mourelatos, 1987: 136). So, according to this linguistic interpretation, Gorgias has an interest in questions such as «do words acquire their meaning by their association with external things, ideas, or behaviour?» or «what do words stand for?»

The third part of the *On What is not* (DK B3 83-87; MXG 980a18-b19) refers to «incommunicability» (But even if they are known, how could anyone communicate them to another? (MXG 980 a19-20; cf. 979a11-14; DK B3 65). According to Mourelatos, Gorgias here deals with «the inability of *logos* (speech) to communicate reality to another person, (...) [for] *logos* cannot furnish, constitute, or represent the external reality» to the effect that communication is undercut (Mourelatos, 1987: 138). Gorgias states «incommunicability» as follows: «Thus (...) since the existent subsists externally, it will not become our speech; and not being speech it will not be made clear to another person» (DK B3 84).

But does Gorgias actually undercut intelligent verbal communication? This is not actually the case, since for Mourelatos «if both speaker and listener have seen (or heard, as the case may be), the thing to which the speaker's words refer communication should be perfectly possible after all» (Mourelatos, 1987: 139-140). For Kerferd as well, «such communication is impossible *unless the listener has himself seen the visible object*, [so that] one man can learn from another» (Kerferd, 1981b: 324).

However, as Gorgias put it, «*and the speaker speaks, but he does not speak a colour or a thing. Anything, then, which a man has not in his own consciousness, how can he acquire it from the word from another, or by any sign which is different from the thing except by seeing it if it is a colour, or hearing it if it is a sound?*» (MXG 980 b2-8 tr. Hett). Gorgias herewith attacks a referential theory of meaning according to which «if words are to have meaning, they must refer to things in the real (at least extra-linguistic and perhaps also extramental) world» (Mourelatos, 1987: 151). He says bluntly that «the speaker speaks not a colour nor a sound, nor any other thing; he speaks **logos** (combining lines 980b2-3 and b6). Blunt, even simplistic though the formulation may be, the argument is by no means trivial. As an elenchus of the referential conception, the argument has fully as much force as the refined modern version of it: we do not eat the meaning of 'cake'» (Mourelatos, 1987: 153). Furthermore, in Mourelatos' view, Gorgias herewith objects to an empiricist, ideational conception of meaning according to which words have or acquire their meaning by «some sort of tie or pairing with perceptions (sensory impressions or mental images or thoughts)» (Mourelatos, 1987: 146, 151).

For Mourelatos (1987: 145), Gorgias possesses the concept of mental image or sensory impression (*Helen* 17: image of the things that are seen); on the basis of this concept, Gorgias uses the argument from perceptual identity or perceptual sameness to show that sensations, sensory impressions or mental images (**eikones**) are not the same to different observers and in different perceptual conditions. So given the assumption that meaning is mental image, «there would always be doubts as to whether a given word has the same meaning when used by different speakers, or when used by the same speaker at different times» to the effect that intelligent verbal communication would be impossible (Mourelatos, 1987: 154).

That is, as Wittgenstein would put it, if words acquire their meaning by their association to sensations, we are under the spell of a misguided metaphysics, for «when we think about the relation of objects to our experiences of objects (...) we are tempted to conceive of two distinct kinds of worlds the mental and the physical (...) It is against such temptations that the private language argument is directed. But concentrating on S (sensation) while enunciating 'S' (a word) does not bring it about that I will remember that 'S' means S, unless concentrating on S will transform the [verbal] sound 'S' into the expression of a concept. If it does not, then subsequent enunciations of 'S' will be empty noises, 'whatever is going to seem right to me is right' (*Phil. Inv.* 258), for no standard has been established by reference to which the subsequent use of 'S' can be evaluated as correct or incorrect» (Hacker, 1972: 223-4).

That is to say, sensations as such are meaningless. Consequently, if words acquired their meaning by their association to sensations, we would be in a state of «incommunicability» and «meaninglessness» in which, according to Gorgias: «Even if anything is apprehensible, yet of a surety it is inexpressible and incommunicable to one's neighbour» (DK B3 65), and «no one could say anything false» (MXG 980a11). In the same way, Lazerowitz and Mourelatos use

Wittgenstein's lectures on private experience and sense data in interpreting Gorgias (Lazerowitz, 1968: 37, Mourelatos, 1987: 154-5, n. 45).

Gorgias also emphasises the common sense fact that it is not necessary for many observers to see exactly the same thing at the same time (MXG 980b 9-19). In this case, he says that there is nothing to prevent a thing from seeming different to many persons (MXG 980b12), since these persons are supposed to be neither exactly the same nor probably on the same vantage ground (MXG 980b12-13). He simply points out that two persons can perceive the same object differently, and thus there may be two different appearances of the same object (MXG 980b10-11). This Gorgianic argument is a simple formulation of the argument from illusion according to which: «a physical object may at the same time look A to one observer and B to another observer; but it cannot be both A and B, for that would be a self contradiction» (Hirst, 1959: 46).

In this case, Gorgias says, it is difficult for someone to have exactly the same sense-experience with somebody else's sense-experience of the same thing. As Gorgias put it, «*for it is impossible for the same thing to exist in several separate persons; for the one would be two ... there is nothing to prevent it from not being the same in them all, seeing that they are not in every way alike, nor in the same place; for if anything were this, it would be one and not two [...] so that one man can hardly perceive the same thing as another*» (MXG 980b10-18).

In modern terms, what Gorgias says is simple. That is, if one physical object has two different appearances, when perceived by two different persons, then what could explain its phenomenal duplication is the possibility of there being two objects with two different appearances — which, as he says, is absurd, since the one thing in question would be two different things (MXG 980b12-13). Therefore, what changes appearances should not be the object itself, but the sensible things (**aistheta**), which vary from man to man. Sensible things are as many as the percipients, they are subjective, private to their owners, unobserved by others, and consequently they cannot be identified with the unique thing. In consequence, our subjective sense-experiences of a single thing and the thing itself are regarded as two separate items (Bakaoukas, 2001).

To interpret this passage Mourelatos adopts a phenomenological reading. He uses the argument from perceptual identity. For Mourelatos, «the 'one' and the 'same' which cannot become 'two' and 'different' is not an external third thing; it is simply the perception or experience or thought» (Mourelatos, 1987: 143). So Gorgias formulates an epistemological puzzle: «even we should allow that the same external thing should somehow also *be* 'in' two knowing subjects, it need not *appear* the same to them, because the two subjects are differently constituted and differently placed» (Mourelatos, 1987: 143).

So the question raised by Gorgias is «how can two minds have the same perception?» or «is perceptual identity or sameness possible?» According to Mourelatos, to solve this puzzle Gorgias uses a metaphysical device. He says that two different subjects do not have the same perceptions (**tauton**) but similar ones (**homoion**). He substitutes similar (**homoion**) for numerically the same (**tauton**). So, as Mourelatos put it, «and since similarity admits of degrees (...) perceptions may not be exactly similar, after all» (Mourelatos, 1987: 144). So, for Mourelatos, if we assume that meaning is mental (or sensory) image, «there would always be doubts as to whether a given word has the same meaning when used by different speakers, or when used by the same speaker at different times» (Mourelatos, 1987:154). This

phenomenological reading is justified by Kerferd's view that what concerns Gorgias is «the status of objects of perception ... with primary reference to phenomenal objects» (Kerferd, 1955: 5, 24).

For Mourelatos, «Gorgias has denied the proposition that language has the function of 'representing' or 'exhibiting', or 'setting forth'» (**parastatikos**) something that is extra-linguistic (first half of the concluding statement in section DK B3 85)» (Mourelatos, 1987: 160). The Sophist does not espouse an ideational theory of meaning either. So what is left is a behavioural theory of meaning. In Mourelatos' words, «it is rather uncanny how closely the vocabulary of section 85 resembles the vocabulary of modern behaviourist theory. External objects (...) 'fall upon us' or 'make an impact on us' or 'impinge upon us' (**prospiptonton, hypoptoseos**)» (Mourelatos, 1987: 163).

So, according to Mourelatos, Gorgias espouses a behavioural conception of meaning. Gorgias believes that a word has effect on other speakers of the language. For example, he says: «in response to the happy and unhappy occurrences affecting things and bodies, the soul comes itself to experience a certain emotion, through **logos**» (*Helen* 9, tr. Mourelatos, 1987: 156-7). For Mourelatos, this Gorgianic position «is an illustration of the conception of words as substitute stimuli (Mourelatos, 1987: 157). Furthermore, Gorgias compares the power of **logos** with that of drugs («just as different drugs draw different humours from the body (...) so too with **logoi**» *Helen* 14 tr. Mourelatos, 1987: 157). As Mourelatos put it, «if only we changed the archaic expression 'drawing out humours' to the behaviourist idiom of 'eliciting a physiological reaction' this sentence could just as well have been written by such advocates of the stimulus-response conception of meaning as Leonard Bloomfield, or B.F. Skinner, or C.L. Stevenson» (Mourelatos, 1987: 158).

Furthermore, Kerferd interprets DK B3 83-85 as follows: «communication is exclusively by means of speech or words, and the externally existing objects are not words. There is no possibility of converting things into words, and as a result there is no possibility of communicating things through, or by means of, words. This sets up an unabridged and unbridgeable gulf between words and things» (Kerferd, 1984: 218. Cf. Mazzara, 1983: 130 ff.). The text speaks clearly about words being ontologically different from things (Kyrkos, 1993: 299 — Jaekel, 1988 — Rodríguez-Adrados, 1981). Such a gulf or difference implies that a referential theory of meaning is ungrounded or at least that words are not «related to things as proper names — **onomata**» (Kerferd, 1984, 218). For Kerferd this passage proves (a) that Gorgias rejects «a referential theory of meaning- the view that words possess meaning, because they refer to (externally existing) things», and (b) that words, according to Gorgias, could not be used to communicate information about objects outside us, so that the possibility of communication by means of **logos** is eliminated (Kerferd, 1984: 218).

In addition, Gorgias says: «if anything exists, it cannot be known, and if it is known, no one could show it to another; because things are not words, and because no one thinks the same thing as another» (MXG 980b 17-19). For Kerferd this Gorgianic view posits a gap between the **logos** and the sense impressions or thoughts (Kerferd, 1981 [b]: 324). So, in Kerferd's view, «Gorgias has introduced a decisive breach into the relation between words and things, and by so doing also between words and sense-impressions. Yet from Parmenides onwards it was part of the received wisdom that words must refer to something (...) all thinkers in the fifth century BC were still imprisoned in the constraints imposed by the search for a referential theory of meaning; (...) in default of any other possible objects of reference for words [Plato] ended up by proposing fresh entities, the Platonic Forms. No such solution was

available to Gorgias. The furthest that he was able to go was to suppose that it is thoughts in our minds which function as objects of reference» (Kerferd, 1981 b: 325-6).

To conclude, as shown, scholars have to deal with multiple frames of interpretation before they can offer any settled account of what Gorgias meant to say to his audience.

BIBLIOGRAPHY

Primary Sources

- Bekker, I. (1960), *Aristotelis Opera*, ex rec. I Bekker-Gigon O. (ed.), 1960.
- Buchheim, Th. (1989), *Gorgias von Leontinoi. Reden, Fragmente, Tesimomien*, Hamburg, 1989.
- Bury, R.G. (trs.), *Sextus Empiricus*, I-II, Loeb, Cambridge-Massachusetts, 1935.
- Cassin, Barbara (1980), *Si Parménide: Le traité anonyme De Melisso Xénophane Gorgia, Édition critique et commentaire*, Lille, 1980.
- Diels, H.-Kranz, W. (DK), *Die Fragmente der Vorsokratiker*, vol 2., Berlin, 1960-1.
- Hett, W.S. (trs.), *Aristotle, Minor Works*, Loeb, Cambridge-Massachusetts, 1936.
- Loveday, T.-Forster, E. S. (trs.), *De Melisso Xenophane Gorgia*, The Works of Aristotle under the editorship of Ross, vol. 6, Oxford, 1913.

Secondary Sources

- Angelini, Marco, (2001), *Interpreting the Elder Sophists*, Radical Academy, The Philosophy Resource Center, 2001, 1-12.
- Bakaoukas, Michael (1995), I. 'The Argument from Illusion in Gorgias' Treatise On What is Not. II. 'Phroneisthai': Gorgias on Perception and Knowledge. III. Interpreting Gorgias from an Epistemological Point of View, MSc Diss., The University of Edinburgh, Edinburgh, 1995.
- (2001), *Gorgias' treatise on what is not or on nature. An epistemological analysis. The ontological controversy between the Eleatics and the Atomists and the intervention of Gorgias*, Ph.D. Diss., Univ. of Athens, Department of Methodology, History and Theory of Science, Athens, 2001.
- Bochenski, I.M. (1951), *Ancient Formal Logic*, Amsterdam, 1951.
- Brocker, W. (1958), «Gorgias contra Parmenides», *Hermes* LXXXVI (1958), 425-440.
- Bux, E. (1941), «Gorgias und Parmenides», *Hermes* 76 (1941), 393-407.
- Calogero, G. (1977, 1932), *Studi sul Eleatismo*, Florence, 1977 (1st edition Roma, 1932).
- Cascardi, A.J. (1983), «The Place of Language in Philosophy; or the Use of Rhetoric», *Philosophy and Rhetoric* XVI (1983), 217-227
- Chiapelli, A. (1890), «Per la storia della Sofistica Greca», *Archiv fur Geschichte der Philosophie* 1890, 1-21
- Crivelli, P. (1996), «The Argument from knowing and not knowing in Plato's Theaetetus (187E5-188c8)», *Proceedings of the Aristotelian Society* XCVI (1996), 177-196.

- Engnell, R.A. (1973), «Implications for Communication of the Rhetorical Epistemology of Gorgias of Leontini», *Western Speech* 37 (1973), 175-184.
- Enos, R. Les (1976), «The Epistemology of Gorgias' Rhetoric: A Re-Examination», *The Southern Speech Communication Journal* 42 (1976), 35-51
- Gomperz, Th. (1901), *Greek Thinkers*, trs. L. Magnus, vol. I, London, 1901.
- Gomperz, H. (1912), *Sophistik and Rhetorik*, Leipzig/Berlin, 1912.
- Gronbeck Bruce, E. (1972), «Gorgias on Rhetoric and Poetic: a Rehabilitation», *The Southern Speech Communication Journal* 38 (1972), 27-38.
- Grote, G.A. (1869), *A History of Greece from the Earliest Period to the Close of Generation Contemporary with Alexander the Great*, vol. VIII, London, 1869.
- (1875), *Plato and other Companions of Sokrates*, vol. I, London, 1875
- Guthrie, W.C.K. (1965, 1969), *A History of Greek Philosophy*, Cambridge, vol. II, 1965 and III, 1969.
- (1971), *The Sophists*, Cambridge, 1971.
- Hacker, P.M.S. (1972), *Insight and Illusion. Wittgenstein on Philosophy and the Metaphysics of Experience*, Oxford, 1972.
- Hamberger, P. (1914), «Die Rednerische Disposition in der alten *technē rhetorikē* (Korax-Gorgias-Antiphon)», *Rhetorische Studien* 2 (1914), 46-62.
- Hays, Steve (1990), «On the Sceptical Influence of Gorgias», *Journal of the History of Philosophy* 28 (1990), 327-337.
- Hirst, R.J. (1959), *The Problems of Perception*, London-New York, 1959.
- Jöel, K. (1921), *Geschichte der antiken Philosophie*, vol. I, Tübingen, 1921.
- Kerferd, G.B. (1955-6), «Gorgias on nature or that which is not», *Phronesis* I (1955-6), 3-25.
- (1981a), *The Sophistic Movement*, Cambridge, 1981
- (1981b), «The Interpretation of Gorgias' Treatise *On Nature or on what is not*, *Deukalion* 36 (1981), 319-327.
- (1984), «Meaning and Reference: Gorgias and the Relation between Language and Reality», in *The Sophistic Movement*, 1st International Symposium on the Sophistic Movement, Greek Philosophical Society, Athens, 1984, 215-222.
- Kyrkos, A. Bas. (1993), *The Ancient Greek Enlightenment and the Sophistic Movement*, Athens, 1993 (2nd edition — in Greek).
- M. Lazerowitz, M. (1968), *Philosophy and Illusion*, London, 1968.
- Levi A. (1941, 1963), «Studi su Gorgia», *Logos* XXVI (1941), 154-191 [reprinted in Levi, A., *Storia della Sofistica*, Morano, Napoli, 1963, 223 ff].
- Maier, H. (1913), *Sokrates, sein Werk und seine geschichtliche Stellung*, Tübingen, 1913.

- Mansfeld, J. (1985), «Historical and Philosophical Aspects of Gorgias' On What is Not», in Montoneri-Romano (eds), *The proceedings of the International Congress on Gorgias*, Catania, 1985, 1985, 243-271.
- (1988), «De Melisso, Xenophane, Gorgia. Pyrrhonizing Aristotelianism», *Reinisches Museum für Philologie XXXI* (1988), 239-276.
- Montoneri, L.-Romano, F. (eds) [1985], *Gorgia e la Sofistica: Atti del convegno internazionale*, Lentini-Catania 12-15 dicembre 1983, Catania, 1985 [33 papers presented at a conference on Gorgias].
- Mourelatos, P.D.A. (1987), «Gorgias on the function of Language», *Philosophical Topics XV* (1987), 135-170 [expanded version of a paper published in Montoneri-Romano (eds), *The proceedings of the International Congress on Gorgias*, Catania 1985. An abridgement, in modern Greek, of this paper was also published in the proceedings of the 1st International Symposium: *The Sophistic Movement*, Athens, 1984, 223-231].
- (1993), *The Pre-Socratics*, Princeton University Press, 1993 (2nd ed)
- Nestle, W. (1922), «Die Schrift des Gorgias 'über die Natur oder über das Nichtsein'», *Hermes* 57 (1922), 551-562.
- Newiger, H.J. (1973), *Untersuchungen zu Gorgias' Schrift über das Nichtsein*, Berlin, 1973.
- Reinhardt, K. (1916), *Parmenides und die Geschichte der griechischen Philosophie*, Bonn, 1916.
- Rodríguez-Adrados, F. (1981), «La teoría del signo en Gorgias de Leontinos», in Geckeler, H. (ed.), *Logos Semantikos. Studia linguistica in honorem Eugenio Coseriu*, Berlin-Madrid, 1981.
- Schiappa, H.-Hoffman, St. (1994), «Intertextual Argument in Gorgias' On What is Not: A Formalization of Sextus, Adv. Math. 7.77-80», *Philosophy and Rhetoric* 27 (1994), 156-161.
- Thom, P. (1986), «A Lesniewskian Reading of Ancient Ontology: Parmenides to Democritus», *History and Philosophy of Logic*, 7 (1986), 155-166
- Untersteiner, M. (1954), *The Sophists*, trs. Kathleen Freeman, Oxford, 1954.
- Walters, D.F. (1994), «Gorgias as a Philosopher of Being: Epistemic Foundationalism in Sophistic Thought», *Philosophy and Rhetoric* 27 (1994), 143-155
- Wittgenstein, L. (1953), *Philosophical Investigations*, trs. G.E. M. Anscombe, Oxford, 1953.
- Zeller, E. (1963), *Die Philosophie der Griechen*, Erster Teil, Zweite Abteilung, Hildesheim, 1963.
-

Acknowledgements: This paper was originally written at the University of Edinburgh and at the University of Athens under the supervision of Dr Theodore Scaltsas and Prof. Basilios Kyrkos to whom I am indebted.

Michael Bakaoukas MSC, PHD

University of Piraeus

6 Poseidonos, Hagios Nikolaos, 18900, SALAMIS, GREECE

GREECE

Tel/fax.: 03014651273

bakaoukas_michael@lycos.com

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349

Issue #13 — October 2001. Pp. 90-98

Privacy, Individuation, and Recognition

Copyright © by SORITES and Michael Ming Yang

PRIVACY, INDIVIDUATION, AND RECOGNITION

Michael Ming Yang

In this paper I examine Wittgenstein's private language argument and Ayer's counter argument. (1) I indicate that central to the language-game approach in general and the private language argument in particular is the thesis that social praxis constitutes the irreducible metaphysical reality from which the meaning of linguistic signs originates. I argue that the language-game approach is of *transcendental* character in the sense that it concerns the logical structure of human activity, which underlies concrete linguistic practices as well as operations of consciousness. Failure in recognizing this results in much confusion. (2) I demonstrate that the key issue concerning private language is, insofar as the argument goes,¹ *not* the problem of *correctness* of identification as commonly believed, but the social nature of individuation. (3) I conclude that sensation can only be recognized through the network of human action, and if one believes that sensation (assisted by memory) can be the *sole* basis upon which language and knowledge are maintained, then certain preferences on privacy as well as skepticism seem to be unavoidable.

§1. The Transcendental Characteristic of the Approach

«What has to be accepted, the given, is... *forms of life*» (226e²). The term «form of life» signifies the modes or structure of systematically understood social *behavior*, of human *activity*.³ Form of life, according to Wittgenstein, is the ultimate reality upon which all the possible modes of intentionality should be grounded. Language-game, in Wittgenstein's use of the term, is «the whole, consisting of language and the *actions* into which it is woven» (7). «The term 'language-game' is meant to bring into prominence the fact that the speaking of the language is part of an *activity*, or of a *form of life*» (23).

It should be emphasized from the outset that the language-game approach is of *transcendental* character. Wittgenstein never intends to turn linguistic analyses into an

¹ There are some reasons due to which private language is logically impossible. For instance, both language and consciousness are social constructs, which do not allow of any elements of logical privacy. But this is not a part of Wittgenstein's original argument.

² In making references to the first part of *Philosophical Investigations*, I shall just indicate section numbers; e.g. (23) means section 23. To the second part, I shall give page numbers. E.g., (226e) means page 226 of Anscombe's English translation (Basil Blackwell, Oxford 1963). Italics mostly are mine.

³ *Not*, as alleged by Kripke, «*the set of responses* in which we agree, and the way they interweave with our activities» (Kripke 1982, p.96).

empirical study like pragmatics, although the latter itself is a worthwhile topic for philosophers. He is never interested in, e.g., what is speaker's intention in uttering a word/sentence. What is of primary philosophical significance for the language-game approach, is the transcendental that underlies variegated uses of linguistic signs. Here the term «transcendental» does *not* signify any a priori principle or any innate linguistic capacity⁴. What should be reckoned as transcendental, the only thing that is given and has to be accepted, is the logical structure of human social activity, i.e. form of life. There are numerous ways in which one word/sentence can be employed and numerous speakers' meanings or intentions that can be attached onto a word/sentence. What is interesting for the language-game approach is how uses of words are embedded in form of life⁵, and it is a problem concerning the relation between a usage and the logical structure of human activity, not the usage itself, which is of empirical nature. I call the former inquiry «transcendental» in the sense that it concerns, given a form of life, the preconditions for possible applications of linguistic signs and possible operations of consciousness⁶.

What signifies the transcendental characteristic of the approach, in Wittgenstein's own terminology, is «depth grammar». At an early date of his later period (in June 1931) he entitled one of his book «*Philosophical Grammar*»⁷. According to him, the language-game approach «is directed not towards [linguistic] phenomena, but ... towards the 'possibility' of phenomena ... Our investigation is therefore a grammatical one» (90). «These are, of course, *not empirical problems*; they are solved, rather, by looking into the *workings of our language*» (109), note, not by probing into the *workings of speaker/hearers*, not by digging into past, present or future *intentions of utterers*. This transcendental approach is concerned with the depth grammar (as opposed to surface grammar, i.e. the way words are used in the construction of sentences) (664). Nevertheless, it is not a Kantian or *Tractarian* one; it is *not* «a final analysis of our forms of language» (91).⁸ The depth grammar, i.e., what is

⁴ In *Prolegomena* section 39: «*Of the System of the Categories*» Kant tells us that the discovery of transcendental categories «presupposes neither greater reflection nor deeper insight, than *to detect in a language the rules of the actual use of words* generally, and thus *to collect elements for a grammar*. In fact both researches [grammar and epistemology] are very nearly related, even though we are not able to give a reason why each language has just this and no other formal constitution, and still less why an exact number of such formal determinations in general are found in it».

⁵ This is a question Wittgenstein raised at the beginning of his later period. «Is meaning then really only the use of a word? Isn't it the way this use meshes with our life?» Wittgenstein 1974, section 29, p.65.

⁶ I'm definitely *not* the first one who uses the term «transcendental» in this fashion. K-O. Apel, for example, used it in a similar way. See his *Towards a Transformation of Philosophy*, Routledge & Kegan Paul Ltd., 1980.

⁷ See *Note in Editing*, Wittgenstein 1974, p.487.

⁸ See also 92 and 97. «We ask: 'What is language?', 'What is a proposition?' And the answer to these questions is to be given once for all; and independently of any future experience»(92). «Thought is... the a priori order of the world: that is, the order of possibilities, which must be common to both world and thought ... It is prior to all experience, must run through all experience»(97). These are (roughly) Kantian and *Tractarian* views.

transcendental for individual employment of words, is itself contingent upon particular form of life⁹.

§2. The Problem of Individuation

Let us now follow Quine and say, «there is no entity without identity.» Sensation words, like any other kinds of words, come into language through communal praxis and they do not have any private reference. If there are any objects of words that can be private, then the question of identity criteria inevitably arises. Wittgenstein addressed this question in 253.

However Wittgenstein's emphasis on the *correctness* of identification in 258 overshadowed the 253 argument. This emphasis is misplaced.

Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign 'S' and write this sign in a calendar for every day on which I have the sensation. — I will remark first of all that a definition of the sign ['S'] cannot be formulated. — But still I can give myself a kind of ostensive definition. — How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation — and so, as it were, point to it inwardly. — But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign. — Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connection between the sign and the sensation. — But «I impress on myself» can only mean this process brings it about that I remember the connection *right* in the future. But in the present case I have no criterion of correctness. One would like to say whatever is going to seem right to me is right. And that only means that here we can't talk about «right». (258)

Here for the correctness argument we have, (A) «I remember the connection right in the future», (B) «There is no criterion of correctness», (C) «Whatever is going to seem right to me is right», and (D) «We can't talk about 'right'».

Suppose I name my sensation S_{t_1} (occurs at time t_1) «S», and I then apply the name «S» to my sensation S_{t_2} (occurs at time t_2), then to $S_{t_3}, S_{t_4} \dots S_{t_n}$ (occur at $t_3, t_4 \dots t_n$), all of which are similar to each other. In the process of applying «S» to $S_{t_2}, S_{t_3}, S_{t_4} \dots S_{t_n}$, is there any *connection* between «S» and S_{t_1} that I have to remember rightly? Wittgenstein would say it is the naming relation between «S» and S_{t_1} . At any rate (A) means that I memorize S_{t_1} and recall it in the future accurately.

In order for understanding (B), we have to know what this *criterion of correctness* could be. In public language the conventional rules, including identification criteria and paradigms, are criteria of correctness. In private language, if I apply «S» to $S_{t_2}, S_{t_3}, S_{t_4} \dots S_{t_n}$, there is no publicly accessible paradigm by which we can check each of these applications is made right or wrong, in terms of how similar S_{t_n} is with S_{t_1} . (B) and (C) virtually assert the same thing and they indicate that 258 assumes that S_{t_1} is *analogously* a private paradigm. The fact that we cannot apply this private paradigm or private criterion of correctness to later occurrences of sensation amounts to that we simply don't have it as a paradigm.

⁹ This is a thesis that cannot be explored here. Ayer once declared: «[I do not] seek to deny that, as a matter of fact, one's references to one's private experiences are made within the framework of a public language. What I am querying is Wittgenstein's assumption that this is a logical necessity» (Ayer 1985, p.74). The private language argument is a *transcendental* argument, not a logical argument in the ordinary sense. A short answer to Ayer's above challenge is, the only thing that is given and has to be accepted is form of life. Cf. the above note 1.

The correctness argument fails on two accounts. Firstly, it relies on contingency of human memory capacity. Secondly and more importantly, it misrepresents the logical structure of recognition act.

If Robinson Crusoe (assuming he was left alone, having not yet learned to speak, nurtured by some animal) possesses a very extraordinary capacity in memorizing sensation, then he should be fully capable of inventing a private language. Whether or not other people have access to his private paradigm would make no difference: he doesn't need public criteria of correctness. One may argue that «it is not possible to obey a rule 'privately'» (202). Yes, but in the current context this is the very thesis that one should establish and therefore it cannot serve as a premise.

For the second point, let us come back to one of Wittgenstein's 1930's texts, where he performed a meditation on recognition.

- [1] How do I know that the color of this paper, which I call 'white', is the same as the one I saw here yesterday? By recognizing it again; and recognizing it again is my only source of knowledge here. In that case, 'That it is the same' means that I recognize it again.
- [2] Then of course you also can't ask whether it really is the same and whether I might not perhaps be mistaken; (whether it is the same and doesn't just seem to be.)
- [3] Of course, it would also be possible to say that the color is the same because chemical investigations do not disclose any change.
- [4] Recognition is what is primary and identity what is secondary.¹⁰

[4] indicates the nature of this discussion: it belongs to the *transcendental* discourse: seeing from the *empirical* angle, identity is primary and recognition is secondary. From the transcendental perspective, it is community's recognition that decides an object's identity criteria, which in turn, makes individual's (empirical) recognition possible.

[3] indicates that the community's *agreement* on choice of paradigm, which establishes identity criteria, is primary. For example, we can choose, or not choose, chemical investigation as the final decision procedure of identification.

[1] concerns situations in which a sample, an instance of paradigm, is not available, e.g. the white paper placed here yesterday has been destroyed, and we only have memory to rely upon. Or, a sample still exists but according to our memory the color has changed (56). Under these circumstances, as [2] makes explicit, it *does not make sense*, to talk about right and wrong or talk about whether or not two color-appearances are *really* the same. And they are not even needed: identification hangs on practical purposes. [2] is a straightforward denial of the correctness argument.

Whether or not a publicly accessible paradigm is available is largely irrelevant, and even if it is available, S_{12} , S_{13} , S_{14} ... S_{1n} still have to be identified through senses.¹¹ This undermines the correctness argument. Note that the correctness argument rests upon the assumption that there is no publicly accessible paradigm by which we can check each of applications of «S»

¹⁰ Wittgenstein 1975, sections 16 and 19, pp.60, 61. Numbering is mine.

¹¹ *Only* in this connection Ayer is correct in saying that the distinction between public and private is idle.

is made right or wrong, *in terms of how similar each S_m is with S_{11}* . Here the problem is rather the range of similarity from S_{t1} to S_{tn} and this is a problem of individuation. Whether S_{t1} , S_{t2} , S_{t3} , S_{t4} ... S_{tn} are private or public makes no difference.

The (*transcendental*) question is, *how* a rule, or an identification criterion, is decided. Central to the private-language case is the problem concerning the *social* nature of individuation. We don't know how a private criterion of identity is generated.

«Another person can't have my pains.» — Which are *my* pains? What *counts* as a criterion of identity here? ...I have seen a person in a discussion on this subject strike himself on the breast and say: «But surely another person can't have THIS pain!» — The answer to this is that one does not define a criterion of identity by emphatic stressing of the word «this». Rather, what the emphasis does is to suggest the case in which we are conversant with such a criterion of identity, but have to be reminded of it. (253)

The decision on identification criteria is an essential part of a linguistic community's form of life. What counts as an individual, the whole body of the rabbit, or each part of it, or a temporal slice of it, depends upon our *collective operational relation* with the object, depends upon the object's status in our way of life. In the case of sensation, e.g., as long as doctors treat your pain and my pain in the same or even similar way, two pains should count as the same (kind). What counts a kind, a type, an identity criterion, is hinged upon our way of life.

Only «if I assume the abrogation of the normal language-game with the expression of a sensation, I need a criterion of identity for the sensation; and then the *possibility* of error also exists» (288). Otherwise «to use a word without a justification does not mean to use it without right» (289). Given the premise that the argument is of transcendental nature, section 288 provides an exact answer to Kripke's question (regardless of what kind of interpretation he is imposing on Wittgenstein):

How can I possibly have any difficulty identifying my own sensation? And if there were a difficulty, how could 'public' criteria help me?

Surely I can identify these [sensations] after I have felt them, and any participation in a community is irrelevant!

It seems to me that we have sensations or sensation qualia that we can perfectly well identify but that have no 'natural' external manifestations.¹²

If I have no difficulty in identifying my sensation, it is due to the fact that at the *empirical level* sensation has already been individuated by our form of life! I agree with Kripke in that the view that an inner process always has outward criteria is «*empirically false*»,¹³ e.g. my feeling of pain may well be more reliable as an indicator of my illness than a CT scanning result is. But public criteria belong to the *transcendental* makeup of sensation language. It is in the transcendental sense that Wittgenstein says «That is not agreement in opinions but in form of life» (241), and «What people accept as a justification — is shown by how they think and live [i.e. what their form of life is like]» (325). Essential to the debate, as I shall show in the following section, is the difference in metaphysical position.

¹² Kripke 1982, pp. 60, 80, 103.

¹³ Ibid., note 83, italics are original, p.103.

§3. The Metaphysical Status of Sensation

In 1959 Ayer claimed that having tried to construct a language all of whose words refer to private things he believed that in any language which allows reference to individuals there must be criteria of identity which make it possible for different speakers to refer to the same individual.¹⁴ By Wittgenstein's standard this amounts to abandonment of the private language view, but Ayer didn't withdraw from the battle and in fact in his lifetime he never did.¹⁵ In 1973 Ayer declared that in the empiricist construction program (1) the observer was *not* permitted to conceive of the data with which she works as private to her and (2) the observer was *not* identified either with her or with any other person.¹⁶ If it is so, what is left with respect to privacy for what he termed as «a reformed Robinson Crusoe approach», i.e. a construction program developed by a single observer? Ayer says that this Robinson Crusoe approach is supposed «to do justice to the fact that any knowledge of the world which anyone acquires is bound to be based upon his *own* experiences [i.e. his own sensation and memory].»¹⁷

Ayer's central thesis is this: since the ultimate ground for language use is the individual's judgment upon her own sensation and memory *only*, the distinction between public and private objects is unfounded in the first place. This is Ayer's final position in his lifelong campaign against Wittgenstein's private language argument¹⁸. The thesis is so pivotal to the metaphysical foundation of those empiricist programs that it merits a careful exploration.

The crucial fact which it seems to me that Wittgenstein persistently overlooks is that anyone's significant use of language must depend sooner or later on his performing what I call an act of *primary recognition*. In Wittgenstein's example [265], it is supposed not to be sufficient for someone to check his memory of the time at which the train is due to leave by visualizing a page of the timetable. He has to check the memory in its turn by actually looking up the page... But unless he can trust his eyesight at this point, unless he can recognize the figures printed in the table, he will be no better off. If he distrusts his eyesight, as well as his memory, he can consult other people, but then he must understand their responses. He needs to identify correctly the signs that they make. The point I am stressing is not the trivial one that the series of checks cannot continue indefinitely in practice, even if there is no limit to them in theory, but rather that unless it is brought to a close at some stage the whole series counts for nothing. Everything hangs in the air unless there is at least one item that is straightforwardly identified.

¹⁴ Ayer 1959, p. 78.

¹⁵ See Ayer 1988, p.16. Ayer died in 1989.

¹⁶ Ayer 1973, p.98.

¹⁷ Ibid.

¹⁸ See above note 15. Kripke in his influential book on private language devoted two full pages to endorse this view, see Kripke1982, pp.60-2. E.g., «If I really were in doubt as to whether I could identify any sensations correctly, how would a connection of my sensations with external behavior, or confirmation by others, be of any help? Surely I can identify that the relevant external behavior has taken place, or that others are confirming that I do indeed have the sensation in question, only because I can identify relevant sensory impressions (of the behavior, or of others confirming that I have identified the sensation correctly). My ability to make any identification of any external phenomenon rests on my ability to identify relevant sensory (especially visual) impressions». Ibid., p. 61.

If this is correct, Wittgenstein is wrong in taking that the corroboration of one memory by another is an inferior substitute for some other method of verification. *There is no other method.* Whatever I have to identify, whether it is an object, an event, an image or a sign, *I have only my memory and my current sensation to rely on.* The difference lies only in the degree to which the memories are cross-checked.¹⁹

In this *primary recognition* argument Ayer commits two mistakes and both of them are essential to the empiricist vision of language and knowledge. The first is that he assumes that memory is capable of working independently and internally without any *external* assistance and corroboration, and the second is that he writes as if all the multifarious employment of memory and senses enjoy *equal* credibility and significance.

In the example of a timetable, if my eyesight is not reliable I will naturally appeal to that of others. This decision *presupposes* my trust upon my ear and understanding. If both my eyes and ears are not dependable I should e.g. purchase auxiliary equipment, and that again *presupposes* the effectiveness of modern technology, the reliability of manufacturer, and so on. But in no time should I belabor the universal doubt. It is not that a series of checks has to stop somewhere by virtue of our special trust upon any particular employment of senses, it is simply that no such an obsessive scrutiny is needed to be conducted by the assumed judges of senses. Only when the primacy and sufficiency of sense/memory are assumed such a scan and a primary recognition should be called for.

First, experience is a symphony of coordinate *actions* rather than a solo of successive sensations. Success, failure, stableness, anomaly, smoothness and incoherence etc. are the most salient parts in the organic whole of our experience. Empiricists are tempted to say that these are nothing but groups of sensory indices, the objection is that they *dominate* the whole web of experience in which sense/memory inquests are only scattered and dependent episodes. Action is metaphysically primary to sensation, and sensation *ultimately* should be understood through action, not the other way around.

Second, certitude pertaining to multifarious uses of senses is varying and thus should not be assigned indiscriminately. In the example of the timetable, the credibility of memory is different *in category*, rather than in degree, with that of senses. If I am weak in senses, then the credibility of them should be different *in category* with that of the aid of sensing-equipment plus my neighbor or colleague's contribution plus my capability of understanding. So on and so forth, reliability can never be accredited *equally* to all the contemporaneously active sectors.

It is the sensualist empiricist's dogma that human experience *only or ultimately* consists of sensation and memory. The fact that our judgments upon sensation and memory are coordinated and corroborated with our every-minute fulfillment in action and the dynamic equilibrium of the whole system of experience is so familiar to us that these factors are easily overlooked. When Kripke declares that «Surely I can identify these [sensations] after I have felt them, and any participation in a community is irrelevant (!)»²⁰ doesn't he commit to this

¹⁹ The first formulation appears in Ayer 1954. My quotation is from Ayer 1985 p.76 and 1982 pp.151-2. It is not surprising that Kripke's comment (on the 1954 version) is «the objection seems cogent». See Kripke 1982, p. 62.

²⁰ Kripke 1982, p. 80.

Ayerian weak form of private language view? It is indeed a difference between two metaphysical positions.

For sensualist empiricists social reality is nothing but a total sum of individual sense experiences. Ayer thus wrote in 1985:

The practice of the community is supposed to bestow meaning on my utterances. But *what is the community except a collection of persons?* And if each of those persons is supposed to take his orders about meaning solely from the others, it follows that none of them takes any orders. The whole semantic house of cards is based upon our taking in each other's washing, or would be if there were any laundries to wash. On this interpretation, Wittgenstein's argument, so far from proving that private languages are impossible, proves that they are indispensable.²¹

Ayer here appeals to the *empirical* notion of community. Indeed, a university, a nation, etc., consists of persons. But this is irrelevant to our question because the notion of social praxis is of *transcendental* one and should not be reduced to that of individual's practice. The *empirical* causal chain of learning doesn't count because the meaning of word (1) is inculcated into an individual's mind through her everyday-and-lifelong participation in social life and (2) has to be cashed out in every-minute social transaction.²²

Sense data do contribute a substratum to the linguistic meaning, and how these data are processed and integrated into the connotation of words should be properly accounted for.²³ But the question here is that if one chooses sensory input *alone* as the metaphysical basis for meaning, then certain commitment to privacy seems to be inevitable, and then skepticism as well seems to be unavoidable.²⁴

REFERENCES

Ayer, Alfred Jules

- 1954: *Can there be a Private Language*, in *Wittgenstein*, University of Notre Dame Press 1966.
- 1959: *Privacy*, in *The Concept of A Person*, St. Martin's Press 1963.
- 1973: *The Central Questions of Philosophy*, Penguin Books 1976.
- 1982: *Philosophy in the Twentieth Century*, Unwin Paperbacks 1984.
- 1985: *Wittgenstein*, Random House 1985.
- 1988: *A Defense of Empiricism*, in *A. J. Ayer Memorial Essays*, the Press Syndicate of the University of Cambridge 1991.

²¹ Ayer 1985, p. 74.

²² It's true that individuals always contribute novel meanings and ideas to a language but this is not a part of the problem.

²³ To solve this problem, a *constructive* theory of language formation is needed. Such a constructive theory can be found, e.g. in *Transcendental Idealism and Pragmatism*.

²⁴ In my development of the views on Wittgenstein's philosophy I owe very much to Dr. Arthur Falk of Western Michigan University for his seminars during 1998 and 2000.

Kripke, Saul

1982: *Wittgenstein on Rules and Private Language*, Harvard UP 1982.

Wittgenstein, Ludwig

1974: *Philosophical Grammar*, Basil Blackwell, 1974.

1975: *Philosophical Remark*, Basil Blackwell 1975.

Michael Ming Yang

312 Watson Hall, Queen's University

Kingston, Ontario, Canada K7L 3N6

<mchly@juno.com>

SORITES

An Electronic Quarterly of Analytical Philosophy

ISSN 1135-1349

COPYRIGHT NOTICE AND LEGAL DISCLAIMER

© 1996 The SORITES Team

Please, read!

- (1) **SORITES** is not in the public domain. In accordance with international Law (especially the Berne Convention for the Protection of Literary and Artistic Works and the Universal Copyright Convention), this issue of **SORITES** is Copyright-protected throughout the Planet.¹
- (2) The Copyright of this issue of **SORITES** taken as a whole is held by the electronic publisher (the SORITES team).
- (3) The Copyright of the papers published in **SORITES** is retained by the individual authors, except that: (i) no part of any such paper may be printed or displayed elsewhere or incorporated into a book, an anthology or any other publication of any sort until **SORITES** has accorded the author(s) permission to that effect [which will be done routinely and quickly, provided **SORITES** is therein clearly and explicitly mentioned as the primary source]; and (ii) the authors agree to abide by the other terms and conditions contained in this Copyright Notice.
- (4) The authors of the included papers and the electronic publisher, the **SORITES** team, — whether jointly or separately, as the case may be — hereby reserve all rights not expressly granted to other parts in this Copyright Notice.
- (5) In compliance with Spanish Law, this issue of **SORITES** has been legally registered, three diskette-copies being deposited with the competent authorities, namely the «Deposito Legal» office of the Autonomous Community of Madrid, c/ Azcona 42. (Legal Deposit Registration: M 14867-1995.)
- (5) A licence is hereby granted without fee for anybody to freely make as many unmodified copies as they wish of this issue of **SORITES** IN ITS INTEGRITY, give such copies to anyone, and distribute this issue of **SORITES** via electronic means or as printed copies, PROVIDED no part thereof is altered or omitted, and especially **NEITHER THIS COPYRIGHT NOTICE NOR THE COPYRIGHT BOXES ON TOP OF THE DIFFERENT PAPERS ARE REMOVED, AMENDED, OR OBSCURED.**
- (6) In this context, *the issue of SORITES as a whole* is meant to consist in: either (i) a single file (be it its official version as a WordPerfect 5.1 document or any unofficial version released by the **SORITES** team as an undivided file); or (ii) a collection of files produced by slicing one of the entire-file versions in order to facilitate handling, browsing or downloading. In the latter case, the conveyor is bound to distribute the whole collection. (In this context printed copies of this issue of **SORITES** are taken to be equivalent to electronic copies, their distribution being subject to the same conditions.)
- (7) This issue of **SORITES** may be sold for profit or incorporated into any commercial material only with the previous explicit consent granted by the SORITES team. Otherwise, no fee may be charged for its circulation. An exception is granted to non-profit organizations, which are hereby authorized to charge a small fee for materials, handling, postage, and general overhead.
- (8) Private copying of single papers by any lawful means is allowed only when done in good faith and for a fair use, namely for purposes of teaching, study, criticism or review; but no part of this issue of **SORITES** may be

¹ The reader may find an excellent discussion of copyright-related issues in a FAQ paper (available for anonymous FTP from rtfm.mit.edu [18.70.0.209] /pub/usenet/news.answers/law/Copyright-FAQ). The paper is entitled «Frequently Asked Questions about Copyright (V. 1.1.3)», 1994, by Terry Carroll. We have borrowed a number of considerations from that helpful document.

conveyed to another individual or to a gathering — whether in writing or through oral teaching or by any other means — unless the source is clearly and explicitly acknowledged.

- (9) In particular, no part of this issue of **SORITES** or of any paper therein included may be conveyed to others by means of reproduction, quotation, copy or paraphrase, without a clear and explicit acknowledgement of the issue of **SORITES** and its date, the author's name and the paper's full title. Whenever the quotation occurs within a publication, it is also mandatory to mention the official pages (as shown within the Copyright box on top of the paper), the ISSN (1135-1349) and the official home site of electronic display, namely <http://sorites.net>.
- (10) Any perpetration of, or complicity with, unfair use of copies or partial copies of this issue of **SORITES**, or of papers therein included, especially forgery or plagiarism — being, as it is, an infringement of the authors' and the electronic publisher's rights — is in any case a civil tort, but may even be a crime under current legislation.
- (11) This issue of **SORITES** is provided «as is», without any guarantee of any kind. The electronic publisher, the **SORITES** team, disclaims all warranties, whether expressed or implied, including, without limitation, the implied warranties of fitness for any particular purpose with respect to the papers included in this issue. By furnishing this document, the **SORITES** team does not grant any license or endorses any commitment except in so much as explicitly set forth in the present Copyright Notice.
- (12) The electronic publisher, the **SORITES** team, does not necessarily agree with the authors' views or arguments. The electronic publisher cannot certify the accuracy of any quotations or references contained in the papers.
- (13) Each author vouches alone for the originality of the papers they submit to **SORITES** and for their compliance with established Copyright laws. Acceptance of a manuscript is done in good faith under the assumption the originality claim is truthful. The electronic publisher — i.e. the **SORITES** team — does not pledge itself for the accuracy of such declarations.
- (14) The **SORITES** team cannot be responsible for any real or imaginary damages suffered as a result of downloading, reading, using, quoting or circulating any materials included in this issue of **SORITES**. The user assumes, at their own risk, full responsibility for the proper use of this issue of **SORITES**.
- (15) Downloading, reading or in any other way using this issue of **SORITES** or any part thereof entails full acceptance of the hereinabove stated terms and conditions. If, after downloading a file containing this issue of **SORITES** or a part thereof, a user fails to agree to the conditions and terms contained in this Notice, they must discontinue using the material and irrecoverably erase or destroy the downloaded file, so as not to occasion any third-part's unfair use thereof.
- (16) Although, thanks to a permission kindly granted by the system's administrators, this electronic journal is displayed at the internet host <www.ifs.csic.es/sorites/> (hostname whose current official IP is 161.111.210.19), which belongs to the Spanish institution CSIC, the journal is not published or sponsored or endorsed by the CSIC, the only owner and publisher being the **SORITES** team.
- (17) A specific licence is hereby granted for this issue of **SORITES** — and all future issues of the journal as well — to be freely displayed by any BBS and any Internet node or site, provided all conditions stated above are fully adhered to. No previous consent of the **SORITES** team is required for such a display, which may be in the form of FTP, Gopher, http-WWW or any other electronic means.

Madrid. April 10, 1995

(Updated: December 31, 1996; April 27 1998)

The **SORITES** Team

RELEASE NOTICE

This issue of *SORITES* is made available in several formats, but its only official version is that released with filename:

sorite13.wp

which is the only file within the archives: **sorite13.zip**. A PDF file (**sorite13.pdf**) is also released; it has been generated from the file **sorite13.wp**.

Although each version, whether official or not — as initially released today (2001-10-25) by the **SORITES** team — is an entire seamless file, it may be splitted down into chunks in order to facilitate downloading, browsing, transferring or e-mailing. In such cases, the unity of this issue of **SORITES** as a whole must be preserved by keeping the ensuing collection intact.