Access to Electronic Thesis

Author:	Matthew Alfred Joseph Jones
Thesis title:	Semantic Scepticism: Normativity and Naturalism
Qualification:	PhD

If this electronic thesis has been edited by the author it will be indicated as such on the title page and in the text.

<u>**Title Page**</u>

Semantic Scepticism: Normativity and Naturalism

Matthew Alfred Joseph Jones

A Thesis Submitted in Partial Fulfilment of the Requirements for the degree of

Doctor of Philosophy

(Philosophy)

Department of Philosophy, University of Sheffield, UK

October 2011

Thesis Summary

The present thesis is a response to Kripke's sceptical argument for the conclusion that there are no facts of the matter concerning what any speaker means by any expression. This conclusion gives rise a paradox: if true, it is meaningless, and if it is meaningless it cannot be true. I focus on two main topics. The first of these is the putative normativity of meaning, which is taken by Kripke to provide an *a priori* reason to reject any putative naturalistic or reductive theory of meaning. The second is the *naturalisation* of meaning. The sceptical challenge is ultimately to account for semantic facts in terms of non-semantic ones. It seems that this would be achieved through a successful naturalisation of the semantic relations.

I argue against the claim that meaning is normative in any sense that could provide an *a priori* argument against factualism about meaning. I then consider the most prominent naturalistic responses to Kripke's argument and find all of these to be unsatisfactory. I then attempt to provide a partial dissolution of the sceptical paradox by arguing that there are reasons to expect naturalising project in semantics to fail other than the truth of Kripke's sceptical conclusion. First, I contend that we are not in a position to know what the eventual resources of a naturalised semantics would be. Since our current science is incomplete and potentially subject to revision, it would be a methodological mistake to view any putative naturalising base for semantic facts as metaphysically fixed. Second, I argue that the semantic facts themselves do not constitute a tractable domain for scientific theorising. I conclude that we should neither be too perturbed by, nor draw sceptical conclusions too swiftly from, the failure of such naturalising projects.

<u>Contents</u>

For my parents

Acknowledgements

Abbreviations

PI   Wittgenstein, *Philosophical Investigations* (2001).

WRPL   Kripke, *Wittgenstein on Rules and Private Language* (1982).

Introduction

In our everyday lives, the idea that there are facts of the matter about what we mean by our expressions is commonplace. The consequences of ascribing determinate meanings to our own and others' utterances are by no means trivial. We base a great many important decisions on ascriptions of meaning. Consider the doctor who says 'you will need a course of antibiotics to get rid of that', the PhD supervisor who says 'that argument is invalid as it stands', or the electrician who says 'the job will cost around £250'. We take it that there is a fact of the matter about what is meant by these utterances, and act accordingly. Call this view – the idea that there are facts of the matter about what we mean – 'meaning-factualism'.

However, in his highly influential essay, *Wittgenstein on Rules and Private Language*, Kripke produced an astonishing argument to the outrageous conclusion that this is all illusory: there simply *are no* facts of the matter concerning what we mean by our expressions. The argument begins with the observation that if, say, Jones means *addition* by '+'[1], then there must be some fact about Jones that makes this the case. Call this a meaning-constituting fact. However, Kripke charges, on inspection no such fact may be found.

Importantly, Kripke places no restrictions on the domains in which we are allowed to look for the kind of fact at issue. Thus his scepticism is more radical than that of Quine (1960; 1969; 1979) who, in his discussion of the 'indeterminacy of translation', restricted the admissible evidence for an ascription of meaning to *behavioural* evidence. Quine argued that, were one to translate a foreign speaker's expressions into English, there would always be some empirically adequate yet incompatible translation manual that one could have used. His claim is that since there is no empirical fact that could determine which translation manual was the 'correct' one, meaning is therefore indeterminate. While one might retreat into talk of purely mental or introspectable facts to escape Quine's scepticism, such a move will not work against Kripke. For Kripke allows

---

[1] I will adopt the convention of specifying meanings/contents in italics.

any fact to be considered in our attempt to account for how meanings are constituted. Thus his claim is that *even if* an omniscient God looked into my mind, he would find nothing that constituted what I mean by the words I am now writing. Thus Kripke's scepticism is more damaging, and more radical than Quine's.

Kripke's argument proceeds by elimination under epistemological idealisations. First, full epistemological access is granted to one's past behaviour, and the entire history of one's 'inner' mental life. The crucial point is that if the disputed kind of fact cannot be found under such idealising conditions, the only possible reason for this is that there simply was no such fact to be found. Kripke then examines and rejects a number of possible candidate facts. Having found all of the responses he considers to be wanting, Kripke concludes that there are no facts of the matter about what we mean. This leaves us with the 'sceptical paradox': 'there are no facts of the matter about what we mean', if true, is meaningless; and if it is meaningless, it can't be true.

Kripke then proceeds to offer what he calls (following Hume) a 'sceptical solution' to the paradox. This begins by conceding that the sceptic's negative arguments are unanswerable. However, it then aims to show that 'our ordinary practice' of ascribing meanings, and our beliefs concerning meanings, are justified as they do not 'require the justification the sceptic has shown to be untenable' (Kripke [1982] 2003: 66). By contrast, a straight solution aims to directly answer the sceptic's negative contentions – it proves the scepticism to be 'unwarranted' (Ibid). In Kripke's sceptical solution, our practice of ascribing meanings, and our belief that expression have meanings, are justified, but *not* by there being facts of the matter about meaning (the justification shown to be untenable by the sceptical argument). Rather, beliefs about and ascriptions of meaning are justified by the role and utility they have in our everyday lives. Kripke argues that such ascriptions of meaning serve to admit speakers into one's 'linguistic community' and to mark them as *trustworthy* in their use of expressions.

A crucial feature of the sceptical argument is the implicit (and un-argued) *reductionism* within it (McGinn 1984: 150 f.). That is, at core the sceptical challenge is to account for meaning *in terms of something else*. There are two

related questions to be answered here. First, why should meaning by viewed as inherently 'suspect' – something in need of reduction, and, second, what are the kinds of facts that are 'respectable' enough to serve as a 'reduction base' for meaning? One answer is that of the metaphysical naturalist – roughly, someone who holds that all entities, relations, properties, and so on must ultimately be characterised in terms of the methods and ontology of the natural sciences. They would hold that semantic facts must reduce to such 'natural' facts to be ontologically respectable. They might also voice this worry: if we could not show how semantic facts are to be brought within the purview of a naturalistic metaphysics, then reality would not be, as it were, uniform. In fact, it is hard to see why one would be a reductionist about meaning if one were not also a naturalist. That is, if one had no particular metaphysical commitments concerning what could enter into one's ontology; one would presumably have no reason to view meaning-facts as ontologically suspect.

This thesis is an investigation into the prospects for a naturalistic response to the sceptic, and thus for a naturalistic theory of meaning in general. *Prima facie,* the principal barrier to a naturalised theory of meaning is an objection raised by Kripke in the course of the sceptical argument. This is the putative *normativity* of meaning. Exactly what the normativity of meaning might amount to is a major focus of this study. As a rough characterisation, we may say that the normativity thesis involves the claim that statements about the meanings of expressions, or, more accurately, about what *speakers mean* by expressions, either involve, or immediately imply *prescriptive* statements concerning how one *ought* to use one's words (Whiting 2007). The sceptical claim against the naturalist is that *no* natural fact has such implications.

There have been several suggestions in the philosophical literature as to how the claim that meaning is normative might be cashed out so as to provide some kind of *a priori* argument against factualism about meaning. I will argue that meaning is not normative in any of these senses. It might be thought that, once this barrier is removed, the way is open for a successful naturalistic theory of meaning. On inspection, however, this optimistic conclusion is unwarranted: even without the claim that meaning is normative, the sceptic has ample resources to shoot down *all* of the extant naturalistic theories of meaning on offer. From this

result, the pessimistic conclusion that the sceptic has won could be thought to follow.

However, I will argue that once one reflects further on the nature of metaphysical naturalism, on semantic facts, and on the nature of reduction more generally, it becomes clear that a naturalistic theory of meaning is likely to fail for reasons *other than* there simply being no facts of the matter about meaning. That is, the prospects for a naturalised semantics look pretty bleak, but this should, in fact, be expected. As it turns out, metaphysical naturalism is not a particularly coherent or stable position. In particular, the idea that semantic facts must be reduced to more respectable 'natural' facts is dubious. On the one hand, we do not have a clear conception of what the 'natural' facts are. That is to say, we do not have a firm 'naturalising base' to which meaning is to be reduced. On the other hand, semantic facts – particularly facts about the truth-conditions of declarative sentences – do not seem themselves to be of the right type to enter into *scientific* theorising. Once one draws the conclusion that semantics is not a scientific endeavour, calls to reduce semantic facts to scientific categories seem misplaced.

What implications does this have for a response to the meaning-sceptic? On the one hand, we should not be *too* perturbed by our failure to provide a workable naturalistic theory of meaning – there may be good reasons for this besides the truth of semantic non-factualism. On the other hand, this obviously does not constitute a philosophical account of the peculiar and distinctive properties of meaning. So, while I hope to remove some of the anxiety felt at the failure to bring semantic facts within purview of a naturalistic metaphysics, I do not claim to thereby provide a refutation of the sceptical argument.

In Chapter 2, I will outline the sceptical argument in more detail. We will see how Kripke utilises the claim that meaning is normative as a block to the promising dispositional response to his scepticism. I will also draw attention to the reductionism that is implicit in the sceptical argument, and note the requirements on a satisfactory *non*-reductionist response. While I do not claim to provide such a response, it is important to bear in mind what would be required in order to refute the sceptic. It should be stressed that my account of reasons behind the failure of naturalising projects in semantics does not constitute such a

response. I will also develop Kripke's sceptical solution in more detail, and argue that it is not coherent. Thus some other kind of response to the sceptical argument is still required.

In Chapter 3, I turn to the principle objection to a naturalistic theory of meaning: the claim that meaning is normative. I survey a number of possible readings of the claim and argue that meaning is not normative in *any* of the proposed senses. Furthermore, I argue that no compelling sceptical argument from normativity to non-factualism is forthcoming. We will consider, *inter alia*, claims to the effect that 'means' implies 'ought', that meaning is inherently motivating, and that meanings are essentially capable of justifying action. All of these claims turn out to be false.

In Chapter 4, we turn to the most prominent naturalistic/reductive accounts of meaning in the literature. I will argue that they all fail to solve the sceptical problem, even in the absence of the claim that meaning is normative. We will see that there is a pattern to such failures: either such accounts cannot provide determinate extensions for our expressions, or in so doing they presuppose the semantic notions that they are trying to explain.

In Chapters 5 and 6, I attempt to diagnose the failure of naturalising projects in semantics. Specifically, I attempt to show that such projects fail for reasons *other than* the truth of the sceptic's claims. In any attempted reduction, there will be a *reduction base* – those entities, relations, and so on to which some phenomena – the *reduction target* – is intended to be reduced. In Chapter 5, I turn to the reduction base: the 'natural' facts. I will argue that we do not, in fact, have a clear conception of what these are. First and foremost, one would presumably only wish to reduce troublesome phenomena – indeed, one would only *be able to* reduce such phenomena – to the entities, relations, and so on of *true* theories of the natural sciences. However, this raises two pressing problems: first, how do we know which theories are the true theories? And, second, the history of science provides excellent empirical evidence that many, if not all, of our current theories will turn out to be false. For these reasons, among others, projects of naturalising semantics seem at best premature.

In Chapter 6, I turn to the other side of the proposed reduction: the semantic facts. I will raise a number of concerns as to the scientific status of semantics in general. This is important for the following reason: reduction is a process that happens *within* science. One does not reduce the ontology of common-sense to the sciences – indeed I will present reasons to think that one *cannot* do so. Thus if semantics turns out not to be a scientific domain, then calls to naturalise semantics seem to require that the non-scientific reduce to the scientific in order to be ontologically respectable. I will argue that this line of thinking is fundamentally misguided. The upshot of all of this is that the failure of projects of naturalising semantics is consistent with the truth of meaning-factualism.

The Sceptical Argument

2.1 Introduction

In this chapter I will outline the sceptical argument, paying close attention to the role of semantic normativity within it. The core of the sceptical argument is a challenge to cite a fact about oneself that could plausibly be said to *constitute* one's meaning – call such a fact a meaning-constituting fact. As we will see, semantic normativity plays a crucial role in blocking some of the (seemingly) more promising responses. I will also address some questions as to the *form* of the sceptical argument. As Kripke presents it, the sceptical argument proceeds by elimination under certain idealised conditions; specifically, omniscient access is granted to certain domains in our search for a meaning-constituting fact. This is crucially important: the sceptic's claim is that if a suitable fact cannot be found under *those* conditions, then there simply was no fact to be found. Thus the idealisation is of paramount importance for the sceptic to prove his negative result. However, Hattiangadi (2007) has suggested that the argument could be *strengthened* from a by-elimination to an *a priori* argument against any putative account of meaning. This *a priori* argument is possible, she argues, only if the sceptic is granted the assumption that meaning is normative.

2.2 The Sceptical Argument

Kripke's argument centres round one example, though nothing is specific to the example and nothing turns on our changing it. Suppose one is faced with an addition problem, such as '68+57=?' and that one has never performed this particular addition before. Suppose further that one has never dealt with numbers of this magnitude before (as Kripke points out, given the infinitude of the natural numbers, surely such an addition problem exists). Now, as Kripke ([1982] 2003: 8) points out, if I were to answer '125', then my answer could be characterised as the correct one to this addition problem in two importantly different senses. Firstly, it is correct in the 'arithmetical' sense (i.e. 68+57 really does equal 125). Secondly, and more importantly for our present purposes, '125' can be

characterised as correct in the 'meta-linguistic' sense 'that "plus", as I intended to use that word in the past, denoted a function which, when applied to the numbers that I called "68" and "57" yields the value 125' (Ibid). That is, given that I meant *plus* by 'plus' (or '+'), the answer I should *now* return, if I am to accord with my past meaning, is '125'.

Kripke's sceptic begins by questioning one's certainty about the answer '125' in what we have just called the meta-linguistic sense. Perhaps, the sceptic suggests, 'as I used the term "plus" in the past, the answer I intended for "68+57" should have been "5"!' (Ibid). *Ex hypothesi*, I have never performed this particular addition before, or, indeed, any addition of numbers of such magnitude. Therefore, I could not have given myself 'explicit' instructions regarding this particular addition (Ibid). If there is to be a determinate answer about what value I should now give, then this can only be as a result something more general; specifically, my applying 'the very same function or rule that I applied so many times in the past' (Ibid).

But now, the sceptic continues, 'who is to say what function this was' (Ibid)? Perhaps the function that I was applying in the past was not the *addition* function, but the *quaddition* function, where x quus y is given by:

x quus $y = x + y$ if $x, y < 57$; otherwise $= 5$.

Who is to say, the sceptic asks, that this was not the function I previously meant by 'plus'. The sceptic's claim is that in returning '125' to our addition problem, I am 'misinterpreting my own previous usage': by 'plus' I in fact meant (and always meant) *quus* (Ibid: 9).

Surely this wild hypothesis is false. But, Kripke points out, 'if it is false, there must be some fact about my past usage that can be cited to refute it' (Ibid). Thus the sceptical challenge is to cite just such a fact. The challenge takes two forms. First, the sceptic 'questions whether there is any *fact* that I meant plus, not quus, that will answer his sceptical challenge'. Second, the sceptic questions whether I have any reason to be confident about what I should answer *now* to '57+68=?' (Ibid: 11). I wish to make a brief but important point here concerning

the character of the sceptical challenge: while there is an important epistemological question to be answered by any satisfactory account of meaning, the deeper challenge is *not* epistemological, but metaphysical (Ibid: 21). We might best see the epistemological challenge as acting as an important *constraint* on answers to the metaphysical one. Any putative meaning-constituting fact must deliver the right kind of first-person, authoritative epistemology that we take ourselves to have of our meanings. (We will return to this point presently.)

It is important to note that the challenge to account for one's confidence in one's present response is not purely epistemological; it also concerns the *nature* of the putative meaning-constituting fact: such a fact 'must, in some sense, show how I am justified in giving the answer "125" to "68+57=?"' (Ibid: 11). Thus Miller (2007: 168) writes: 'it has to be possible to *read off* from [the] fact what constitutes *correct* and *incorrect* use of the "+" sign'. It is in this sense (i.e. accounting for correct and incorrect use) that a satisfactory response to the sceptic will show 'why the answer to "68+57=?" is [meta-linguistically] justified'. Therefore the challenge concerning justification is, in fact, one of the earliest formulations of the normativity of meaning that one finds in Kripke – something that we will return to in due course.

Before embarking on the quest to cite a suitable fact, Kripke lays down some ground rules for conversing with the sceptic. First, there is to be no dispute concerning what I *now mean* by my words. Of course, if the sceptic is correct, then his challenge generalises: if there was no fact of the matter about what I meant *in the past*, then clearly there is no fact of the matter about what I mean *in the present* either. However, in order to converse with the sceptic, this generalisation of his argument is to be postponed. Second, there are to be 'no limitations, in particular, no *behaviourist* limitations, on the facts that may be cited to answer the sceptic' (Kripke [1982] 2003: 14). This 'no limitations' rule is to be understood very strongly; the sceptic asserts that 'whatever "looking into my mind" may be…even if God were to do it, he still could not determine that I meant addition by "plus"' (Ibid). Thus we are afforded unlimited epistemic access to both our external behaviour (linguistic and non-linguistic), and the entirety of our 'internal mental histories' (Kripke [1982] 2003: 21). The sceptical claim is

thus that if one cannot find a suitable meaning-constituting fact *in those areas and under such conditions*, then there simply is no such fact to be found.[1]

Specific Instructions

With these ground rules in place, let us review the various suggestions for a meaning-constituting fact that Kripke considers. First of all, we have noted that *ex hypothesi*, the addition problem that Kripke utilises is one that the subject of the sceptic's attack – call him Jones, for convenience – has never performed before. So Jones' meaning cannot consist in explicit instructions that he gave himself regarding *this particular* addition problem – by hypothesis he did no such thing. Nor would it suffice for Jones to work out a few example addition problems and give himself something like the following instruction: 'Let "+" be the function instantiated by the following examples…' as no finite number of examples can instantiated any *unique* infinite function (Kripke [1982] 2003: 15).

General Instructions

Perhaps, Kripke suggests at this point, the sceptical problem only arises 'because of a ridiculous model of the instructions' Jones gave himself regarding the 'addition' function. Rather than extrapolating from examples, we might say that Jones learned 'and internalised instructions for' a *rule* 'which determines how addition is to be continued' (Ibid). Well, what rule might that be? Kripke writes:

> To take it in its most primitive form: suppose we wish to add *x* and *y*. Take a huge bunch of marbles. First count out *x* marbles in one heap. Then count out *y* marbles in another. Put the two heaps together and count out the number of marbles in the union just formed. The result is *x+y* ([1982] 2003: 15).

Although rather primitive, the above response may be characterised as a kind of 'algorithm' that Jones learned and internalised (Ibid). Will this response suffice to rebut the sceptic? Sadly not: it will only work if 'count' refers to the act of *counting*. However, Jones has applied 'count', like 'plus', 'only to finitely many cases'. Thus the sceptic can claim that by 'count' Jones formerly meant *quont*,

---

[1] Thus, as we noted in the introduction, Kripke's scepticism is thus more radical than that of Quine (1960) who admits only behavioural evidence in his discussion of the "indeterminacy of translation".

where to 'quont' a heap is 'to count it in the ordinary sense, unless the heap was formed as the union of two heaps, one of which has 57 or more items, in which case one must automatically give the answer "5"' (Ibid). Thus, as Miller (2007: 169) points out, citing a 'general thought or instruction' as a response to the sceptic won't work: the sceptic 'can always respond by giving a deviant interpretation of the symbols of the general thought or instruction itself'.

Kripke ([1982] 2003: 16 fn. 12) rules out a related thought for much the same reason. This is the idea that the '+' sign figures in certain arithmetical laws that are definitive of addition (for example, the law that x plus zero equals x and x plus the successor of y equals the successor of (x plus y)), and that Jones accepts these laws. Alas, the sceptic can merely re-run his challenge at the level of the 'signs used in these laws': each of these can be given its own 'non-standard' interpretation such as to fit, for example, a bizarre quus-like interpretation of '+'.

Dispositions

The next response is perhaps one of the more promising, and it is here that the normativity of meaning comes into its own. Kripke notes that the kinds of mental states just considered – those of entertaining some rule or instruction – are what we might call *occurent*, something that 'comes before the mind' at a given moment, as it were. Perhaps, he suggests, what the failure of such responses shows is that 'the fact that I meant plus (rather than quus) is to be analysed *dispositionally*, rather than in terms of occurent mental states' ([1983] 2003: 22). Thus to mean *plus* by '+' is 'to be disposed, when asked for any sum '*x+y*' to give the sum of x and y as the answer' (Ibid). The dispositional account is promising as one's dispositions do at least *point* to cases beyond the immediate one. In other words, they do speak to what I would do in as yet unconsidered cases.

As promising as it may be, the sceptic has several powerful objections to the dispositional account. The first is the 'finitude' objection. Kripke's thought is that not only my past linguistic performances, but also my dispositions are finite. Thus, for example, my disposition to add does not extend to huge numbers – 'those simply too large for my mind – or my brain – to grasp' ([1982] 2003: 27).

Thus it may be that by '+' I mean *quaddition*, defined thus: x quus $y = x+y$ if x, y are small enough for me to handle. Otherwise, =5.

The second objection we might call the epistemology objection. The thought is this: when I mean something, I tend to know what I mean privately, 'immediately' and with 'fair certainty' (Kripke [1982] 2003: 40). It is hard to see how my dispositions could deliver this kind of knowledge: they are as open to any observer as they are to me, and in order to make a determination of my meaning on the basis of them would require me to investigate them over time, perhaps entertaining 'hypotheses' concerning what I mean, or drawing inferences to my meaning on the basis of my dispositions. None of this seems right as an analysis of the epistemology of meaning (cf. Kripke [1982] 2003: 21-2).

The third and perhaps most powerful objection is the normativity objection. The core of this objection, as Boghossian (1989a: 513) observes, is that the dispositional account gives the wrong analysis of the relation between meaning and use. Kripke writes:

> Suppose I do mean addition by '+'. What is the relation of this supposition to the question how I will respond to the problem '68+57=?' The dispositionalist gives a *descriptive* account of this relation: if '+' meant addition, then I will answer '125'. But this is not the proper account of this relation, which is *normative* not descriptive. The point is *not* that, if I meant addition by '+', I *will* answer '125', but rather that, if I intend to accord with my past meaning of '+', I *should* answer '125'. Computational error, finiteness of my capacity, and other disturbing factors may lead me not to be *disposed* to respond as I *should*, but if so, I have not acted in accordance with my intentions. The relation of meaning and intention to future action is *normative*, not *descriptive* (Kripke [1982] 2003: 37).

A further point is that in misrepresenting the relation between meaning and use, the dispositional view collapses the distinction between correct and incorrect, acting correctly and making a mistake, between 'seems right' and 'is right' (Ibid: 23-4). In other words, if being an answer I am disposed to give is what *makes it the case* that such an answer is correct, then we can make no sense of making a linguistic mistake: everything I am disposed to do is 'correct'.

Perhaps, however, the dispositional account under consideration is too crude. It takes it for granted both that I am disposed to do whatever I, in fact, do and that *all* of my dispositions are meaning-constituting. What if we modify the

account to make it more sophisticated? Due to suggestions such as these, Kripke adds a *ceteris paribus* clause to the dispositional account. According to sophisticated dispositionalism, then, to mean 'plus' by '+' is to be disposed, *ceteris paribus*, to return the sum of a pair of numbers.

Alas, the sceptic has some powerful responses to this suggestion. These concern certain difficulties that the dispositionalist faces when she tries to fill out the *ceteris paribus* clause. The first difficulty is circularity. We can imagine many different possible idealisations of one's dispositions – which is the 'right' one? As Kripke ([1982] 2003: 28) points out, one cannot choose one idealisation over another by assuming which function is intended or meant. This would both render the proposal circular and violate the terms of discussion with the sceptic. A second problem concerns *systematic* mistakes (Ibid: 30-2). Suppose Jones systematically forgets to carry when answering addition problems even in what we might call ideal conditions – he is not tired or distracted, the numbers are not too large for him to grasp, and so on. So, if these optimal or ideal conditions give the content of the *ceteris paribus* clause, even *ceteris paribus* Jones is not disposed to return sums to addition problems. But now, one what basis do we say that Jones is adding incorrectly, rather than, say, *quadding* correctly? To assume one answer over the other is to render the proposal circular again. This concludes my brief discussion of the dispositional account. I will take the issue up again in chapter 4.

Simplicity

Kripke now turns to another suggestion: Jones means *plus*, not *quus*, by '+' as that is the simplest hypothesis on offer as to what Jones meant. There are several problems with this suggestion. Kripke passes over some of these quite quickly, as he detects a deeper problem, but I will mention them here anyway for completeness. First, simplicity may be 'relative'. Second, it is hard to give a precise definition of what simplicity is. Finally, what is simple for us may not be so for a Martian, who might find the quus function simpler than the plus function (Kripke [1982] 2003: 38). However, the deeper problem is this: the simplicity response, Kripke says, misunderstands the nature of the sceptical problem. To repeat, the problem is not epistemic – it is not, given these two hypotheses about

what Jones means by '+', how do we know which is correct. Simplicity considerations *might* help us to answer that question. However, the real sceptical problem is not to help us choose between hypotheses, but to 'tell us what the hypotheses *are*' (Ibid). This response will be important to bear in mind when we consider Kripke's own 'sceptical' solution to the paradox.

Qualia

Kripke next turns to the idea that some qualitative mental state might do the trick. There are two separate suggestions here. The first is that meaning might consist in some familiar kind of qualitative state, such as an image or a feeling. The second is that meaning might consist in a unique kind of qualitative state. As Miller (2007: 170) points out, the sceptic has damaging arguments to the effect that such a state of the former type is neither necessary nor sufficient for meaning something by an expression. Consider the necessity first. Suppose that Jones means *green* by 'green'. It is quite conceivable that nothing qualitative comes before his mind when he means or understands the expression 'green'. Even supposing that there is some 'empirical regularity' such that some greenish image comes before his mind when he understands 'green', we could still '*conceive* of someone understanding the expression in the absence of that item' (Miller 2007: 170).

Turning now to sufficiency, the basic problem with the idea that mental images, 'pictures' and the like could serve as the meanings of expressions is that *they must themselves be interpreted* in order to provide a correctness-condition for the expressions. As Miller puts it:

> The essential point is that the picture does not by itself determine the correct use of the associated word, because the picture thus associated is really just another sign whose meaning also requires to be fixed (2007: 170).

To take a variant of one of Kripke's examples, suppose something greenish comes before Jones' mind when he understands 'green'. What does the greenish image

signify? Perhaps it signifies the property of being *grue*, where an object is grue iff it is green before the first of July, 2011, and blue thereafter (cf. Goodman 1973). To take another example, Kripke ([1982] 2003: 41-2) writes:

> Suppose I do in fact feel a certain headache with a very special quality whenever I think of the '+' sign. How on earth would this headache help me figure out whether I ought to answer '125' or '5' when asked about '68+57'?

In brief, any quale would have to be interpreted in order to yield a determinate meaning for an expression. But at precisely this point of interpretation, the sceptical challenge may be run again.

What about unique qualitative states? Sadly, the sceptic has some powerful objections here as well. First, the postulation of such a state is rubbished by the sceptic on the grounds that such a move is 'desperate' and that it leaves the nature of the putative qualitative state 'completely mysterious'. As Kripke ([1982] 2003: 51) points out, such a state is not open to introspection (if it were, it would have 'stared one in the face' and 'robbed the sceptic's challenge of any appeal), and yet each of us is supposedly aware of it to some fair degree of certainty. How could this be? Furthermore, how could such a finite state, held within a finite mind, 'reach out' to a possible infinitude of future uses of an expression so as to determine which uses of that expression were correct (Miller 2007: 171)? There is certainly more than a hint of paradox here and, 'as it stands', the current account 'gives no clue as to what the nature of this relationship is, or how it is forged' (Ibid). For these reasons, the sceptic concludes that the 'qualia' response, taken in either form, provides no firm foundation for the meanings of expressions.

Sui generis meaning-facts

Kripke ([1982] 2003: 51) next suggests that 'we may try to recoup' by arguing that states of meaning are *sui generis* – primitive states 'not to be assimilated to sensation or headaches or any "qualitative" states, nor to be assimilated to dispositions, but a state of a unique kind of [their] own'. Kripke concedes that such a move 'may in a sense be irrefutable' but argues that:

> It seems desperate: it leaves the nature of this postulated primitive state…completely mysterious. It is not supposed to be an introspectable state, yet

we are supposedly aware of it with some fair degree of certainty whenever it occurs…Such a state would have to be a finite object…which could *not* be interpreted in a quus-like-way[.] How could that be? ([1982] 2003: 51-2).

The postulated state indeed has an air of paradox about it – finitely contained, yet speaking to an infinitude of possible uses in such a way as to not be open to sceptical interpretations; not introspectable, yet something we are aware of 'immediately' and 'with fair certainty'. How could this be? It seems that anyone wishing to defend the *sui generis* view owes the sceptic an account of the answers to these seemingly intractable questions.

Platonism

Finally, Kripke considers a kind of Platonism about language, such that the addition function, taken as, for example, a set of triples contains within itself all of its own instances (such as the triple (57, 68, 125)). On such a Fregean/Platonic view the 'sense' of an expression would determine its 'referent': it is simply the nature of a sense to do so. We might then see the sense of the expression 'plus' as determining the addition function as its referent. While this is all very well, it is not of any great help in dealing with the sceptic. Indeed, the sceptical problem could simply be recast in Platonic/Fregean terms: what *constitutes* my grasping a sense? As Kripke points out, the Platonic response is of no real help here: it is simply an 'unhelpful evasion' of the real problem of explaining 'how our finite minds can give rules that are supposed to apply to an infinity of cases' ([1982] 2003: 54). While the addition function, construed Platonically, can contain within itself its entire domain and range, my mind clearly cannot do so. How then can I 'grasp' the meaning of '+'?

The Sceptical Paradox

As we have seen, none of the accounts the sceptic considers is able to meet his challenge. Thus we are left with the shocking conclusion that there are no facts of the matter concerning what Jones meant in the past. Of course, this now generalises to the present as well: if there were no facts constituting Jones' *past* meaning, then there are no facts constituting his *present* meaning either. And, of course, this doesn't just go for Jones, it goes for everybody: no-one ever means

anything by any expression! This leaves us with the following 'sceptical paradox': 'no-one ever means anything by any expression', if true, is meaningless, and if it is meaningless it cannot be true (Kusch 2006: 23). I will outline Kripke's sceptical solution to the paradox in a moment. I will then argue that it does not seem to be coherent. First, however, I wish to offer some brief comments on the sceptical argument that will be important for what is to follow in the rest of the thesis.

Some Comments on the Roles of Normativity and Reductionism in the Sceptical Argument

First of all, note that as it is usually read – indeed, as I have presented it here – the sceptical argument proceeds by elimination under epistemic idealisations. One may thus offer the nice point that, strictly speaking, the sceptic is not entitled to claim that there is *no* possible meaning-constituting fact – perhaps we have just not been clever enough to think of one. After all, it is one thing *imagining* oneself to have unbounded epistemic access to the potentially relevant facts, and quite another thing to *actually* have it.

However, Hattiangadi (2007) has argued that the sceptical argument can be strengthened to an argument with full *a priori* status against *any* putative theory of meaning. This result is only secured, she argues, *if* the sceptic is granted the claim that meaning is normative. In the next chapter, we will consider in some detail how such an argument may go. However, I will provide a sketch here of how Hattiangadi sees the argument running. Essentially, Hattiangadi argues that the semantic realist's position is analogous to that of the moral realist, and thus Ayer's argument against moral realism may apply, *mutatis mutandis*, to the semantic case (2007: 39). Ayer (1936) argues that the moral realist faces a dilemma: either 'the putative moral facts are said to be "natural" facts' in which case the realist commits Hume's 'naturalistic fallacy', or the moral facts are non-natural, hence 'unlike ordinary empirical facts and unknowable' (Hattiangadi 2007: 39). In the next chapter, we will discuss how arguments from Humean views about motivation and normative facts, and what we might call Mackiean views about 'queerness' could be developed. I will argue that these arguments are unsuccessful as meaning is not normative in the sense required for such arguments to run.

The second feature of the sceptical argument that I wish to draw attention to is the implicit reductionist assumption made by Kripke that meaning and intentionality must be accounted for in terms of some *other* kind of fact. Colin McGinn (1984) has taken issue with this. He writes:

> The sceptic is assuming that unless semantic facts can be captured in non-semantic terms they are not really facts; but why should this assumption be thought compulsory? So the question for Kripke is why we cannot give the truth conditions of 'he means addition by "+"' simply by *re-using* that sentence, frankly admitting that no other specification of truth conditions is available – precisely because semantic statements cannot be *reduced* to non-semantic ones (1984: 151).

Wright ([1989] 2002: 112-3) has some sympathy with this, but takes McGinn to task for not recognising the need for the non-reductionist to provide an account of those paradoxical aspects of meaning-states mentioned above: our non-inferential authoritative knowledge of them that is obtained even though they do not appear in qualitative consciousness, and their ability to speak to a potential infinity of as yet unconsidered situations despite having to fit within a finite brain.[2]

I agree with McGinn that we do not have to reduce semantics to some other domain in order for there to be facts of the matter about meaning. However, my aim in this thesis is not to provide a non-reductionist account of meaning. Rather, I seek to make a plausible case that we should, in fact, *expect* reductive accounts of meaning to fail – but not because there simply are no facts about meaning to capture. Rather, reductionism about meaning fails because of the *nature* of the semantic facts, of reduction, and of the resources available to the reductionist.

It is also worth noting here that the desire to reduce the semantic/intentional often goes hand in hand with a commitment to metaphysical *naturalism*; roughly, the view that only those entities that appear in theories of the natural sciences are 'real', and therefore that the only facts are 'natural' (usually, ultimately physical) facts. For instance Fodor (1987: 97) writes:

---

[2] Such an account, of course, is exactly what Wright (1989) attempts to provide by taking intentions as basic and attempting to account for their contents.

> I suppose that sooner or later the physicists will complete the catalogue they've been compiling of the ultimate and irreducible properties of things. When they do, the likes of *spin*, *charm*, and *charge* will perhaps appear on their list. But *aboutness* surely won't; intentionality simply doesn't go that deep…If the semantic and intentional are real properties of things, it must be in virtue of their identity with (or maybe their supervenience on?) properties that are themselves neither intentional *nor* semantic. If aboutness is real, it must be really something else.

Indeed, it is hard to see why one would *want* to be a reductionist about meaning and intentionality, *unless* one had some such metaphysical commitments concerning which entities are ultimately 'real'. We will return to questions of naturalising semantics in chapter 4. For now, let us return to Kripke and his sceptical solution to his paradox.

2.3 The Sceptical Solution

The sceptical argument concludes that there are no facts of the matter in virtue of which we mean something in particular by our words, and thus that there are no facts of the matter about what we mean. How, then, are we to get out of this morass? Kripke ([1982] 2003: 66) distinguishes two ways that the sceptical argument may be responded to. A *straight* solution 'shows that on closer examination the scepticism proves to be unwarranted'. By contrast, a *sceptical* solution concedes 'that the sceptic's negative assertions are unanswerable' but holds that, nevertheless, 'our ordinary practice or belief is justified because…it need not require the justification that the sceptic has shown to be untenable'. As Miller points out, in the case of meaning, this would amount to an argument to the effect that our 'practice involving sentences which ascribe meaning and understanding' does not require the 'justification which the sceptic demands', that is, 'an account of the facts that would render [such sentences] true or false' (Miller, 2007: 175).

Instead of focusing on the truth-conditions of meaning ascriptions, Kripke turns his attention to two questions. First, 'under what conditions can [ascriptions of meaning] be asserted (or denied)?' Second, 'what is the role, and the utility, in our lives of asserting (or denying) [such ascriptions] under these conditions?' ([1982] 2003: 73). The answer to the first question is, roughly, when a speaker, *considered as a member of a linguistic community*, has performed 'satisfactorily often enough' with an expression (Miller, 2007: 177). As Miller points out, this is

taken to involve nothing more than such a speaker using the expression 'satisfactorily often enough' in the same way(s) that 'his fellow speakers in the community' are disposed to use it (Ibid). The utility or role of meaning ascriptions in such conditions is thus to mark the acceptance of the speaker into their midst, and to mark the fact that that speaker is trustworthy with respect to his or her use of the expression (Ibid). The utility of ascriptions of meaning to speakers is thus to distinguish those who can be trusted in their use of expressions from those who cannot (Ibid). Thus meaning ascriptions play a useful role in our lives, and thus our practice of ascribing meaning is justified.

A corollary of this view is that we can make no sense of ascriptions of meaning to speakers *considered in isolation*. The conditions under which meaning would be ascribed to such a speaker are those where *that speaker himself or herself* would assert, say, 'I mean addition by "+"'. As Miller points out, this leaves no room for a distinction in the assertion conditions between, for example 'Jones *believes* Jones means *addition* by "+"' and 'Jones means *addition* by "+"' (2007: 176). This collapsing of the distinction between 'seems right' and 'is right' 'only means that here we can't talk about "right"' (Wittgenstein *PI*: §202). Matters stand differently when Jones is considered as a member of a linguistic community: the assertability conditions for 'Jones means *addition* by "+"' here make essential reference to community linguistic inclinations. Thus we can make sense of a correctness-condition for Jones' use of the plus sign: Jones' uses are 'right' iff they accord with the linguistic inclinations of his fellow speakers.

What are we to make of this sceptical solution? Is it coherent? Zalabardo (1989) offers some compelling reasons to think not. As Miller (2007: 184-5) points out, Zalabardo's first objection to the sceptical solution is remarkably similar to the objection Kripke himself raises regarding simplicity considerations. Recall that Kripke dismissed the simplicity response as betraying a misunderstanding of the sceptical challenge. The challenge is not really to make a choice between two competing underdetermined *hypotheses* about what a speaker means. If that were the challenge, then simplicity considerations could indeed come into play. Thus we might make the determination that Jones means *plus*, and not *quus*, by '+' on the basis of the fact that the addition function is the simpler of the two.

Now, consider the sceptical solution in this light. In other words, let us suppose, for the moment, that the sceptical problem is to resolve an *indeterminacy* regarding what Jones means by an expression. Zalabardo (1989: 37) puts the matter this way: Suppose it is indeterminate which meaning Jones attaches to the '+' sign. Now, Jones has certain inclinations to use the '+' sign, but these cannot be characterised as 'correct' or 'incorrect' since, because the meaning Jones attaches to '+' is indeterminate, there is no fact of the matter concerning which of Jones inclinations accord with his meaning and which do not. Suppose now that we wish to know when we are justified in characterising those inclinations as 'right' or 'wrong'. Here the sceptical solution is quite pertinent: we can take Jones' inclinations to use '+' *in the same way as the other members of his community* as correct, and those inclinations to use '+' in other ways as incorrect.

So, problem solved? Well, no, sadly not. As Zalabardo (1989: 36) rightly points out, the result of the sceptical argument 'does not concern an indeterminacy as to which rule is being followed, but the notion of rule itself'. Thus the sceptical solution should not be viewed as a way to resolve an 'indeterminacy' concerning *which* rule a speaker is following. That is, it should not be intended to pick between two competing hypotheses concerning what Jones means by '+': there are no such hypotheses on the table, as it were. Rather, the solution should be aimed at rehabilitating the whole idea of following a rule or meaning something by an expression.

How does the sceptical solution fare when viewed in this light? Recall that Kripke's account of the meaning of meaning-ascriptions makes crucial reference to the conditions under which such an ascription may be asserted. Thus the question is whether 'Kripke's description of the assertability conditions of content' could provide the meanings of ascriptions of content or meaning. There are two principal problems here.

The first problem, as Zalabardo points out, is that the notion of content that emerges on such a view is 'hardly recognisable' with all judgements collapsing into judgments about the linguistic inclinations of one's community (Zalabardo 1989: 39). I will explain the reasoning behind this conclusion now.

Zalabardo begins with the observation that whatever constitutes what a speaker means by an expression must contain a determination of that state of affairs that fits the content of that expression (1989: 38). Thus, for example, whatever constitutes what Jones means by 'grass is green' must determine the state of affairs that fits the content of this expression (*viz.* grass being green). Elsewhere, Zalabardo puts the thought slightly differently: it is the 'determination of conditions of fit that constitute content' (1989: 39). Suppose we take the expression 'green'. We can give the meaning or content of this expression by stating its 'conditions of fit'; that is, by stating its correctness-conditions. Thus:

> Jones applies 'green' to *x* correctly iff *x* is green.

If there were content-constituting facts, it is precisely such facts that would determine the conditions of fit, and thereby constitute content. However, according to the solution of the sceptical argument, *there are no content-constituting facts*. Given this result, we are entitled to wonder: if content is constituted by the determination of conditions of fit, then, in the absence of content-constituting facts, what determines conditions of fit? And what kind of notion of content emerges on such a view?

So Zalabardo's leading question is this: once the sceptical conclusion goes through, what determines the conditions of fit that constitute content? Now, as we have seen, the sceptical solution offers us *assertability conditions* for ascriptions of content: a content ascriber – call him Smith – can assert that Jones means, say, *green* by 'green' iff Jones' performance with 'green' is satisfactorily similar to Smith's use of 'green' satisfactorily often. In other words, all Smith can do is to 'provide his own blind inclination to act as a pattern of correctness' for Jones' performance with the expression 'green' (39).

Now, Zalabardo argues, suppose that such assertability conditions for the ascription of meaning or content to Jones' expression 'green' could give us the *meaning* of such an ascription such that:

> 'Jones means *green* by "green"' means that Jones' inclinations to use 'green' fit the community's inclinations to use it.

Alternatively:

> 'Jones' expression "green" has the meaning/content *green*' means that Jones' inclinations to use 'green' fit the community's inclinations to use it.

This gives us a new kind of correctness-condition for 'green':

> Jones applies 'green' to *x* correctly iff Jones' linguistic community are inclined to apply 'green' to *x*.

From this we can give a correctness-condition for any *assertion* that Jones may make regarding the greenness of an object:

> Jones' assertion that *x* is green is correct iff Jones' linguistic community are inclined to assert that *x* is green.

Now, here is the problem: if, as it is natural to suppose, the identity conditions for contents are given via specifications of the correctness-conditions of those contents, then, as Zalabardo points out, it seems that all of Jones' judgments will ultimately be judgements about the linguistic inclinations of his community (1989: 38-9). Here is why. When we make a judgement – say, the judgement that some object is green – we in effect judge that the correctness-condition for the content of the judgement obtains. On the classical picture that the sceptical argument undermines, such a correctness-condition is just the *truth*-condition of that content. Thus:

> Jones' judgement that *x* is green is correct iff *x* is green.

So:

> Jones' judgement that *x* is green has the content that x *is green*.

However, on our current sceptical account, we have it rather that:

Jones judgement that *x* is green is correct iff Jones' linguistic community are inclined to judge that *x* is green.

So:

Jones judgement that *x* is green has the content that *Jones' linguistic community are inclined to judge that x is green*.

Now, of course, there are some glaring problems with this. First, there is no scope for so much as making a judgment about the world on such a view (Ibid: 39). Second, making a judgement about the linguistic inclinations of my community simply does not seem to be what I do when I judge, say, that *grass is green*. Third, such a view fails to make sense of some fairly robust intuitions about linguistic correctness. For instance, Blackburn writes:

> If my community all suddenly started saying that 57+68=5, this fact does not make me wrong when I continue to assert that it is 125. I am correct today in saying that the sun is shining and daffodils are yellow, regardless of what the rest of the world says (1984: 294).

Thus, Zalabardo concludes, the notion of content that emerges on Kripke's sceptical solution is barely recognisable (1989: 39).

The second problem with viewing Kripke's assertability conditions as giving the meanings of meaning-ascriptions is that, viewed in this light, the sceptical solution in fact, looks like a *straight* solution (Zalabardo 1989: 39-40; Miller 2007: 186). This would be of roughly the form:

Jones means *plus* by '+' iff the assertability conditions for 'Jones means *plus* by "+"' are met.

In other words,

Jones means *plus* by '+' iff Jones performs satisfactorily often enough with '+' in a similar ways to those employed by the other members of his linguistic community.

The main difficulty here, as Miller (2007: 186) points out, is that such an account of the meanings of meaning ascriptions *itself* falls prey to the sceptical argument. Consider:

> Smith associates with 'Jones means *plus* by "+" the assertability condition that *Jones has performed satisfactorily often enough with '+' in ways that are similar to other members of the linguistic community.*

Crispin Wright (1984: 770) has clearly spelled out how the sceptical argument can be run against such an account:

> Could it yesterday be true of a single individual that he associated with the sentence 'Jones means addition by "+"' the sort of assertion conditions Kripke sketches? Well, if so, that truth did not consist in any aspect of his finite use of that sentence, or its constituents; and, just as before, it would seem that his previous thoughts about that sentence and its use will suffice to constrain within uniqueness they proper interpretation of the assertion conditions he associated with it only if he is granted correct recall of the content of those thoughts – exactly what the sceptical argument does not grant. But would not any truth concerning assertion conditions previously associated by somebody with a particular sentence have to be constituted by aspects of his erstwhile behaviour and mental life? So the case seems to be no weaker than in the sceptical argument proper for the conclusion that there *are* no such truths; whence, following the same routine, it speedily follows that there are no truths about the assertion conditions that any of us presently associates with a particular sentence, nor, *a fortiori*, any truths about a communal association.

Thus not only does Kripke's account of the meaning of meaning-ascriptions yield a bizarre notion of content, it also falls prey to his own sceptical argument!

Of course, it may be that Kripke's assertability-conditions for ascriptions of content are *not* supposed to provide the meanings of such ascriptions. However, if this is the case, then, as Zalabardo (1989: 39) points out, we are inclined to wonder what in the sceptical solution *does* give us the meanings of content ascriptions. It is hard not to agree with Zalabardo's contention that 'Kripke gives us no help in this respect' (Ibid). But if this is the case, then:

> We are forced to admit that [while Kripke] describes the conditions under which [content] is attributed…[he] provides no indication of what the content of the attribution might be. But this amounts to saying that the main problem posed by the sceptic – that the concepts of meaning and intending one function rather than another make no sense – remains unsolved (Zalabardo 1989: 39).

Furthermore, as Miller (2007: 187) points out, if the purpose of Kripke's appeal to assertability-conditions for content is *not* to give the meanings of such ascriptions, then it is utterly unclear what role those assertability-conditions are playing in the sceptical solution.

In summary, we may say that Zalabardo presents Kripke with a kind of dilemma: either the assertability-conditions for ascriptions of content give the meanings of those ascriptions, or they do not. If they do, then, firstly, the notion of content that arises is utterly unrecognisable and leads to some highly undesirable consequences and, secondly, the sceptical solution itself seems vulnerable to the sceptical argument. If they do not, then it seems that the sceptical problem is left untouched by the sceptical solution. In any event, it seems that the sceptical solution fails to satisfactorily solve the problem it was intended to solve.

Zalabardo's argument seems pretty damaging to the sceptical solution. However, perhaps we have been too hasty. In recent work, several commentators on Kripke[3] have questioned exactly what claims about meaning are advanced by *the proponent of a sceptical solution*, as opposed to by *the sceptic himself*. That is, as we noted above, a sceptical solution to Kripke's paradox would begin by accepting that the sceptic's negative assertions about meaning are unanswerable. However, as Kripke's sceptic in fact makes more than one negative assertion about meaning, we immediately face the problem of determining which of the *sceptic's* negative assertions are accepted by *the proponent of the sceptical solution*. In what follows I will focus on Wilson's ([1994] 2002) discussion of this problem. I will argue that, while Wilson makes an important distinction between the sceptic's two main negative claims, his attempt at a 'factualist' sceptical solution to Kripke's paradox is unsuccessful.

Wilson reads the proponent of the sceptical solution as undermining the 'classical realist' (CR) picture of meaning. According to CR, if a speaker means something by an expression, then there is some property (or set of properties) that governs the correct use of that expression for that speaker. Thus, for example, if Jones means *dog* by 'dog', then there is some property (or set of properties) that governs the correct use of 'dog' for Jones (in this case the property of doghood).

---

[3] See, for instance, Wilson [1994] 2002; Davies 1998 Kusch 2006.

The sceptical argument charges that if CR is true, then there must be some fact about Jones that constitutes the fact that it is the property of doghood that governs the correct application of his expression 'dog' and not some other property, say, dog-or-cathood. Wilson calls this the 'grounding constraint': 'the existence of conditions of applicability for a term must be intelligibly grounded in facts about the speaker's psychological and/or social history' ([1994] 2002: 239). We can express this claim as the following conditional: if CR is true, then there must be a fact about a speaker S that constitutes some set of properties as conditions that govern S's use of an expression E. Or, more briefly, if CR is true, then classical semantic factualism (CSF) is true.

Wilson now distinguishes two negative assertions about meaning that are made in the course of *WRPL*: the rejection of classical semantic factualism and the assertion of *semantic nihilism*:

(~CSF)   There are no facts about a speaker S that constitute any set
      of
          properties as conditions that govern S's use of an
      expression E.

(SN)   No one ever means anything by any expression.

Wilson contends that, while the *sceptic* asserts *both* ~CSF and SN, the proponent of the sceptical solution only accepts ~CSF. Thus the proponent of the sceptical solution jettisons CR on the back of a rejection of SN, and places an account in terms of communal dispositions and assertability conditions for meaning ascriptions in its place. We may thus distinguish the positions of the sceptic and the proponent of the sceptical solution as follows:

Sceptic: CR → SF, ~SF (SNF) ⊢ SN

Proponent: CR → SF, ~SF, ~SN ⊢ ~CR

This raises the intriguing possibility that rejecting a classical realist account of semantic factualism does not necessarily commit one to rejecting semantic

factualism *tout court*. Perhaps there are some *other* kinds of semantic facts. As Miller puts it, perhaps ascriptions of meaning do report facts, and therefore have truth-conditions, just not 'the kind of truth-conditions that the sceptical argument attacks' (2007: 191). Regarding Zalabardo's argument, we may wonder (i) whether Zalabardo requires the proponent of the sceptical solution to endorse SN in order for his argument to go through; and (ii) whether there is a tenable form of meaning-factualism that escapes his argument.

A kind of sceptical meaning-factualism is advanced by Wilson himself. His account centres around what he calls 'normative conditionals' of roughly the form:

(NC)   Jones means *green* by 'green' → Jones ought to apply 'green' to an object only if that object is green.

According to classical realism, such conditionals are made true by facts about Jones: it is CSF that makes NC true. Now, of course, the proponent of the sceptical solution has *rejected* CSF. So they need a different account of conditionals such as NC. Wilson replaces the notion of a truth-condition for NC with an *assertability condition* for it: roughly, the idea is that if Jones applies 'green' to green objects satisfactorily often enough, we may assert that Jones means *green* by 'green'. Conversely, if Jones applies 'green' to non-green objects persistently and often enough, then we lose our right to assert that Jones means *green* by 'green'. Thus the sceptical solution is really a form of factualism about meaning: 'Jones means *green* by "green"' has the truth condition that the assertability conditions for 'Jones means *green* by "green"' are met. Note that this condition for ascribing meaning to Jones' use of 'green' provides us with a correctness-condition for Jones' use of that expression: Jones applies 'green' correctly iff he uses it in ways that do not lead us to withdraw the ascription of meaning to him. Wilson puts the thought in this way: our standard criteria for asserting that Jones means something by an expression enforce and engender requirements that govern the correct application of that expression for Jones.

As promising as this line of thought may seem, Miller has offered some powerful reasons for thinking both that the sceptical solution cannot be a form of

factualism about meaning and that Wilson's account of the sceptical solution does not resist the slide from the rejection of classical realism to semantic nihilism. First, recall that, on Wilson's construal of the sceptical solution, our standard criteria for asserting that Jones means something by an expression enforce and engender requirements that govern the correct application of that expression for Jones. Now, Miller argues, we can apply an analogous grounding constraint that that placed upon the classical realist to Wilson's solution. That is, if such requirements govern Jones' applications of an expression, then there must be a fact about Jones that *constitutes* the fact that it is *those* requirements that do so. Now here is the problem: the sceptic will argue that *there is no fact about Jones (or his speech community) that could constitute the fact that it is just* those *requirements being enforced and engendered by our standard criteria for ascriptions of meaning, and not some other requirements*. To give a specific example, nothing constitutes the putative fact that it is the requirement that 'dog' ought to be applied to x only if x is a dog that governs Jones' use of 'dog' and not, say, the requirement that 'dog' ought to be applied to x iff x is a cat or a dog (2007: 197). In other words, Wilson's sceptical solution itself falls prey to the sceptical argument. From this it follows that there are no 'sceptical' meaning-facts of the envisaged sort. So Wilson's sceptical solution does not prevent the sceptic from drawing the inference from the denial of 'sceptical' semantic realism to semantic nihilism.

Note that Wilson's account seems to find itself impaled on one of the horns of Zalabardo's dilemma: his account of the assertability conditions for ascriptions of meaning gives the *truth-conditions* for those ascriptions. As Miller (Ibid: 198) points out, the sceptical solution thus loses its distinctive sceptical character, becoming instead another attempt at a straight solution, and, as we saw above, it is thus itself susceptible to the sceptical argument. At this point it is hard to see why at least most of Zalabardo's argumentation against the sceptical solution fails to go through. Even if we leave aside his argument concerning the nature of content that emerges on the sceptical solution, his arguments concerning the loss of the sceptical nature of the sceptical solution and the possibility of running the sceptical argument against the sceptical solution seem to rule out the possibility of a successful *factualist* sceptical solution. If instead the proponent of

the sceptical solution eschews factualism about meaning *of any kind*,[4] then it is at best difficult to see how we are to resist the slide from non-factualism about meaning to semantic nihilism, in which case all of Zalabardo's arguments seem to go through.

2.4 Conclusion

Kripke's sceptical argument is powerful and startling. However, his sceptical solution is unsatisfying, so the search for a different line of response to the sceptical paradox is back on. As we saw in this chapter, the claim that meaning is normative is utilised by Kripke as a block to the promising dispositional response to his meaning scepticism. Perhaps, then, a straight response could be rehabilitated were the claim that meaning is normative to be shown to be false. In chapter 3, I will outline several ways that the claim that meaning is normative may be understood. I will argue that meaning is not normative in any of these senses, except perhaps in the weakest sense – the sense that meaningful expressions have conditions of correct application. Furthermore, I will argue that there is no plausible argument from the normativity of meaning to semantic non-factualism. In chapter 4, I will examine the most prominent reductive/naturalistic accounts of meaning and argue that, even relaxing the claim that meaning is normative, none of them is successful.

---

[4] Miller (2007: 200-1) argues that the proponent of the sceptical solution is best read as rejecting *both* classical realism and *any* form of factualism about meaning.

Semantic Normativity

3.1. Introduction

In chapter 2 we saw that the sceptic utilises the claim that meaning is normative as a block to the seemingly promising dispositional account of meaning. As Zalabardo ([1997] 2002) rightly notes, such a move potentially threatens several 'naturalistic' accounts of meaning as well, particularly those that treat meaning in information-theoretic terms, as these may be seen as broadly dispositional (roughly, they work in terms of what *would* cause a mental representation to be tokened in ideal circumstances, and are thus relevantly similar to the sophisticated dispositionalist view discussed by Kripke). Furthermore, we noted Hattiangadi's claim that the whole sceptical argument may be strengthened to become an *a priori* argument against *any* putative reductive/naturalistic account of meaning *if* the sceptic is granted the assumption that meaning is normative. (This contrasts with the more standard 'by-elimination' reading of the argument.) In this chapter, we turn to examine the claim that meaning is normative, and how it is intended to block 'naturalistic' theories of meaning, in much more detail.

One difficulty we face is that there is a lively and ongoing debate about just what the normativity of meaning amounts to. If we are to assess the success of Kripke's semantic scepticism, we must therefore attend to all of the prominent suggestions on offer. There are several questions to which we must then address ourselves regarding each conception of semantic normativity. These are as follows:

1. On such a conception of semantic normativity, what does the claim that meaning is normative amount to?
2. Does this conception of semantic normativity provide the sceptic with an argument from normativity to non-factualism about meaning? If so, what is the most charitable reading of such an argument?
3. Is meaning normative in this sense?
4. Does the argument from normativity to non-factualism go through?

Obviously, our answer to question 3 may well inform our answer to question 4. That is, if it can be shown that meaning is not normative in the intended sense, then any argument from such normativity to non-factualism collapses. Conversely, if the arguments from the intended sense of normativity to non-factualism entail bizarre or implausible consequences, then this would constitute a reason for holding that meaning cannot be normative in the intended sense. So the claim that meaning is so normative may be attacked indirectly, as it were, by examining the consequences of the arguments to non-factualism. Of course, if the argument from normativity in the intended sense to non-factualism fails, then it becomes less important, in a sense, to show that meaning is not so normative. As far as we are concerned in this chapter, normativity is only problematic insofar as it allows the sceptic to mount arguments against factualism about meaning.

It is my contention that meaning is not normative in any sense that has been suggested so far in the literature. Accordingly, I will argue that none of the corresponding arguments from normativity to non-factualism is successful. In what follows, I shall set out the various accounts of normativity that are on offer, and attempt to show that none leads to a plausible argument from normativity to non-factualism. This conclusion lends weight to Hattiangadi's (2007) claim that the *a priori* reading of the sceptical argument trades on an equivocation between 'normativity' and 'norm-relativity', where the former term implies *prescriptive and action guiding*, whereas the latter implies merely *relative to a standard*. Hattiangadi concludes that meaning is not normative in the first sense, but it is a platitude that it is 'norm-relative'; that is, it is platitudinous to most philosophers that meaningful expressions possess correctness-conditions. Her claim is then that, while extant responses to the sceptic cannot account for correctness-conditions (a view that I will argue in favour of in chapter 4), this does not constitute an *a priori* argument against meaning-factualism (as strong normativity does). In what follows I will expand and defend this general thesis.

In this chapter, we will look at the four principle ways in which the claim that meaning is normative has been cashed out in the literature. These may be broadly categorised by using the following short taxonomy:
1. Normativity and Trans-Temporality;
2. Normativity and Correctness-Conditions;

3. Normativity and Motivation;
4. Normativity and Justification.

 (As we will see this taxonomy is a little crude as a number of subtly different views get subsumed under (2), and certain views of normativity consist of a combination of components which fall under different headings, for example Boghossian's view of normativity combines (2) with (3).)

In §3.2, I will outline Colin McGinn's trans-temporal account of the normativity of meaning. We shall see that, while this reading is certainly suggested by Kripke's text, it cannot be the conception that Kripke had in mind as it poses no serious obstacle to the dispositional account of meaning to which Kripke raises normativity as an objection. In §3.3, we will consider the claim that any statement of an expressions meaning gives a correctness-condition for that expression's application, and that 'correct' here is a normative notion. Much of the literature on the topic has focused around the claim that 'correct' implies 'ought to be done'. The argument from correctness-conditions to semantic normativity – called the 'Simple Argument' for normativity by Glüer and Wikforss (2009b) – originates in Boghossian (1989a). We will see, however, that there is no clear way of formulating such obligations that are clearly semantic, thus the putative normativity of meaning is really a number of hypothetical obligations, which arise from communicative pressures, in disguise. We will also question whether 'correctness' as it appears in 'correctness-conditions' is really a normative notion, or whether it rather performs a categorising or sorting function. Finally, we will assess to what extent the correctness of an utterance is purely a matter of semantics, and argue that pragmatic concerns are equally important here. From this is follows that correctness does not flow from meaning alone, and this gives us a reason to reject the claim that meaning itself is normative. At the end of the section, we will argue against Weiss' (2010) attempt to rehabilitate this 'correct implies ought to be done' notion of the normativity of meaning.

In §3.4 we will examine another influential suggestion by Boghossian: that the normativity of meaning is closely bound up with the idea of motivation. Roughly, his thought is that if it is correct that I apply an expression only to a certain class of things, then I will be *motivated* to do so. Following Miller (2007; forthcoming) we will see that any argument to non-factualism from such a view is

premised on quite implausible claims about motivation, and thus meaning is not normative in this sense. Finally, in §3.5, I will outline a proposal that has received surprisingly little critical attention given its prominence in Kripke's text: the idea that meaning-constituting facts may be appealed to in order to *justify* our use of expressions. I will first consider the most straightforward way of understanding this contention, which is that meanings provide *justifying* or *normative reasons* for linguistic usage. I will outline two opposed accounts of justifying reasons from Smith (1994) and Dancy (2000). I will argue that on Smith's account, it is highly implausible that meanings could provide justifying reasons, whereas on Dancy's account there is no possibility of mounting an argument from justification by meanings to semantic non-factualism. I then consider two arguments from Gampel (1997) and Zalabardo ([1997] 2002) which take the justificatory nature of meaning to pose a problem for reductive naturalistic accounts of meaning. I will argue that neither amounts to a convincing *a priori* argument against the possibility of meaning-facts. In §3.6, I conclude the chapter by considering the remaining challenges for a reductive/naturalistic response to the sceptic which follow from the weakest notion of the normativity of meaning: that expressions have correctness-conditions.

3.2. Trans-temporality

The first attempt to provide an answer to the question of what semantic normativity might be is due to Colin McGinn (1984). McGinn's characterisation of the normativity thesis has rightly been universally dismissed but the reasons are instructive, so I shall outline his view briefly here. It should be said that what is truly baffling about McGinn's view is not that it is so obviously incorrect, but that he could come to believe that it is the idea that Kripke is working with *given* what he (McGinn) says elsewhere in his book. The characterisation of normativity that McGinn ends up with seems to be totally at odds with his earlier remarks on the topic, which are largely on target as an exegesis of Kripke.

So let us begin with the earlier remarks. Given the amount of flak that McGinn's view has received (as we will soon see), I think it only fair to quote these longish remarks in their entirety. McGinn writes:

According to Kripke...Wittgenstein focuses attention upon the normative notion of an application of a sign being (linguistically) correct, i.e. in accordance with its meaning. (This is not the notion of *factual* correctness, i.e. stating a truth about the world; it concerns the question of *which word* is linguistically appropriate to the facts. Thus, for example, suppose I believe truly that this object is red; the question of linguistic correctness is then which word expresses this belief: is 'red' the word I ought to use to state the fact in which I believe?) We ordinarily think that some uses of words are correct and some are incorrect, some uses correctly express the fact we want to state and some do not: Wittgenstein's question is supposed to be what this distinction consists in. What makes it right to use words in one way rather than another? It is clear that this normative property of words depends upon their having a determinate meaning...Therefore we need to make sense of the idea of a word meaning one thing rather than another if we are to give content to the notion of correct (or incorrect) use of language. To put it differently...we must be able to read off from any alleged meaning-constituting property of a word what is the correct use of that word. The normativeness of meaning thus functions as a condition of adequacy upon any account of what meaning is (McGinn 1984: 60-1).

There are several important strands to pull apart here. I am not in full agreement with McGinn on the first: I think McGinn is right to hold that the notion of correctness in question is linguistic correctness.[1] However, I don't think he gives quite the right account of what it is to use a word in accordance with its meaning – at least so far as Kripke's argument is concerned. The picture McGinn suggests is this: we have the expression, 'red', it has a meaning, *red*, therefore if I want to express the belief that some object is red, the word I ought to use is 'red' – i.e. 'red' is the word that is linguistically appropriate to the expression of my belief. However, for Kripke the picture is more individualised: *I* mean something, say *red*, by an expression, 'red', thus the linguistically appropriate word to express my belief that this object is red is 'red'. The important difference comes when my meaning is 'deviant' – such as were I to mean *red* by 'green'. Now, in Kripke's terms the linguistically appropriate expression for me to use in expressing my belief that this object is red is 'green'! This, in fact, is akin to the first statement of the sceptical challenge – perhaps, the sceptic suggests, the answer I *ought* to have given should have been '5'! (Cf. Kripke [1982] 2003: 8) Thus the question is not: which word is appropriate to the facts, but rather, what is the appropriate use of a word for me to make *given what I mean by it*. This distinction will be of crucial importance when we discuss Boghossian's view in §3.3.

---

[1] There is a further move required to get us from the fact that a use of an expression is linguistically correct to the claim that such a use therefore *ought* to be made. I leave this aside for now as we can adequately discuss McGinn's view without worrying about whether 'correct' implies 'ought'. We will return to this at great length, however, in §3.3.

However, McGinn's next remark is absolutely on target: meaning is indeed supposed to demarcate correct from incorrect use; but this again needs to be expressed in terms of an individual to fully capture Kripke's view. Thus *what I mean* by an expression is supposed to determine *which of my uses of it are (in)correct*. And here (in)correct means: in accordance (or not) with the meaning I have 'attached' to the expression. McGinn's third point is, again, completely correct as an exegesis of Kripke *if* read at the level of the individual: we must be able to read off from any alleged meaning-constituting fact *in virtue of which a speaker means something by a word* what is the correct use of that word ('correct' here is again shorthand for 'correct for that speaker *given* their meaning').

McGinn's foregoing remarks come from the second chapter of his *Wittgenstein on Meaning* which offers an exegesis of Kripke's reading of Wittgenstein. As I have argued above, these remarks are largely on target as an exegesis of Kripke, and I include them here to be maximally fair to McGinn. Now, where McGinn seems to go seriously wrong is in his chapter four, which offers an assessment of Kripke's arguments. According to McGinn-in-chapter-four, the notion of normativity in play in the sceptical argument is a *trans-temporal* one: 'the idea of present use being in accord with past meaning' (1984: 174). Thus we have accounted for normativity when we have provided:

> (a) an account of what it is to mean something at a given time, and (b) an account of what it is to mean the same thing at two different times (McGinn 1984: 74).

McGinn's own response to the sceptical paradox is to invoke the notion of a capacity. Thus he meets the two conditions above by stipulating that (a) to mean *plus* by 'plus' at time t is to associate the capacity to add with 'plus' at t, and (b) to mean the same thing (*plus*) by 'plus' at t^* is to associate the same capacity to add with 'plus' at t^* (Ibid).

Now, in fairness McGinn's reading is naturally suggested by the framing of the sceptical challenge: is not the whole sceptical argument framed in terms of a *change* in meaning (from *quus* to *plus*)? Are we not challenged to demonstrate that such a change has not taken place? Moreover, he is right to say that normativity concerns present use being in accord with past meaning. Where he goes seriously wrong is when he imagines that this can be accounted for by citing

a fact about the speaker which is present at two arbitrary times. In fact, what needs to be characterised is how the fact that constituted my meaning in the past can determine that my present use is (or is not) in accord with my past meaning. In other words, what determines whether my present use is in accord with my past meaning is *my past meaning* (and, accordingly, the fact which constitutes it), and *not* my present meaning (nor the fact which constitutes that)! The shape of a straight response to the sceptic ought to be of the following form: assume that no change of meaning has taken place; cite a fact in virtue of which you meant *plus* in the past; ensure that that fact is such as to determine that *now* you ought to answer '125'. Thus the claim that there is a meaning-constituting fact about me which obtains at t and t^* is the *starting assumption* in such a response, and not its end result.

McGinn's notion of normativity is faulty for two principal reasons, which Boghossian (1989a, 1989b) has pointed out. Recall that the sceptic raises the normativity of meaning as an obstacle to a dispositional theory of meaning. The trouble is, on McGinn's construal, normativity is no obstacle to a dispositional theory at all![2] For suppose we were to insist that (a) to mean *plus* by 'plus' at t is to be disposed to respond to addition queries with sums at t, and (b) to mean *plus* by 'plus' at t^* is likewise to be just so disposed. Thus a dispositional response can meet the sceptic's challenge as easily as the capacity response can. If a dispositional response can escape the normativity argument so easily, why on earth would Kripke think that all of his objections to dispositionalism 'boil down' to it? (Kripke [1982] 2003: 24)

Secondly, and perhaps worse, there is Boghossian's extension of this first objection: the normativity thesis that McGinn presents not only allows dispositionalism to meet the normativity requirement, it allows *any* candidate theory to meet it. This is because:

> Any candidate theory of meaning that provided an account of what speakers mean by their expressions at arbitrary times – however crazy that theory may otherwise be – would satisfy McGinn's constraint (Boghossian, 1989a: 513).

---

[2] Wright [1989] 2002, p. 110 cf. Boghossian 1989a; 1989b.

So suppose I were to insist that a speaker meant *green* by 'green' at *t* if that speaker is warm blooded at *t* and that the same speaker meant *green* by 'green' at *t** if that speaker is warm blooded at *t**. According to McGinn's reading, that would be a sufficient response to the normativity argument. Thus McGinn's reading in fact 'places no substantial constraints on the choice of a meaning constituting fact' (Miller 2007: 222).

So, clearly McGinn's account of semantic normativity is wrong – or, at least, not what Kripke is talking about. Now with regards to our four questions above: it is clear that, given the ease with which a dispositional account can meet McGinn's constraint, there is no argument from normativity to non-factualism to be had here. That, in effect, accounts for questions 2 and 4. With regards to the question of whether meaning could be said to be normative in McGinn's sense, however, matters are not so clear cut.

While I think the above criticisms are decisive against the normativity requirement that McGinn presents, I also think that the aspect of trans-temporality which underlies his conception should not be completely dismissed. It seems to me that McGinn is right to draw our attention to the trans-temporal dimension of normativity; sadly, he misconstrues the nature of it. According to McGinn, the trans-temporal aspect of normativity is accounted for when we are able to cite a meaning-constituting fact which is present at any given time. This is incorrect. The true trans-temporal aspect of normativity is this: if a speaker means something by an expression at a given time *t*, then, provided they not wish to change their meaning, certain uses of that expression will be correct at any given time *t**. Consider the case of Jones who means *plus* by '+' at *t*. Then:

> If asked: '57+68=?', then, providing no change in meaning occurs for Jones: It would be correct for Jones to answer '125' at *t*, and it would be correct for Jones to answer '125' at t_1, and it would be correct for Jones to answer '125' at t_2…and it would be correct for Jones to answer '125' at t_n.

This is the trans-temporal aspect of normativity that the sceptic focuses on: if I mean something by an expression at a time, then only certain uses of that

expression will be correct for me to make *at any given time*, *unless* I intend to bring about a change in my meaning. Thus Kripke writes:

> [The sceptic] merely questions…whether I am *presently* conforming to my *previous* linguistic intentions. The problem is…"How do I know that '68 plus 57', as I *meant* 'plus' in the *past* should denote 125?" (Kripke [1982] 2003: 12).

What we are after, then, is a fact which implies that '125' is the linguistically correct answer for me to give *whenever* I am faced with the addition query '68+57=?'. In this sense it would have an 'infinitary normative character' to use Boghossian's phrase (1989a: 542).

Now, given all of this, I think it becomes clear that trans-temporality enters into the account of normativity in this way: if I mean something by an expression *e* at a time *t*, then it is correct for me to use that expression in certain ways at *t* but also, provided I do not intend to make a change in my meaning, at any arbitrary *t** in the future. And this is a very natural thought: it would seem remarkably odd to suggest that a meaning-constituting fact implied *only* that my *current* use was correct, but had nothing to say about past and future uses. So there remains, I think, something for any theory of meaning to account for: the notion that meaning demarcates correct from incorrect use, and not just in the current case, as it were, but in any uses of an expression that I might make. This thought – that meaningful expressions possess conditions of correct application that outstrip my current use and, indeed, all uses of an expression that I might make – is in fact shared by normativists and non-normativists alike; in fact, it is regarded as platitudinous by many philosophers of language. However, as we shall see, this platitude is the starting point for our next normativity argument.

3.3. 'Correct' implies 'Ought'

The second view to be considered stems from Boghossian's influential reading of the normativity constraint. As I noted in the introduction, Boghossian's view comes in two parts. The first is recognition of the fact that meaningful expressions possess conditions of correct application (correctness-conditions), and the claim that this generates 'normative facts' about how such expressions *ought* to be used. The second is to argue that we are motivated to apply expressions only

in those ways that are correct. We'll consider the second strand of Boghossian's view in §3.4. For now, let us attend to the first strand.

Boghossian writes:

> Suppose the expression 'green' means *green*. It follows immediately that the expression 'green' applies *correctly* only to *these* things (the green ones) and not to *those* (the non-greens). The fact that the expression means something implies, that is, a whole set of *normative* truths about my behaviour with that expression: namely, that my use of it is correct in application to certain objects and not in application to others. This is…a relation between meaning something by [an expression] at some time and its *use at that time*. The normativity of meaning turns out to be, in other words, simply a new name for the familiar fact that…meaningful expressions possess conditions of *correct use* (Boghossian, 1989: 513).

He goes on to spell out how he thinks such a thesis is intended to play out in the sceptical argument:

> Kripke's insight was to realise that this observation [that is, that meaningful expressions possess correctness-conditions] may be converted into a condition of adequacy on theories of meaning: any proposed candidate for the property in virtue of which an expression has meaning, must be such as to ground the "normativity" of meaning – it ought to be possible to read off from any alleged meaning-constituting property of a word, what is the correct use of that word (Boghossian, 1989: 513).

A few elucidatory comments about this are in order. First, note that when Boghossian talks about the correct use of a word, he really means the correct application of it. We will look at this distinction in much more detail shortly; for now, suffice it to say that 'application' is essentially a cover-all term for semantic relations such as *reference*, *denotation* and *predication*. By contrast, 'use' is a broader term that covers all uses of language, some (perhaps many) of which will not be applications (cf. Millar 2002: 158 f). Secondly, note that Boghossian frames the normativity of meaning as a relation between *an expression's meaning* and the *applications of that expression that it is correct to make*. Now, this is not clearly the notion operating in Kripke's dialectic. There Kripke argues that it is *what a speaker means* by an expression that carries the normative implications. Thus we must be careful to distinguish putative norms arising from *expression-*

meaning from putative norms arising from *speaker-meaning*.[3] (We will return to this distinction later.)

Note that, thus far, Boghossian has made no mention of *prescriptions*; that is, (in this case semantic) obligations that are typically captured in 'deontic vocabulary' (Glüer & Wikforss 2009a). Thus it is initially tempting to respond, as Fodor does, with simple puzzlement about what 'grounding the normativity of meaning' really amounts to, beyond offering a successful naturalistic account of such semantic relations as reference, extension and so on. He writes:

> Requiring that normativity be grounded suggests that there is more to demand of a naturalised semantics than that it provide a reduction of such notions as, say, *extension*. But what could this 'more' amount to? To apply a term to a thing in its extension *is* to apply the term correctly; once you've said what it is that makes the tables the extension of 'table', there is surely no *further* question about why it's *correct* to apply 'table' to a table. It thus seems that if you have a reductive account of the semantic relations, there is no job of grounding normativity left to do (Fodor, 1990: 135-6).

Now it may be that, say, a dispositionalist account, or an information-theoretic account of semantic relations such as Fodor favours, *cannot*, as a matter of fact, provide a reduction of semantic relations such as extension. But, of course, this has to be *shown* via a detailed consideration of such programs – they cannot be knocked down by the *observation* that meaningful expressions possess correctness-conditions, because this is something they are intended to account for! Thus if there is to be a quick argument from normativity to non-factualism, or an argument that broadly dispositional theories of meaning cannot account for the normativity of meaning *in principle*, then it surely cannot consist simply in this observation.

Now clearly Boghossian is aware of this – as we have seen he adds an element of motivation to his account – but there is a more immediate line of

---

[3] One cautionary note here: I do not intend speaker-meaning as necessarily being a matter of, say, Gricean intentions. The notion is rather to be understood as something like 'the literal meaning of an expression in a particular speaker's idiolect'. Thus, for example, suppose that Jones really does mean some bizarre things by his expressions such that by 'neat' he means *impeccable* and by 'handwriting' he means *dress-sense*. Now suppose that he is asked to comment on Smith's application for a lectureship in philosophy and says, 'He has neat handwriting'. In the sense of speaker-meaning that I intend here, Jones means that *Smith has impeccable dress-sense*. Of course, he may further *imply* by this that *Smith is not a very good philosopher* – but this is not the intended sense of speaker-meaning here.

argument that he suggests. With reference to the sophisticated dispositional account of meaning, he writes:

> Kripke seems to think that even if there were a suitably selected disposition that captured the extension of an expression accurately, that disposition could still not be identified with the fact of meaning, because it still remains true that the concept of a disposition is descriptive whereas the concept of meaning is not (Boghossian, 1989: 532).

One might plausibly understand this passage as implying something like the following: there is a sense in which meaning is *prescriptive*; thus one would expect ascriptions of meaning to carry *prescriptive* consequences. Following this line of thought, Glüer and Wikforss (2009a; 2009b) construct the following argument (the 'Simple Argument' for normativity):

The Simple Argument

P1: For any speaker S, and at any time t: if 'green' means *green* for S at t, then it is correct for S to apply 'green' to an object x at t iff x is green at t.

C: For any speaker S, and at any time t: if 'green' means *green* for S at t, then S ought to apply 'green' to an object x at t iff x is green at t.

They make the important observation that this argument rests upon two further assumptions: first, that C follows from P1 'directly', that is, 'without the help of any further (substantive) premises'; and, second that the simple argument depends upon it 'being a conceptual truth' that the notion of correctness employed in P1 is normative (Glüer and Wikforss 2009b: 35).

I will return to the Simple Argument shortly. First, let us consider the corresponding argument from normativity to non-factualism about meaning – call it the 'Crude Argument' – which would then run as follows:

P1: Ascriptions of meaning to speakers' expressions carry prescriptive consequences about how such expressions *ought* to be used.

P2: Meanings can have prescriptive consequences only if the facts which constitute them likewise have such consequences.

C1: Any candidate meaning-constituting fact must therefore carry such prescriptive consequences.

P3: No candidate meaning-constituting fact carries such consequences.

C2: There are no meaning-constituting (and thus no meaning-) facts.

Now if supplemented with a couple of assumptions, this argument *does* tell against the ability of such theories as Fodor's to meet the sceptical challenge. For one might think, following Hume, that no number of facts about what *is* the case will allow us to infer a conclusion about what *ought* to be the case. So *even if* a theory of meaning manages to account for the semantic relations successfully, if that theory carries no *prescriptive* consequences about how speakers *ought* to use their expressions, then that theory will have failed to satisfactorily account for meaning.

In fact, one may be able to bolster this crude argument by looking further afield. In recent work, Miller (2010) has investigated how the meaning-sceptic might adapt Mackie's (1977) error theory of moral judgment in order to mount an argument against semantic realism. Let us consider how such an argument might go now. First, we will need to attend to the distinction between hypothetical and categorical norms. This distinction will be useful not only in understanding Mackie's argument, but also in drawing out some important implications for the normativist. Consider:

(1) You ought to take the 52 bus to Crookes.
(2) You ought to help another human being in distress.
(3) You ought to apply the word 'green' to an object only if that object is green.

Now, there are important differences between these statements. Sentence (1) expresses a hypothetical norm. One can, to use Miller's phrase, 'release oneself from the scope of the "ought"' in (1) by 'citing some contingent fact about one's

actual desires' – for example, 'But I don't want to go to Crookes!' If you don't desire to go to Crookes, then it is not the case that you ought to get on the number 52! However, the 'ought' in (2) is not like that; if (2) is true, then we cannot release ourselves from our obligation simply by citing contingent facts about our desires (Miller, forthcoming). We may say that the 'ought' in (1) is *hypothetical*, whereas the 'ought' in (2) is *categorical*. So an initial question is this: is the 'ought' in (3) like that in (1), or that in (2)?

For the normativist's thesis that 'correct' implies 'ought' to amount to anything substantial, it must be the case that the 'ought' in (3) is categorical like the 'ought' in (2). Why is this? Well, first of all, as Boghossian points out, if semantic 'oughts' turned out to be merely hypothetical – perhaps contingent on our desires to communicate successfully, or tell the truth, or whatever, then meaning is normative in the same way that *every* fact is potentially normative:

> Of course, we can say that, if you mean addition by '+' and have a desire to tell the truth, then, if you are asked what the sum of [68 and 57] is, you should say '125.' But that is mere hypothetical normativity, and that is uninteresting: every fact is normative in that sense. (Compare: if it's raining, and you don't want to get wet, you should take your umbrella.) (Boghossian 2005: 207).

Secondly, as Hattiangadi points out, the 'oughts' involved in hypothetical norms are not *genuine* (2006: 228). That is, you cannot incur a genuine obligation to, say, invade a country by *desiring* to do so (this problem is known as 'bootstrapping'.) This is important in connection with P3 of the crude argument I sketched above.

As it stands P3 of the argument carries little general plausibility. It just isn't clear that *no* account of meaning-constituting facts construes such facts as having prescriptive consequences. However, we may limit this premise to *naturalistic* accounts of meaning. We will be concerned with exactly what 'naturalism' amounts to in chapter 5, but here we may note that, as commonly understood, naturalism is the doctrine that all phenomena may be accounted for in terms of the entities, properties and relations invoked in our theories of the natural sciences (Loewer [1997] 2006: 108). In any event, naturalists 'do not think that meaning, intentionality, or normativity can be among the building blocks of the universe' and thus seek reductive accounts of these notions (Hattiangadi 2006:

220). Now such a view of naturalism is clearly at odds with existence of facts with prescriptive consequences *so long as* those prescriptions are categorical and not hypothetical:

> The distinction [between hypothetical and categorical 'ought's] is decisive, since many hypothetical 'ought's pose no difficulty for naturalism. The reason is that on our usual, normative interpretation of 'ought', hypothetical 'ought'-statements seem to be plainly false. How can it be that someone's desire or intention to do something makes it the case that she ought to do it…Hence, as R. M. Hare suggested, the force of many hypothetical 'ought'-statements must be descriptive, rather than prescriptive…Thus it is important that the 'oughts' foisted on the naturalist are not hypothetical oughts, and thus descriptive (Hattiangadi, 2006: 228).

In the cited passage, hypothetical 'ought'-statements are said to be false because of the bootstrapping problem (the oughts are not genuine), and thus they merely describe means to ends rather than being truly prescriptive. So it is clear that, if normativity is to be a problem for naturalism, it has to be categorical normativity with prescriptive force. Now, let us return to Mackie and construct an explicit argument from this kind of normativity to non-factualism about meaning. This is Mackie's argument from queerness.

In Mackie's terms, moral facts provide us with objective, categorical obligations. Thus if an agent S judges that Φing is good, and S's judgement expresses a belief, then it follows that S has a reason – in Kantian terms, a categorical reason – to Φ. As Miller (2010) notes, in such a case S would be unable to release himself from his obligation to Φ by reference to any of his intentions or desires. Now, what is the argument from such a view to non-factualism about morals?

According to Mackie, there simply *are no* categorical reasons for action. He gives two arguments in support of this bold claim, one metaphysical, and one epistemological. The metaphysical objection is that moral facts (were they to exist) would provide them who knew about them with 'both a direction and an overriding motive' as such facts would have 'to-be-pursuedness' or 'to be doneness' somehow built into them (Mackie 1977: 40). The problem here, as Miller points out, is that when we ordinarily look into states of affairs in the world (or, we might say, when we consider 'naturalistic' facts), 'they do not seem to have such demands built into them' (Miller 2003a: 117). Thus moral facts would

be 'of a very strange sort, utterly different from anything else in the universe' (Mackie 1977: 38). The epistemological objection is that, given their peculiar metaphysical status, we could only come to know about moral facts via 'some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else' (Ibid.).

Given that there are no categorical reasons for action, Mackie concludes that all beliefs pertaining to categorical reasons are uniformly and systematically false. This is known as an error theory of moral facts. Now, if meaning-facts provide us with categorical reasons for action, as the normativist must maintain for reasons outlined above, then it seems clear that Mackie's error theory would apply, *mutatis mutandis* to meaning as well. So, we have here perhaps a stronger argument to the effect that, if meaning is normative (read, categorically prescriptive), then there are no meaning-facts.

So, we have seen that the first strand of Boghossian's view consists in the observation that meaningful expression possess conditions of correct application, and that these carry the implication that certain applications of expressions ought (or, perhaps, ought not) to be made. This only follows if one accepts that it is a conceptual truth that the notion of correctness in play here is a normative (prescription-giving) one. (As it happens, we have not yet been given a reason to think that it is.) We have then constructed a crude argument that purports to show that no theory of meaning is able to account for the normativity of meaning (thus construed) and that, therefore, non-factualism about meaning follows. We then bolstered this argument with a variant on a stronger one from Mackie. The question is: how are we to resist the slide from normativity to non-factualism?

As I noted above, it is open to a non-naturalist to simply deny P3 of the crude argument – they might try to offer an account of meaning that does make use of semantic/intentional notions, and thereby ground some kind of normativity. However, since I do not think that meaning is normative, I am not going to concern myself with arguing against P3 – I merely wanted to highlight the fact that such a response is clearly open to the non-naturalist. Instead, I want to attack P1 (and thus C1) of the Crude Argument, essentially by challenging the inference from P1 to C in the Simple Argument. If the move from 'correctness' to 'oughts'

or 'prescriptions' is unwarranted, then neither the Crude Argument, nor the stronger argument from Mackie, should give us cause for concern.

In what follows, I will offer four lines of response to the Simple Argument. The first questions the scope of the deontic operator 'ought' in the conclusion of the argument. We will see that a *narrow* reading of the operator either yields unmanageable obligations or else fails to place substantive constraints on one's use of expressions. A *medium* scope reading of the operator similarly places unmanageable obligations on speakers, while on a *wide* scope reading, one's semantic obligations are too easily met. The second line of response questions whether 'correctness', as it appears in the Simple Argument, is really a normative notion. I offer some considerations for thinking that it is not. The third response extends the notion of correctness-conditions to sentences, and shows that the normativist seems to have to commit to the counter-intuitive view that one ought to always speak the truth, or that lies, irony, humour and so on are semantically forbidden. The fourth response charges that our very assessments of the (in)correctness of utterances of sentences depend crucially on pragmatic features of the context in which they are uttered, and thus cannot be seen as issuing from meaning itself. I will then, for good measure, offer a direct response to the Mackiean argument. Finally, I will consider Bernhard Weiss' attempt to rehabilitate the 'correct' implies 'ought' construal of semantic normativity. I will argue that his attempt is unsuccessful.

First Response: Questions of Scope

The first problem has to do with the scope of the deontic operator 'ought' in the conclusion of the Simple Argument. The basic worry here is that any plausible unpacking of the conclusion will yield only obligations that are so weak as to be unable to support the full blown normativity thesis (Glüer and Wikforss 2009a; re. Bykvist & Hattiangadi 2007: 280). Let us examine more closely how this line of thinking is intended to work.

Recall the conclusion of the Simple Argument:

C: For any speaker *S*, and at any time *t*: if 'green' means *green* for *S* at *t*, then *S* ought to apply 'green' to an object *x* at *t* iff *x* is green at *t*.

Now, as it stands, the consequent of the conditional in C is ambiguous with respect to the scope of the deontic operator. We can unpack it in three ways, giving various degrees of scope to the 'ought':

The Narrow Scope Reading

If 'green' means *green* for *S* at *t*, then (*S* ought to [apply 'green' to an object *x* at *t*] iff *x* is green at *t*.)

The Medium Scope Reading

If 'green' means *green* for *S* at *t*, then (*S* ought to bring it about that [*S* applies 'green' to an object *x* at *t* iff *x* is green at *t*.])

The Wide Scope Reading

S ought to bring it about that [If 'green' means *green* for *S* at *t*, then (*S* applies 'green' to an object *x* at *t* iff *x* is green at *t*.)]

As Glüer and Wikforss (2009a) note, in the debate over semantic normativity, all three of these construals may be found. However, all three face some considerable difficulties. Let us consider them in turn.

The Narrow Scope Reading at least has the advantage that the 'ought' in question can be 'detached'. If we expand the biconditional, we have it that:

(*x* is green at *t* → *S* ought to [apply 'green' to *x* at *t*]) & (*S* ought to [apply 'green' to *x* at *t*] → *x* is green at *t*)

So from '*x* is green at *t*' it follows that there is a genuine obligation on *S*: *S* ought, in this case, to apply 'green' to *x*. Thus the 'ought' in question is detachable from the conditional. This captures the intuition that one's semantic obligations regarding 'green' can only be discharged in one way – by applying 'green' to green objects (Bykvist & Hattiangadi 2007: 283).

However, Hattiangadi points out that there are two fairly glaring problems with this principle. First, it means that anyone who means *green* by 'green' is obliged to apply 'green' to *every green object in existence*. This clashes with the widely held principle that 'ought' implies 'can'. Hattiangadi therefore concludes on the basis of this problem that the narrow scope reading of the biconditional is not only too demanding, but also false (Hattiangadi 2007: 180).

Notice also that there is no analogous problem with the *premise* of the Simple Argument. The premise – let us refer to it as *Correctness* for ease of reference – has it that:

Correctness

For any speaker *S*, and at any time *t*: if 'green' means *green* for *S* at *t*, then it is correct for *S* to apply 'green' to an object *x* at *t* iff *x* is green at *t*.

Hattiangadi writes:

According to *Correctness*, if I understand 'horse' to mean *horse*, [it is correct for me to apply 'horse']…to all and only horses. There is no analogous difficulty [to that outlined for the Narrow Scope Reading above] in saying that the correct application of 'horse' outstrips what I can do – 'correct' or 'refers' does not imply 'can' (Hattiangadi, 2006: 227).

The second problem for the Narrow Scope Reading is that the normativist wants to say that *S* has an obligation to use, say, the word 'green' in certain ways because of what she means by 'green'. As Hattiangadi notes, however, any obligation to call something 'green' is conditional, on the narrow scope reading, on some object *x* being green. Put another way, *S* has an obligation to use 'green' in certain ways *only in certain conditions* – specifically, iff there is some green object to which she is applying the term (Hattiangadi 2006: 225 fn. 4). Presumably the normativist does not want our semantic obligations to be conditional in this way.

One reason would presumably be that, on such a view, it would be very difficult to work out when we would ever have a semantic obligation regarding expressions like 'unicorn', 'Pegasus', 'Sherlock Holmes', 'The present king of France' and so on. Surely there is a sense (perhaps more than one) in which these

expressions can be used incorrectly – and, in particular, it certainly seems that they can be applied incorrectly. However, we cannot get a prohibition against applying them incorrectly out of the above conditional, because what follows from, for instance, *x* being non-green is just that it is not the case that *S* ought to apply green to it. Consider:

(*S* ought to [apply "green" to *x* at *t*] → *x* is green at *t*)

~(*x* is green at *t*)

~(S ought to [apply "green" to *x* at *t*]

This conclusion is consistent with it being permissible that *S* apply 'green' to *x* (Bykvist & Hattiangadi 2007: 280). Thus on the Narrow Scope Reading of the deontic operator, we are allowed to apply words to things that are not their referents, or in their extensions. This surely cannot be what the normativist is after.

This is one reason why the strategy of replacing the biconditional in the Narrow Scope Reading with a weaker 'only if' principle (as is advocated by Whiting, 2007) does not provide the normativist with the conclusion they want. Of course, it does at least solve the problem of an unmanageable obligation. If we have it simply that for any speaker *S* who means *green* by 'green':

S ought to [apply 'green' to *x* at *t*] → *x* is green at *t*

then clearly their obligation (such as it is) is at least manageable. However, as Glüer and Wikforss (2009a) point out, it is unclear that this weaker principle is able to support the normativity thesis. First of all, one cannot derive any obligation of the form '*S* ought to Φ' on the basis of it. Secondly, as we have seen, one can't derive something of the form '*S* ought not to Φ' from it either. Glüer and Wikforss argue that, given that it is permissible, on such a view, that *S* apply 'green' to a non-green thing, (thereby doing something semantically incorrect by the premise of the Simple Argument), it does not follow on such a view that doing something semantically incorrect is doing something one ought not to do. Thus '"semantically incorrect" and "ought not" come apart' (Ibid).

At this point it is open to the normativist to rephrase the narrow scope reading such that rather being read as prescriptive or permissive, it is read as *prohibitive*. Thus, for instance, we could replace the conditional:

S ought to [apply "green" to *x* at *t*] → *x* is green at *t*

with the negative one:

x is not green at *t* → *S* ought not to [apply "green" to *x* at *t*]

Notice first that this prohibitive principle is *not* the contrapositive of the prescriptive one. Were we to negate the left hand side of that conditional we would get 'it is not the case that *S* ought to [apply "green" to *x* at *t*]' and this is not what the prohibitive normativist requires.

Now the main problem with the prohibitive reading, as Hattiangadi (2010: 92) points out is that it outlaws any non-literal use of language, such as, for instance, lies, irony, or humour. Thus it implies that there is something *semantically* wrong with:

(4) Paris is in London
(5) Snow is green
(6) Coronation Street is a wonderfully deep examination of the human condition
(7) A duck walks into a bar…[4]

Now, surely, (4) – (7) are perfectly *meaningful*; they aren't just gibberish as (8) – (9) are:

(8) At at the the the if then circle
(9) Sincerity is green

Of course, there is an important difference between (8) and (9): (9) at least obeys the rules of English syntax; but it is nonsense nonetheless. Now, surely the fact

---

[4] No, I'm not going to attempt a joke here.

that (8) and (9) are nonsense whereas (4) and (5) are false and (6) is ironic[5] has at least something to do with the fact that (4) – (7) are semantically in good order; whereas (8) and (9) are not. (It seems to be a precondition for being in semantically good order that an utterance, at least more or less, follows the rules of syntax). Clearly any principle that implies that there is no semantic difference here is false.

In fact, Chomsky (1965: 152) has argued for various semantic principles that explain why certain (syntactically) well formed sentences are nonetheless meaningless in terms of semantic relations *between* words. Consider:

(10)

      a. Sincerity may virtue the boy.

      b. Sincerity may elapse the boy.

      c. Sincerity may admire the boy.

      d. Sincerity may frighten the boy.[6]

As Collins (2008: 165) points out, these sentences range from being perfectly meaningful (i.e. fine from a semantic point of view), to being complete gibberish. He argues that one might have the intuition that (d) is the most acceptable and (a) the least; and that this may be explained in terms of certain semantic relations between the words of these sentences. Thus 'admire' requires an animate subject; 'elapse' is an intransitive verb; and there is no verb at all in (a) (thus the rules of syntax are broken here). The point here is that this has to do with *relations between words* and not *relations between words and things* such as truth and falsity. Thus the prohibitive semantic 'oughts' envisaged above are simply too crude to make room for the idea that lies, irony and humour can be semantically fine (indeed, must be so) in a certain respect, even though what they say of the world is not (literally) true.

As we have seen, a Narrow Scope Reading of the normativity thesis is untenable. What about the Wide and Medium Scope Readings? Gampel (1997: 228) and Millar (2004: 168-69) endorse the Wide Scope Reading since it yields a

---

[5] At least *I* think I'm being ironic here – I won't argue the point with any die-hard Corrie fans!
[6] The examples are from Chomsky 1965: 152.

disjunctive obligation. On such a view, one can discharge one's semantic obligations regarding 'green' either by applying 'green' to green objects, or by not meaning *green* by 'green' (Glüer & Wikforss 2009a). I will discuss the Wide Scope Reading shortly; the crucial feature I wanted to bring up here is the way that it can be discharged: either by applying 'green' thus and so, *or* by not meaning *green* by 'green'. This shows that for any obligation of the form:

S ought to bring it about that $(p \rightarrow q)$

S has two options regarding how to discharge the obligation: either by bringing about q, or by bringing about $\sim p$. What I want to do here is apply this idea to the Medium Scope Reading.

Hattiangadi (2006: 225, fn. 4) argues that the Medium Scope Reading better captures the intuition that meaning involves prescriptions, because such obligations, on this reading, are not contingent on, for example, x being green. However, she spots a further problem that confronts the Medium Scope Reading. Recall that we have it that, for any S who means *green* by 'green' at t:

(S ought to bring it about that [S applies 'green' to an object x iff x is green at t.])

Unpacking the biconditional gives us:

S ought to bring it about that ([x is green at $t \rightarrow S$ applies 'green' to x at t] & [S applies 'green' to x at $t \rightarrow x$ is green at t])

Assuming that 'S ought to bring it about that (A & B)' entails 'S ought to bring it about that A & S ought to bring it about that B', we have it that:

S ought to bring it about that ([x is green at $t \rightarrow S$ applies 'green' to x at t]

The trouble here, as Hattiangadi says, is that it will, in many cases, not be up to S whether x is green. So the only option she has for discharging her semantic obligation is to apply 'green' to all those things that are green and she cannot make not green! Once again, this is too demanding: S cannot meet the obligation, and, since 'ought' implies 'can', the principle is false (2006: 227).

What happens if we replace the principle with a weaker one such that *S* ought to bring it about that they use an expression *only* when it is correct to do so? We might capture this with the *other* side of the biconditional in the Medium Scope Reading thus:

S ought to bring it about that [*S* applies 'green' to *x* at *t* → *x* is green at *t*]

Does this give us a good account of semantic normativity? Recall that, on the Medium Scope Reading, this is supposed to state a necessary condition on meaning 'green' by *green*, thus:

S means *green* by 'green' at *t* → S ought to bring it about that [*S* applies 'green' to *x* at *t* → *x* is green at *t*]

But, Hattiangadi points out; it is simply *false* that it is a necessary condition on meaning, say, *green* by 'green' that one ought to apply 'green' only when it is correct to do so:

> It is not the case that in order for someone to mean something by a term, it is necessary that she ought only to use it correctly. Under some circumstances, I might be obligated to tell a lie, which does not imply that I mean something non-standard by my expressions…[a prescriptive view of semantic standards] makes it a necessary condition of meaning something by a term that a speaker ought to speak the truth. But this requirement is too strong to be a purely *semantic* requirement (Hattiangadi, 2006: 227).

Suppose I see a performance that is absolutely terrible. In my idiolect, the word 'terrible' means *terrible* and the word 'wonderful' means *wonderful*. Now, suppose that a friend of mine was in the performance and asks me what I thought about it. It seems that, for the sake of my friendship, I ought to tell him a lie in this case. Suppose that this is what I do: I say, 'It was wonderful!' My friend would naturally understand my expression as meaning that *the performance was wonderful*. However, since according to our prescriptive principle it is a *necessary condition* on my meaning *wonderful* by 'wonderful' that I ought to apply 'wonderful' correctly, it seems that on this occasion I *fail to mean what I usually mean by 'wonderful'*. This seems flatly false. So the Medium Scope Reading

seems to fare no better than the Narrow one. Now, let us return to the Wide Scope Reading.

Application and Use

In order to properly assess the Wide Scope Reading, we need to return to a distinction that loomed large at the beginning of our discussion of Boghossian, the distinction between *application* and *use*. The Wide Scope Reading of semantic normativity is supposed to capture something like the putative normativity of *use* rather than the putative normativity of *application*. I'll first run through the distinction, and then outline how the Wide Scope Reading is supposed to do this.

Consider the following:

(11)   The tree in that field is an oak.

(12)   The sky is blue.

(13)   That's Gordon Brown.

(14)   Were that an oak, I would have cut it down.

(15)   Is the sky blue?

(16)   Are you Gordon Brown?

In 11-13, the expressions 'an oak', 'blue' and 'Gordon Brown' are used and applied, whereas in 14-16 they are not applied, though they are used. Thus use is wider than application: applications are a class of uses. Crucially, for Millar 'use' covers *interpretative* uses of language. For example, when you say to me 'Go and fetch the cheddar from the fridge' *I* use, but do not apply, the word 'cheddar' in interpreting what you mean (Millar 2002: 60).

As we have seen above, the notion of application is essentially the notion of predication/reference/denotation. A natural thought is that an application of an expression is correct iff the expression is true of the thing it is applied to. It seems to follow from this that a correctly applied sentence is simply a true sentence. (More on this in a later objection). An application may occur, however, within an expression that is not truth apt. Consider:

(17)   Could you pass me the horseradish sauce?

(18)   Is the man who is drinking a martini your supervisor?

In (10) the expression 'horseradish sauce' is applied to whatever the speaker is asking for, and in (11) 'martini' is applied to whatever is in the glass of the man in question.

Now, a misapplication of language, fairly intuitively, is a case where an expression is applied to something that it is not true of. For example, someone predicates a property to x that x does not have, or refers to an x as a y. What is a misuse of language? This is usually understood (following Burge 1979) as a use of an expression which is *not in accordance with its meaning*. What could this mean? Glüer and Wikforss (2009a) offer the following helpful elucidation: misuse of language relates to mistakes concerning 'which expressions are "appropriate" or "suitable" for expressing a certain belief'. This is essentially the notion of linguistic correctness that we saw McGinn appeal to in §3.2. We can make more sense of this idea by considering some examples where misuse and misapplication come apart.

Suppose that Jones means *horse* by 'horse', and that he sees a horse in a field, correctly identifies it as a horse, wants to tell his companion who is called Smith (originally enough), and says, 'Lo! A horse!' In such a case, given his meaning, Jones correctly applies and correctly uses the word 'horse'. Now, if we tweak the example a bit, such that Jones misidentifies a cow in a nearby field as a horse (Jones and Smith having recently vacated their local), we can find a case where Jones, uttering, 'Lo! A horse!' misapplies but does not misuse the word 'horse'. Crucially, he says something *false* about the animal in the field (he applies 'horse' to a cow), but he does use the appropriate word for expressing his belief that there was a horse in the field. We can also tweak the example the other way round as well. Suppose that Jones sees a horse in the field, mistakes it for a cow (they've really been at the drink, this pair), and utters, 'Lo! A horse!' Now here Jones' expression 'horse' is applied correctly to the thing in the field, so this is not a case of misapplication, *but* Jones has not used the appropriate word for expressing his (false) belief; thus he has misused the word 'horse'.

Now I have framed all of these cases in terms of Jones' idiolect – it is not necessary to bring in the idea of a public language to motivate the distinction between misuse and misapplication, although examples naturally suggest themselves. To take Burge's famous case, a patient goes to the doctor with pain in his thigh, which he describes as arthritis. Whether this is a misapplication, a misuse, or both depends upon the perspective we take. From the point of view of the patient, who means *any painful condition of the limbs* by 'arthritis', then it is neither. From the point of view of the doctor, the patient's use is certainly a misapplication. Were the doctor to describe the patient's condition as 'arthritis' it would be *both*, since for the doctor 'arthritis' means *a painful swelling of the joints*. (Still it's not beyond the imagination to suppose that the doctor might jokingly refer to his 'arthritis patient' with another member of staff.)

Millar (2002, 2004) has a specific thesis about what misuse of language consists in. Thus 'misuse', for Millar, describes cases where an expression is used in a way that does not respect its correctness-conditions (Millar 2002: 61). His conception of the normativity of meaning is this: expressions have certain senses. Whenever a certain sense of an expression is 'in play' we commit ourselves to respecting its conditions of correct application (2002: 69). Thus we have the disjunctive obligation either *not to use* the expression in that sense, or, if we do so, to respect its conditions of correct application. Remember that this covers interpretative use of language as well as an expressive uses one might make. So I respect the correctness-conditions for an expression, while not *applying that expression*. Let us illustrate this with an example. Suppose you ask me to get some cheddar from the fridge. If I am not to misuse the expression 'cheddar' when you ask me this (that is, if I am to respect the correctness-conditions of 'cheddar' in this case; to use 'cheddar' in accordance with its meaning), then I am committed to interpreting your request in a certain way (i.e. as pertaining to a certain kind of cheese).

So the general idea is this: we are obliged such that we can either use 'green' with the meaning *green,* thereby incurring a commitment to respecting its correctness-conditions; or we can choose not to mean *green* by 'green'. Now, let us return to the Wide Scope Reading.

Wide Scope Reading

S ought to bring it about that [If 'green' means *green* for *S* at *t*, then (*S* applies 'green' to an object *x* at *t* iff *x* is green at *t*)]

Now, as we have put it, this is not quite an accurate rendering of Millar's position as it is framed in terms of *application* and says nothing about *use*. A better principle would be:

Wide Scope Reading*

S ought to bring it about that [If 'green' means *green* for *S* at *t*, then (*S* commits to respecting the correctness-conditions for 'green' *viz.* 'green' applies to an object *x* iff *x* is green)]

In Wide Scope Reading*, in order to meet her obligation to respect the correctness-conditions for 'green', *S* must, *inter alia*, commit to a certain pattern of *applications* for 'green', just as Wide Scope Reading requires. However, she must *also* commit to *using* 'green' only in certain ways, and this involves uses of language that are *not* applications (such as interpretations).

The problem for either construal of the Wide Scope Reading, however, is that it is all too easy to meet one's semantic obligations on it. Let us return to Kripke's example. Take Jones the quadder who means *quus* by '+'. Now, there is a more usual usage of '+' to denote the addition function. So we have it, on our principle, that:

Jones ought to bring it about that [Jones means *plus* by '+' → (Jones applies '+' only to the addition function)]

To discharge the observation, Jones can either use '+' in ways suitably appropriate to its meaning (i.e. to denote addition) *or* he can not mean *plus* by '+'. The trouble is that Jones the quadder meets this constraint! Since he doesn't mean *plus* by '+', he fulfils his semantic obligations regarding the expression '+' with the meaning *plus*.

There is another worry here. The normativity of meaning is supposed to issue, surprisingly enough, from meaning itself, and it's just not clear that, on the

Wide Scope Reading, it does. It is not a consequence of Jones meaning, say, *green* by 'green' that he incurs an obligation regarding his use of 'green'. Rather, his obligation is prior, as it were, to any fact about his meaning: he's obliged such that *if* he's going to mean *green* by 'green', then he'd better respect 'green''s correctness-conditions. But the obligation is *separate from* his meaning *green* by 'green'. Thus it is hard to see how a meaning-constituting fact is supposed to have this kind of normative consequence.

A third problem is this: it seems that we are heading for a clash of semantic principles over correct use and correct application. Recall that misuse of an expression is said to occur when someone selects the wrong word to express a certain belief that they might have. Millar cashes this out in terms of not respecting the correctness-conditions of that word. The trouble is that in certain cases, such as when I misperceive, any obligations I might incur to *use* and *apply* an expression may come apart. Consider: if I misperceive a horse as a cow, I accurately report my belief that the animal I am currently seeing is a cow if I say 'Lo! A cow!' Thus to avoiding misusing the word 'cow' here, I ought to describe the horse as a cow; however, in doing so I have certainly misapplied the expression. So the basic problem seems to be this: how do I respect correctness-conditions when applying my expressions? Am I obliged to correctly *apply* those expressions and so always speak the truth, *or* am I obliged to always accurately report my beliefs? Either option seems too strong a requirement.

Second Response: Is Correctness a Normative Notion?

In a sense, the first objection regarding exactly how our semantic obligations are best formulated becomes irrelevant if the inference from correctness-conditions to prescriptions does not go through in any event. And, as numerous commentators have pointed out, it simply isn't clear that 'correctness', as it is employed in the simple argument, is a normative notion.[7] Glüer and Wikforss rightly point out that there are plenty of respectable uses of 'correct' that are *non*-normative, such as *categorising* and *sorting*. Thus it is just not obvious that 'correct' as used in P1 is normative. As Hattiangadi puts it:

---

[7] Cf. Glüer 1999; Dretske 2000; Wikforss 2001; Hattiangadi 2006; Glüer & Wikforss 2009a.

> Sometimes to say that something is right [i.e. correct] does *not* imply a prescription; rather, it is to say that it meets a certain *standard*…To say that some use of a term is 'correct' is thus merely to describe it in a certain way – in light of the norm or standard set by the meaning of the term (Hattiangadi 2006: 223-25).

Thus, as Miller puts it,

> To say that something meets a particular standard isn't by itself to endorse (or condemn) that standard: what is said is purely descriptive or value-neutral (Miller 2010: 7).

In fact, if we consider the kinds of uses of language to which the label of 'correctness' is most obviously applicable (i.e. applications), then, as Boghossian points out 'correct' may just mean 'true' and 'there is no obvious sense in which truth is a normative notion' (Boghossian 2005: 207). In other words, correctness-conditions may simply be truth- or satisfaction-conditions, and there is nothing obviously normative about these (Ibid: 208).

This line of thinking against the idea that it is a conceptual truth that correctness is a normative notion has been challenged by Whiting (2007). He argues that while *whether* or not something meets a specified standard is indeed a matter of what descriptions are true of it, *that* it *does* in fact meet the standard 'is clearly a normative matter; in addition to the descriptive statement it also immediately implies a statement about what one ought to or may do' (Whiting 2007: 135). Thus, for Whiting, a statement is normative if it is, 'or immediately implies, a statement about what we ought (not) to or may (not) do' (Ibid: 134).

Now, it is certainly open to challenge that a statement concerning the meaning of an expression *immediately* implies a statement about what we ought (not) to, or may (not) do. Miller (2010) and Glüer (1999) have both expressed the idea that, in addition to a standard being in operation, it is also necessary for us to take a certain sort of attitude towards the standard in order for us to derive a prescription from it. Glüer puts this thought thus:

> That something is used as a standard for…sorting and classifying actions does not imply that this something is used as a normative [i.e. prescriptive] standard. A normative standard would be one that specifies which actions are good, and which actions should be carried out…[To get semantic prescriptions] one would have to show that the speaker must take a prescriptive attitude towards the standard that merely sorts applications into correct and incorrect ones. It follows that even the

concept of *linguistic mistake* is not necessarily a concept with prescriptive content (Glüer 1999: 166-7; as cited in Kusch 2006: 60).

The important point is that one must take a prescriptive attitude towards a standard in order to derive prescriptive consequences from it. As Miller puts it:

> [Normative consequences do not immediately follow from that fact that a standard is "in force" because] we also require some additional consideration which implies that [the standard] expresses a norm that we ought to subscribe to (Miller 2010: 9).[8]

Weiss (2010) finds this line of reasoning utterly implausible. He argues that 'we do not need to support the *ought* derived from a norm [standard] by a claim that we ought to subscribe to the norm' (2010: 145). Consider: if I am playing squash and my serve goes above the out line, then I ought to concede the point to my opponent. According to Weiss, Miller's claim that we require additional considerations which imply that we ought to subscribe to the semantic norms/standards in question as on a par with the question of whether I 'ought' to be playing squash, which simply doesn't arise here: I am playing squash, so I ought to concede the point. Compare: I am speaking English, therefore I ought to speak thus and so.

To respond, it is necessary to return to the distinction I raised above with respect to Boghossian's exegesis of Kripke between *expression-meaning* and *speaker-meaning*. Weiss writes that the important question here is whether '*qua* speaker of [say] English, the correctness-conditions of *that language* deliver any *oughts*' (2010: 145, my emphasis). Now, clearly Weiss is viewing semantic normativity as operating on the level of *expression-meaning*, so the question of whether I ought to subscribe to the relevant norms is the question of whether I ought to be speaking English – not a sensible question at all. However, in Kripke's dialectic it is *speaker-meaning* that is supposed to have normative implications. And here matters are not so straightforward. Consider:

(1) Jones means *quus* by '+' → Jones ought to answer '5' when asked '68+57=?'

(2) Jones means *blue* by 'green' → Jones ought to apply 'green' to x only if x is blue.

---

[8] Page references for Miller 2010 are from a pre-publication manuscript.

Now these are clearly semantic norms that are consistent with Kripke's arguments. The first is simply taken from *WRPL*. Indeed, the possibility of having such an obligation is crucial to the framing of the sceptical argument. But – and this is the crucial point – such obligations cannot arise for, say, Jones *qua a speaker of English*. Why not? Well, because there is no such English word as 'green' meaning *blue*, or 'plus' meaning *quus*. Thus the semantic norms in question do not arise due to the literal meanings of words in public languages. Indeed, the possibility of someone attaching *deviant* meanings to their expressions is crucial for the sceptical argument to get off the ground.

I suggest that Weiss' arguments against Miller only seem persuasive because of a conflation of 'literal' expression-meaning and speaker-meaning. The 'facts' about literal meaning in a language such as English may well imply prescriptions if we are aiming to speak that language. However, Kripke's dialectic works at the level of the individual idiolect of a speaker. And the question is whether the facts about the meanings of words-in-their-idiolect have normative implications. It is correct for Jones, who means *quus* by 'plus' to respond with '5' to '68+57=?'. Ought he to do so? Probably not – and certainly not *qua* speaker of English: *qua* speaker of English, what he ought to do is change the sense he attaches to 'plus', because that phonological pattern *in English* means *plus*, not *quus*.

I think we have a conflation of certain considerations about misuse and misapplication generally here. Philosophers have tended to think that the concept of meaning is meaning-in-a-public-language. Now, I'm not going to raise any objections to this idea here (although I will return to the question of the constitution of public languages such as English, and how these relate to the individual idiolects of speakers in chapter 6.) However, I think it is clear that it is only by conflating the idiolect and the public language pictures that certain ideas about normativity become plausible.

One prominent theme in philosophical discourse is the idea that language is essentially for communicating (see Smith 2004: 150 re. Davidson 1994: 234). When anti-normativists have insisted that one can derive consequences about how

one ought to use expressions *only if* one supplements facts about meaning with, say, pragmatic concerns such as the desire to communicate, this has been met with some suspicion. Isn't it just obvious if you're speaking a public language like English that you incur some semantic obligations whether you want to communicate or not? Well, maybe. But the facts about what constitutes an expression of a public language are not our central concern here. We are concerned with the idiolect of an individual and whether she incurs any genuine obligations from the facts which constitute her meaning.

Consider Jones the quadder again. Now, as I have noted, there is no such expression as 'plus' with the meaning *quus* in English. If anything, we might say Jones is under some pressure *qua* English speaker to mean *plus* by 'plus' – certainly if he wants to go about his shopping successfully, or get on in mathematics. But this is not our question. Our question is whether *qua* speaker of Jones' idiolect, Jones incurs semantic obligations from the meanings of his expressions. Once viewed in this light, I think it becomes fairly obvious that semantic norms become a veritable *barrier* to communication. Consider Jones, who means *horse* by 'horse', and Smith, who means *cow* by 'horse'. In the sense in question, Jones *ought* to apply 'horse' only to horses, and Smith *ought* to apply 'horse' only to cows. This surely does not facilitate communication between Smith and Jones. (Nor does it present an insurmountable difficulty either – Smith and Jones could presumably work out what's going on in a favourable environment – one with a horse or a cow in it, say.) Now I agree that making use of a public language like English requires one to speak in certain ways in order to be sure of being understood. But this has nothing to do with the constitution of speaker meaning *per se*. Thus I think such considerations fail to convince that meaning is essentially normative.

Third Response: Correctness and Truth

We turn now to a problem that looms large for any account of the normativity of meaning that cashes out the notion in terms of correct application. So far, we have focused on the (in)correctness of applications of predicates or singular terms. We have not, as yet, said anything about the (in)correctness of 'applications' of sentences. Now, it would certainly seem to be the case that if

applying an expression correctly is a matter of applying it to its referent or to something in its extension then it follows that the correctness-conditions for sentences would have to be their truth-conditions. Thus to 'apply' a sentence correctly is to *utter a true sentence*.[9]

Consider the sentence 'grass is green'. Now, by the premise of the Simple Argument, we have it that a speaker who means *grass* by 'grass' applies 'grass' correctly only to grass. Similarly, if that speaker means *green* by 'green', then they apply 'green' correctly only to green things. So, on the assumption that an utterance of a sentence is correct iff each term in that sentence is applied correctly, it can only follow that an utterance of 'grass is green' by a speaker who means *grass* by 'grass' and *green* by 'green' is correct iff 'grass' is applied to grass and 'is green' is applied to something that is indeed green (in this case, grass). In other words, an utterance of the sentence 'grass is green' is correct iff *grass is green* – and this is just the *truth-condition* for 'grass is green'.

The specific problem for the normativist here is this: if they insist that one ought to utter a sentence iff one's utterance is correct, then they are committed the claim that you ought to utter all the true sentences of your language or idiolect. As Hattiangadi points out, this violates the generally held principle that 'ought' implies 'can' (2010: 92). If, on the other hand, they hold that one ought to utter a sentence only if such an utterance is correct – that is, only if it is an utterance of a true sentence, then, as we have remarked, they 'commit to the counter-intuitive view that lying, irony, and humour are in some sense, semantically forbidden' (Ibid). Furthermore, recall Hattiangadi's objection that a semantic principle such as:

S means *green* by 'green' *S* → (ought to apply 'green' to *x* → *x* is green)

makes it a necessary condition on meaning *green* by 'green' that one ought to apply 'green' correctly. Now this objection clearly recurs at the level of sentences.

---

[9] I think it is correct to view *utterances* of sentences as being the appropriate object of assessments of (in)correctness rather than sentences themselves. Consider the analogous case for predicates: it makes no sense to ask if 'green' is correct. What one *can* ask is whether a particular *application* (predication, referential use) of 'green' is correct. Or, to put it another way, one may ask 'when is it correct to apply "green" to an object?' Presumably the idea of an *application of a sentence* would be a sincere utterance of that sentence. Thus it seems that it makes sense to ask in the sentential case 'when is it correct to (sincerely) utter "grass is green"'.

That is, if we apply similar principles to that above to all of the component terms in a sentence that a speaker might utter, then it seems that a necessary condition on that speaker meaning what they usually do by those component expressions is that they ought to apply all of those expressions correctly; i.e. that they ought to tell the truth.

There are two problems here. First, as Hattiangadi points out, it just isn't the case that whenever we are not obliged to tell the truth we start attaching deviant meanings to expressions (2006: 227). Consider the case of a novelist who writes fiction. Surely in writing their novel, they are not obliged to write the truth, but it does not follow that every expression used in their novel has a non-standard meaning. The second problem is simply a continuation of the first. As we have noted, if it is a necessary condition on speaking meaningfully that one ought to tell the truth, it seems that, in any situation when we are not so obliged, we fail to speak meaningfully. Thus not only are lies, humour, and so forth semantically *forbidden*, they are also, in certain circumstances, *impossible*: if I am not obliged to tell the truth in circumstances *C*, then I cannot lie in *C* because in order to lie I must first speak meaningfully.

Fourth Response: Correctness and Context

In our second response we questioned whether 'correct' as it appears in the Simple Argument was a normative notion. In very recent work, Hattiangadi (2010) has adopted a different line of attack. She now argues that *even if* correctness *is* a normative notion, the intuitive (in)correctness of an utterance of a sentence is *not* determined purely by the semantic features of that sentence. Thus even if our utterances are truly described as (in)correct, and this is a normative matter, we still do not secure the conclusion that *meaning* is normative. This is essentially because the (in)correctness of an utterance of a sentence is as much a matter of pragmatics as it is of semantics. At best, the semantics of a sentence are a *factor* in determining whether an utterance of that sentence was (in)correct. I will now examine this line of argument in more detail.

In our third response above we argued that the correctness-conditions for utterances of sentences would have to be the truth-conditions of those sentences. This followed from the claims that:

(a) it is necessary for an utterance of sentence to be correct that every term in that sentence be applied correctly; and

(b) were all the terms in a sentence to be applied correctly, the sentence could not but be true.

This led to some fairly awkward consequences for the normativist. However, one might start to wonder whether the claim that an utterance of a sentence is correct iff the sentence is true is quite right. First let us formalise this for ease of reference:

N: *u* of *s* is correct iff *s* is true.

Hattiangadi notes that N actually flies in the face of many of our intuitions about the correctness of utterances of sentences. She appeals to examples from the literature on contextualism to bring this out.

One set of examples involves cases of 'context-shifting' – cases 'in which an utterance of a sentence, resolved for indexicality, ambiguity, vagueness and so forth, still seems to express different propositions in different contexts' (Hattiangadi 2010: 94 f). Consider:

(1) Every bottle is empty.[10]

Now the idea is that the *content* of (1) cannot be determined by its literal meaning. Content here contrasts with *character* (Kaplan 1989): the character of an indexical expression is that aspect of its meaning that remains invariant over contexts; the *content* is what the indexical picks out in a context. Thus the content of 'I am currently typing my PhD thesis' is very different when uttered by *me* than when uttered by *you*, although its character remains the same. However, (1) contains no overtly indexical expressions, so why think that its content cannot be determined by its literal meaning?

---

[10] The example is from Stanley and Szabó 2000: 219-20.

Hattiangadi reasons as follows: intuitively the content of (1) changes according to context, even though the literal meaning of (1) remains fixed. First, we noted that (1) contains no overt indexical expressions that act as 'syntactic markers' of 'context shift'. Thus the content of (1) 'cannot be determined by its literal, linguistic meaning and those features of the context whose relevance to determining content is mandated by overt, syntactic elements of the sentence' (such as indexicals like 'I', 'here', 'now' and so on) (Hattingadi 2010: 95). Thus the content of (1) must be partially determined by further features of the context.

Consider the following example. I am trying to do the washing up, but my family have shoved a load of empty washing up bottles back under the sink instead of throwing them in the bin. (This sort of thing happens all the time.) I may be irritated by this, and when asked what the matter is, I might say, 'I'm trying to do the washing up, but every bottle is empty!' The content of what I say here seems to be: *Every bottle of washing up liquid in the house is empty*. That is presumably the truth-condition of my utterance that would be understood by my interlocutor. It is false iff there is a bottle of washing up liquid in the house that is not empty. As Hattiangadi points out, there is nothing in the 'overt syntactic structure of (1)' that corresponds to the extra conditions in the content of what I say (Ibid). The contextualist argues from this to the conclusion that the context of the utterance must be relevant to determining the '"semantic content" or what is said by [the] utterance' (Ibid).

Now, to come to the point: recall that according to N an utterance of a sentence is correct iff the sentence uttered is true. This follows from the view that we build up to the correctness-conditions of sentences from the correctness-conditions of their component expressions. Thus an utterance of 'every bottle is empty' is correct iff every bottle is empty. However, as Hattiangadi points out:

> If the content of any utterance of (1) is [*every bottle is empty*], then it will be false if there are any non-empty bottles in the universe. And it is plausible to assume that on most of the ordinary occasions on which one might utter (1), there is at least one non-empty bottle in the universe (Hattiangadi 2010: 96).

Thus, for example, in my imagined context above, my utterance is semantically *in*correct. Hattiangadi (2010: 95) considers some further examples which lead to the same conclusion. Suppose that at the end of a party someone utters 'Every

bottle is empty'. Presumably this means that *every bottle of alcoholic drink in the house is empty*. To take another case: suppose that I come round to your house with some bottles to fill up with your home-made sloe gin. You might ask me if I need to empty any of the bottles I brought. If I say, 'Every bottle is empty', then the content of my utterance is: *Every bottle that I brought to your house to fill up with your sloe gin is empty* (Hattiagadi 2010: 95, cf. Stanley and Szabó 2000: 231-32).

Now according to N, *all* of these utterances are semantically incorrect (that is, the sentence that is being uttered is (almost always) strictly speaking *false*), and hence they are all semantically forbidden. However, as Hattiangadi points out, *contra* what N says, it is surely intuitive that these utterances *are* semantically correct in all of the contexts described above. Indeed, asserting the opposite ('every bottle is not empty!') would be a far better candidate for the incorrect thing to say in each of those cases. An utterance of 'every bottle is not empty' in the sloe gin case, for example, would be utterly misleading at best. Thus it seems that we can only conclude that N is false.

How can we remedy the situation? Hattiangadi suggests that one way would be to view the correctness of an utterance of a sentence *not* as issuing simply from the (context invariant) truth-conditions of the sentence, but rather from the content or proposition that an utterance of that sentence expresses in context. We may formalise this thus:

N*: An utterance of *s* is correct iff *p*

where *p* is 'the proposition expressed by the sentence in context'. This avoids the counter-intuitive view that all of the above utterances of (1) are semantically incorrect, at the cost of conceding that correctness rests, in part, on 'features of the extra-linguistic context of [an] utterance'. To repeat: on such a view the (in)correctness of an utterance of a sentence rests not only on the semantic contributions of that sentence's constituents, but also on 'pragmatic features of the context of utterance'. The problem for the normativist, on this view, is that N* does not amount to the thesis that *meaning* is normative: even granting that the (in)correctness of an utterance of a sentence is a normative matter, according to

N* any facts about whether an utterance of a sentence is (in)correct do not immediately follow from the meaning of that sentence (Hattiangadi 2010: 96-7). For meaning to be normative, it would have to be the case that the (in)correctness of an utterance of a sentence followed directly from the meaning of that sentence. As we have seen, however, such a view flies in the face of our intuitions about the (in)correctness of utterances.

Further examples can be brought to bear against the claim that N gives the correct correctness-conditions for utterances of sentences. Hattiangadi considers two further families of cases: incompleteness and inappropriateness cases. Let us briefly consider these in turn. Incompleteness cases are such that the meaning of a sentence fails to determine a full proposition – that is, in such cases the meaning of a sentence is not something truth-evaluable. For example, the sentence 'ibuprofen is better' seems incomplete (Carston 2002: 22) – we need to know *what* it is better than and *for what* it is better in order to get something truth-evaluable. Now, imagine the following case: Jones has a throat infection and his throat is accordingly very swollen. He reaches for some paracetamol. However, Smith (more medically informed than Jones) knows that, in addition to its painkilling properties, ibuprofen is also an anti-inflammatory, so he utters, 'Hold on! Ibuprofen is better.' Now with the context given, we can work out that the content of Smith's utterance is something like *ibuprofen is better than paracetamol for treating cases where there is pain and also some inflammation*.

Now, consider matters from the point of view of the normativist who holds that N is true; that is, that the correctness-conditions for utterances of sentences are determined by the truth-conditions of those sentences. The problem is that 'ibuprofen is better' does not, by itself *have* a truth-condition – we need to supplement it with contextual information to end up with something truth-evaluable. Thus, as Hattiangadi points out, according to N, it is *indeterminate* whether an utterance of 'ibuprofen is better' is correct or incorrect. This seems to be false: Smith's utterance above seems to be perfectly correct. Were he to have made the utterance when Jones was reaching for an antacid to treat his heartburn, it would have been incorrect. Thus, again, N flies in the face of our intuitions.

Inappropriateness cases are those such that, while a sentence may express a full proposition (unlike in incompleteness cases), the proposition that it expresses seems to be at odds with what we would regard as the proposition expressed by an *utterance* of that sentence (Hattiangadi 2010: 99). Consider the following case (adapted from Hattainagdi 2010: 100): Jones calls round to Smith's quite late having had some breakfast, but not had any dinner yet. Smith is preparing food and asks Jones 'Would you like some dinner?' Suppose Jones was to say 'I've eaten'. According to N, this would be a correct utterance: *ex hypothesi,* Jones has eaten breakfast, so *a fortiori* he has eaten at some point in his life, which is all that is required for the truth of his utterance 'I've eaten'. However, this is again highly counter-intuitive: what Jones' utterance seems to express in the context is the proposition that *Jones has eaten recently* – that is, that *Jones has already eaten dinner*, which is, of course, *false ex hypothesi*. N has it that Jones' highly misleading utterance of 'I've eaten' in this context is correct. This can only offer us further reason to reject it as an account of the correctness-conditions of utterances of sentences.

From the four responses to the Simple Argument, it seems that we can draw the following conclusions: first, that the idea that semantic correctness-conditions imply prescriptions is false; second, that any argument from semantic prescriptions to semantic non-factualism accordingly fails to go through; and third, that even if there were any prescriptions regarding how language ought to be used, the (in)correctness of utterances of sentences on which any such prescriptions would rest does not seem to be determined solely by meaning in any event. Thus even if there were such prescriptions, they wouldn't be *semantic* prescriptions. Now, just in case the Mackiean argument from queerness is still worrying you, I will, as promised, offer a direct response to it.

A Direct Response to the Mackiean Argument

Recall that, in the semantic case, the Mackiean argument from queerness is this:

1. If there were any meaning-facts, then these would have to be categorical reasons for action.

2. Categorical reasons for action are metaphysically outlandish and require a far fetched epistemology.

3. There are no categorical reasons for action

4. There are no meaning-facts

How can we resist the conclusion of this argument? I think that Miller (2010) shows us the correct approach, which is to challenge the claim that meaning-facts may be characterised as categorical reasons for action. This, he claims, has 'no plausibility whatsoever'. Meaning facts will only provide us with a reason for action when supplemented by an appropriate desire. As Miller puts it, only when it is given that an agent has 'a desire to communicate, or perhaps a desire to think the truth, or a desire to conform to his prior semantic intentions' does he have a reason to apply an expression in some way (Miller 2006: 109).

Let's consider an example (this one is drawn from Miller (forthcoming)). Suppose that we adopt a 'public language' picture, and say that 'green' means *green* in the public language. Supposedly, this gives Jones a categorical reason to apply 'green' only to green things. (We will say 'only' to green things to avoid Hattiangadi's 'ought' implies 'can' objection.) Now imagine that Smith and Jones are back from the pub, and Smith is having a peculiarly violent episode. He puts a gun to Jones' head, and threatens to shoot him if he refuses to call his red front door 'green'. According to the view of categorical prescriptions outlined above, Jones cannot wriggle out of his semantic obligation to apply 'green' only to green things by citing his desire to keep his brains in. But this seems hopelessly implausible – surely any obligation he may have regarding his use of 'green' simply stems from his communicative intentions and desires, and so is, in any case, irrelevant in this event. Certainly, it would be remarkably odd to view Jones as somehow semantically *blameworthy* for calling the door 'green'.

This contrasts markedly with the moral case. Suppose Smith holds the gun to Jones' head and threatens to shoot him unless Jones shoots their neighbour. Now, we might think, that even *given* his desire to keep his brains in their right place, Jones cannot release himself from his moral obligation not to murder

people; even though we might view Smith's threat as *mitigating* any such action on Jones' part, we would still think that Jones was (at least partially) blameworthy.

I think the argument from queerness loses even more of its appeal when we remind ourselves that the meaning-facts in question are not supposed to be facts *independent* of any speaker (as the moral facts are taken to be independent of any agent), but are rather facts about what speakers mean by their words. Surely it cannot follow from the fact that a speaker of a particular idiolect incurs a categorical prescription to use of the words of his idiolect only in certain ways. I think the foregoing considerations are sufficient to show that the argument from queerness cannot be applied to the semantic case.

Weiss' Attempt to Rehabilitate Normativity

Earlier we saw how Weiss attempted to argue against the idea that we need to take some kind of normative attitude towards standards of correctness in order to derive prescriptions regarding how we ought to use language. It is worth attending to Weiss' attempt to defend the normativity of meaning here in some further detail, as I think a successful response to his line of argument will allow us to block another way in which meaning might be said to involve 'oughts'. Weiss' fundamental contention is that there are certain intentions that are *fundamental* to meaningful use of language. These cannot be explained away (as we saw Miller attempt to do) by invoking further *communicative* desires or intentions of a speaker. Such intentions, according to Weiss, are of the form:

> Meaning-Constituting Intentions
> S intends that their use of an expression *e* be interpretable as meaning that *p*

The idea is that it follows from such intentions that S ought to adhere to a certain *regularity* of use with respect to *e*. And, crucially, this 'ought' is said to be purely linguistic: it is prior to, and separate from, the desire to (say) communicate, speak truthfully, or whatever (Weiss 2010: 136). The intention to communicate successfully is thus analysed as 'a second order intention to be interpreted as one intends' (Ibid). I will now offer three lines of response to Weiss' argument.

There is a largely Kripkean worry lurking in the background here that the intention that one's use of an expression is understandable as expressing a certain meaning already requires some notion of meaning or content to be in play. Thus meaning would be *prior* to the intention, and not a consequence of it. Of course, an intention is not made up of linguistic symbols. However, it does seem that certain (perhaps the more complex of) intentions may require the prior ability to think using a language. This is roughly the idea that one cannot *have* certain complex intentions without being able to articulate them; and this requires a *prior* ability to use meaningful language – at least in thought. Crucially, such meaningful deployment of language surely couldn't be the result of a Weissian intention to be interpretable as meaning that such-and-such as *having the intention* could only be *posterior* to deploying meaningful language. Thus even if such intentions underlie aspects of communication, this would still fall short of securing the conclusion that meaning is normative, as some *non*-intention derived meaning would have to underlie the intentions.

Second Response

Weiss' thought may have an air of plausibility about it regarding spoken language where a speaker is aiming to be understood, but it says little about the meaningful employment of language in thought. It simply isn't at all clear to me that one's linguistic thoughts *get* their meanings in virtue of intentions that one may have. Perhaps Weiss could suggest that my linguistic thoughts get their meanings in virtue of something like the following hypothetical intention: were my thoughts to be articulated, they *would be* interpretable as meaning this-or-that. Perhaps such an intention accompanies my thoughts, but I am unaware of it. We may also ask here: interpretable by whom? If we imagine that my intentions are such that a *radical* interpreter could make sense of what I say (re. Davidson 1984), then it is clear that I do not 'interpret' my own thoughts in the way that a

radical interpreter would: without any recourse to internal/mental evidence.[11] In other words, surely I do not interpret my own meanings on the basis of my external patterns of usage. This opens the possibility that my thoughts may be perfectly comprehensible to *me* – I'm in a privileged position with respect to them, after all – but would *not* in fact be comprehensible were I to simply articulate them. Thus it is not clear that such Weissian intentions are *necessary* for the deployment of language in thought.

Dretske (2000) expresses a similar thought when he writes that a view that tries to derive normative consequences from one's intentions regarding the deployment of concepts is:

> …a confusion between two quite different things – the norm-free concepts (meanings) we intentionally apply, on the one hand, and, on the other, the norm-laden intentions with which…we apply them. All intentional acts, in virtue of being intentional (or goal directed), are subject to norms (Dretske, 2000: 250).

We may even add a caveat to this: intentional acts are subject to *hypothetical norms* and these are *not* genuinely normative. (We will consider this in more detail in our next response.)

Let us consider the following development of Dretske's thought. Here are a couple of conflicting intuitions about meaning. On the one hand, we may say that speakers understand expressions, or *know their meanings*. Imagine that Jones knows that the meaning of 'green' is *green*. Is this fact sufficient for saying that Jones means *green* by 'green'? If so, then it seems that meaning something by an expression – *attaching* a certain meaning to it – is a state of *knowledge*, not an intention at all. And, in fact, one may know the meanings of certain expressions – think slurs and terms of abuse – and have *no intentions whatsoever* to deploy those expressions. I think this even goes for the kind of hypothetical intentions we considered above; those of the form, 'were I to use the expression *e*, it would be interpretable as meaning that such-and-such'. On the other hand, we might view meaning as an intentional *action*; in other words, meaning might consist in *saying that* – in making external *use* of the expression. So even though someone

---

[11] In any case, Piertroski has advanced some compelling arguments aimed at showing a) that human beings are *not* radical interpreters and b) that radical interpreters would *mis*interpret in any case. See Piertroski 2005.

understands an expression as meaning that such-and-such, we might not want to say that *she* means such-and-such by it *unless* she actually deploys the expression in an utterance.

If we go with the former option, and think that someone who understands an expression in a certain way *ipso facto* means something by that expression, then we may draw the distinction between meaning something by (i.e. understanding) an expression, which doesn't seem to require Weissian intentions, and making meaningful communicative *use* of that expression, which perhaps does. In any case, insisting on the need for Weissian intentions in order to make meaningful communicative *use* of language doesn't address the claim that one can understand (i.e. mean something by) an expression without using it in such communicative settings. So I think once again we have a conflation of meaning and communication, understanding and being interpretable. And, once again, we are unable to derive any genuine semantic 'oughts'.

Third Response

Wikforss (2001) has argued against the general idea of drawing normative consequences from intentions. We have already seen that the bootstrapping problem is a block against deriving *genuine* prescriptions from intentions. Wikforss argues further that the view that intentions give rise to prescriptions results from confusing the fact that the relation between an intention and its fulfilment is *internal* with the claim that such a relation is normative. She writes:

> If I intend to eat an apple, then it is not merely an empirical question what will fulfil my intention; rather I *have* to eat an apple (and not, say, a pear) to fulfil this intention. Similarly, if I intend to add, then it is not an empirical question what I should do in order to fulfil my intention; rather I *must* add (Wikforss, 2001: 213).

However,

> The fact that the relation between an intention and its fulfilment is *internal* does not show that it is *normative*. If I intend to eat an apple and I eat a pear instead, then my intention is not fulfilled, but there is not implication that by eating the pear I have failed to do what I *should* do. Similarly, to say that the relation between intending to add and answering '125' to '58+67' is internal is not to say that I ought to answer '125', but that if I do not do so my intention is not fulfilled, that is,

I am not adding. This is not at all a prescriptive claim but a constitutive one (Wikforss, 2001: Ibid).

Is it plausibly *constitutive* of meaning something by *e* that one intends that one's use of *e* is interpretable as meaning that *p*? If we drop any additional claims about any 'oughts' arising from such intentions, this seems to be very close to what Weiss intends to show. However, I am not convinced that it is the case. I think it is perfectly possible to mean something by an expression but have *no* standing intentions whatsoever regarding its use.

3.4. Correctness and Motivation

Boghossian's View Part II: The Motivating Nature of Extensions

It is time to consider the other side of Boghossian's view: the idea that a speaker who means something by an expression will be motivated to use it only in certain ways. Remember that Fodor voiced the worry that giving a naturalisation of relations such as extension ought to be sufficient to account for the 'normativity' of meaning: there is nothing to the notion above and beyond saying what makes it the case that, say, 'green' applies to the green things, and not to the non-greens. Now, there seems to be a key difference in the way that Fodor and Boghossian understand the notion of extension. For Fodor, this is simply a factual matter: 'green' has the green things in its extension, so it is correctly applied to the green things, and it is misapplied to the non-greens. For Boghossian, an extension is something more than this. He writes:

> One might have a thought like this. A proper reduction of the meaning of an expression would not merely specify its extension correctly, it would also reveal that what it is specifying is an *extension* – namely, a *correctness condition*. And this is what a dispositional theory cannot do…To be told that 'horse' means *horse* implies that a speaker ought to be motivated to apply the expression only to horses…[Any successful dispositional theory of meaning] must show how possession of [the relevant dispositions] could amount to something that deserves to be called a correctness condition, something we would be inherently motivated to satisfy (Boghossian 1989: 533-4).

So the 'more' to the normativity thesis that a naturalistic theory of meaning must capture is the idea that meaning is *inherently motivating*. When we understand an expression – or, perhaps better, when we form a belief about what an expression

means – we thereby grasp its correctness-condition, and this in turn *motivates us* to use the expression as the correctness-condition stipulates. This gives us the result that no supplementary desire is necessary to motivate an agent: she will be motivated to speak in certain ways simply because of her *beliefs* about the correctness-conditions of her expressions.

Now one might immediately worry that it is not at all clear how a naturalistic or dispositional theory can account for such a feature of meaning. However, Miller (2007, forthcoming) has explored the possibility that the semantic sceptic can do better still by taking another well know meta-ethical argument, The Argument from Motivational Internalism, and applying it to the semantic case. To see how such a move may be motivated, we will need to first see how the argument runs in the moral case. As a preliminary to this, it will be necessary to sketch the Humean Theory of Motivation, on which the argument partially rests.

The Humean Theory of Motivation

The Humean Theory of Motivation basically consists of the following claims:

(H1)   Beliefs by themselves are causally inert: they cannot cause/move/motivate us to acts.

(H2)   Thus motivation is always a matter of having both beliefs and desires.

(H3)   There can be no necessary or internal connection between any particular belief and desire (beliefs and desires are 'distinct existences').

To see a simple example of this theory at work, suppose that Jones believes that there is beer in the fridge. The idea is that, by itself, this is not sufficient to motivate Jones to leave the comfort of the couch to go in search of beer. Couple this belief with the *desire* for beer, however, and we may well see Jones making his way towards the fridge. Note that there is no internal connection between the belief and the desire: it is not a necessary consequence of Jones' belief that *that's* where the beer is that Jones will desire the beer.

Now, the question is: why might this view of motivation cause a problem for *moral* realism? Miller (2007: 188) provides the following helpful demonstration. Consider:

(P1) The judgement that doing X in C is morally right expresses a belief.

(P2) It is a conceptual truth that: if an agent S judges that doing X in C is morally right, then so long a S is practically rational, S will be motivated to X in C.

(P3) Motivation is always a matter of having both beliefs and desires, where beliefs and desires are 'distinct existences' in the sense that there is no conceptual or necessary connection between any particular belief and any particular desire.

Here P1 expresses a commitment to *cognitivism* – the idea that judgements about what is morally right 'express beliefs, states apt for assessment in terms of truth and falsity' (Miller 2007: 188). Note that any kind of realism or factualism about meaning can be similarly described. The semantic realist[12] holds that 'judgements ascribing linguistic meaning or understanding express beliefs'. This parallel between moral- and semantic-realism will be of crucial importance in what follows. P2 expresses a commitment to motivational internalism – the thought that judgements about what is morally right motivate us to act in certain ways without the need for supplementary desires. (Note again the parallel with Boghossian's view of judgements about meaning.) P3 expresses a commitment to the Humean Theory of Motivation (as outlined above).

Now essentially the problem is this: P2 (Motivational Internalism) stipulates that any practically rational agent who makes a moral judgement will *necessarily* be motivated to act in some way. Thus there is a necessary link between an agent's belief that, say, X is good, and their desire to bring about X. But it is precisely this kind of link that is forbidden by P3 (the Humean Theory). The non-cognitivist thus intends to jettison P1 on the strength of P2 and P3. The cognitivist may hope to respond that making a moral judgement isn't simply a matter of having a belief; rather, it is a matter of forming some kind of belief/desire *complex*. Unfortunately, the Humean has an answer: in order for a judgement to be intrinsically motivating, given that it involves a belief, it must

---

[12] Here just anyone who thinks that there are facts-of-the-matter about meaning, be they a classical realist, or a Dummettian anti-realist.

also involve some motivating state such as a desire. The kind of belief/desire complex envisaged can only fulfil this motivating role if there is a necessary connection between the belief and the desire involved – otherwise it wouldn't be intrinsically motivating. But this connection is *still* forbidden by P3.[13]

It is, of course, not my concern here to argue whether or not this argument holds water in the moral case. However, given Boghossian's remarks about the nature of correctness-conditions above, it does seem that an analogous argument could be run, *mutatis mutandis*, in the case of meaning. Miller (2007: 188-91) runs just such an argument along the following lines: Assume that Jones judges that 'green' means *green* and that this judgement expresses a belief. Then, on Boghossian's account of correctness-conditions, provided that Jones is practically rational, Jones will be motivated to apply 'green' as its correctness-condition stipulates (i.e. to all and only green things), thus:

> Jones judges that 'green' means *green* $\rightarrow$ (provided that Jones is practically rational, Jones will be motivated to apply 'green' to $x \leftrightarrow x$ is green)

But Jones' desire to do apply 'green' thus and so necessarily arose from his belief and this is forbidden by the Humean Theory of Motivation. Thus, by *reductio*, Jones' judgement did not express a belief. Thus semantic realism is false, and the sceptic has a direct argument from normativity to non-factualism about meaning. Now, does it stand up?

I think it is fairly clear that the argument, as applied to the semantic case, does not hold water. The problem, as Miller points out, is that the premise that, having made his judgement, Jones will, as a matter of conceptual necessity, be motivated to use the expression 'green' in certain ways is completely implausible (Miller 2007: 188-91). I deal with a preliminary problem first, and then come on to Miller's argument.

As I've presented the conditional above, we have it that if Jones judges that 'green' means *green*, then Jones will be motivated to go around applying 'green' to each and every thing he sees. This is a natural consequence of taking

---

[13] I am grateful to Bob Hale for pointing out the realist response and the Humean reply.

seriously the idea that we are motivated to apply expressions in the ways that correctness-conditions require. Correctness-conditions can be formulated biconditionally, and will therefore often outstrip the correct applications of an expression that a speaker could possibly make. But it is simply false that one is motivated to point and shout 'green' every time one comes across a green thing (at least, I don't have any such urges!) So to give Boghossian's account any plausibility at all, one must replace the biconditional statement of correctness-conditions in the above conditional with a weaker 'only if' principle:

Jones judges that 'green' means *green* → (provided that Jones is practically rational, Jones will be motivated to apply 'green' to x → x is green)

This at least averts the consequence that practically rational speakers are continually fighting down their desires to apply words to things! However, as we will see, it does not ultimately help the sceptic.

The problem is that, even with this weaker claim in place, it is fairly easy to think of cases where a practically rational Jones will not have any such motivation. To demonstrate this, let us return to the example of Smith's violent episode that we raised as an objection to the argument from queerness. Recall that Smith puts a gun to Jones' head and threatens to shoot unless Jones will call his red front door 'green'. Supposing that Jones judges that 'red' means *red* and 'green' means *green*, then, on our current account, we have it that he is motivated to apply 'red' only to red things, and 'green' only to green ones. However, as Miller points out, it would be crazy in our current case to call Jones' practical rationality into question should he feel no motivation to call the door 'red', or some (perhaps a great deal!) of motivation to call it 'green' (he does prefer his brains where they are, after all). I don't even think that this can plausibly be called a case of one prudential norm trumping a weaker semantic one. Generally speaking, when norms by which one is bound conflict, one feels some measure of discomfort, being pulled in two directions. This does not seem to be the case here. The general point is that Jones' judgement about the meaning of 'red' *simply doesn't motivate him at all*.

It is perfectly possible to have beliefs about what expressions mean while having no desires concerning the use of those expressions whatsoever. Consider again slurs and insults targeted at a certain groups – racist insults, say, or derogatory terms for the disabled. Surely it is false to say that someone who knows what such terms mean is motivated to use them 'correctly'; that is, to apply them only to people in their extensions. A decent person who knows the meanings of such terms is not motivated – that is, feels no 'pull' or inclination – to apply them *at all*. Thus the crucial second premise of the Argument from Motivation Internalism is simply false in the semantic case. As a final note, it is perhaps worth bearing in mind that *even if* it were true, the argument still rests on an acceptance of the Humean Theory of Motivation. It is therefore open to the semantic realist who is also a motivational internalist about semantic judgements to jettison P3 on the basis of P1 and P2 – although, given my anti-normativist tendencies – this is not a strategy that I will pursue here.

3.5. Meaning and Justification

We now move on to our final account of the normativity of meaning: the idea that meanings (and meaning-constituting facts) in some sense *justify* linguistic usage (call this 'justification normativity or JN). Despite the fact that this idea is prevalent in Kripke's text, it seems to have received surprisingly little critical attention in the literature. In this section, I will provide some textual evidence from Kripke to show that some form of JN is clearly being endorsed, and distinguish the notion from the kind of *guiding* or *motivating* normativity that we encountered in §3.4. I will then consider what is *prima facie* the most straightforward way of understanding JN: roughly, that meanings are or provide *justifying reasons for action*. I will contrast two views of such reasons from Smith (1994) and Dancy (2000). I will argue that on Smith's view, it turns out to be quite implausible that meanings are justifying reasons. I will further argue that, on Dancy's view, for meanings to be justifying reasons they would have to be *states of affairs*; thus there is no scope for a non-factualism about meaning arising on the basis of such a view. I will then consider two attempts to mount an argument from a version of JN to put pressure on reductive or naturalistic accounts of meaning from Gampel (1997) and Zalabardo ([1997] 2002). I will argue that neither argument is successful.

A preliminary question is this: how does justification relate to motivation or guidance? It has been remarked that justification and guidance function in Kripke's text as two sides of the same coin (Kusch 2006: 9). In one sense this is quite a natural thought: surely we want to be able to say that, on occasions, agents act for reasons that both motivate/guide them in what they do and justify their action. However, at the same time we must respect the fact that justification and guidance do not always go hand in hand.

Here is a line of argument that suggests that guidance and justification ought not always to be run together. Glüer and Pagin (1999) hold that we are *guided* by reasons that enter into the motivation we have for acting in one way rather than another (1999: 208). They stress that such reasons are what are called 'explanatory reasons' in the literature: I can be guided by a reason (explanatory sense) that isn't a *good* or *justificatory* reason (more on this presently). So, in order for the meaning of an expression – call this *M* – to guide me, it is first necessary that *M* be a reason that goes some way to explaining what I did. However, assuming a traditional belief-desire model of practical reasoning, this by itself is not enough. For the reasons that enter into the standard belief-desire model may be divided into two categories: *motivational* and *doxastic* (211). To play a truly guiding role, *M* must have motivational force: I must act as I do *because of M*. Doxastic reasons, by contrast, play an auxiliary role: effecting 'theoretical transitions' between reasons that *do* have motivational force (218). To illustrate this thought, Glüer and Pagin make use of various practical syllogisms which describe transitions from various pro-attitudes to intentions or from one pro-attitude to another. These take the general form:

(PA)   I want to Φ.

(B)   (This action) φ is a case of Φing.

(I)   So, I shall do φ.

or,

(PA$_1$) I want to Φ.

(B)   I can Φ by Ψing .

(PA$_2$) So, I want to Ψ.

where 'PA' stands for some sort of pro-attitude (canonically desire), 'B' stands for belief, and 'I' stands for intention.[14] The important point for our purposes is that while both pro-attitudes and beliefs are reasons for action, in order to have *guiding* force, a reason for acting in a certain way must occupy a *motivational* position in such a syllogism. That is, it must enter into practical reasoning as a pro-attitude, and not as a belief.[15]

Now, the reason that it might be a mistake to view guidance and justification as naturally running together is that it is all to easy to think of cases where an agent is guided by a reason (in the sense explained above) where that reason *in no way justifies the agent acting as they did*. Consider:

(PA)   I want to vent my frustration.

(B)   Rearranging my squash racquet around my opponent's head is a case of venting my frustration.

(I)   So, I shall rearrange my squash racquet around my opponent's head.

(PA$_1$) I want some heroin.

(B$_1$)   I need money to obtain heroin.

(PA$_2$) I want some money

(B$_2$)   I can obtain the money by robbing the bank.

(I)   So, I shall rob the bank.[16]

In neither of these cases do the motivating reasons that the agent has for acting in accordance with their intentions *in any way justify* their acting in that way. Thus

---

[14] Following Davidson (1963), we may think of the reasons in the first practical syllogism – those that lie immediately behind the intention to act in such-and-such a way – 'primary reasons'.

[15] Glüer and Pagin's immediate concern is to see whether meaning-determining rules (that is, rules that are metaphysically prior to and which serve to determine meanings) could both play the envisaged meaning-*constitutive* role *and* guide linguistic usage. They argue not as no plausible candidate for a meaning-determining rule can 'fit' into a motivational slot in the practical syllogisms.

[16] The former example is based on Watson 1975: 101, and the latter on Frankfurt 1971: 87.

guidance and justification may come apart. I cannot always appeal to what guides me in order to justify what I do. Thus we cannot conclude from the supposed fact that meanings guide use that they also justify it, or *vice versa*.

That guidance and justification can both run together and come apart is, in fact, a puzzle that Smith's (1994) account of normative or justifying reasons is supposed to solve. He writes:

> To the extent that reflection on our normative reasons moves us to act…accepting that we have certain normative reasons must in some way be bound up with having corresponding motivating reasons. But…our desires may come apart from our acceptance of corresponding normative reason claims…The puzzle, then, is to explain how it can be that accepting normative reason claims can both be *bound up with* having desires and yet *come apart from* having desires. In other words, the problem is to explain how deliberation on the basis of our values can be practical in its issue *just to the extent that it is* (1994: 136).

I will now outline Smith's account of normative reasons. and consider JN in this light. I will argue that there is reason to think that meanings are *not* justifying or normative reasons on such an account.

Smith (1994) attempts to find a role for normative reasons in his account of motivation that captures their justificatory force. For Smith, normative reasons are, first of all, *propositions* that express truths of the form:

For any agent A: 'A's Φing is desirable or required' (Smith 1994: 95).

Motivating reasons, by contrast, are 'psychological states' (belief-desire complexes in this case: Smith defends the Humean Theory of Motivation encountered earlier with regard to *explanatory* reasons) that 'motivate actions and which we cite in explaining them' (Lenman 2009: online). However, Smith is radically *anti*-Humean with respect to normative reasons. The crucial step in Smith's theory is the claim that normative reasons do have a role to play in motivating action *via* our beliefs about them (1994: 177-80). As Lenman puts it:

> Normative and evaluative beliefs about what I ought to do or what it would be desirable for me to do [i.e. beliefs about normative reasons] *cause* me, insofar as I am rational, to have relevant desires that then proceed to play their essential role in motivation (Lenman 2009: online).

Let's consider how this works in some examples. We'll take a moral example first, and then see how the account transfers to the semantic case.[17] Consider:

> Here I am believing that it would be desirable for me to Φ. In virtue of having this belief, I am brought to desire to Φ. This desire of mine to Φ together with a belief that Ψ-ing is a necessary means to Φ-ing then *motivate* me to Ψ. The desire of mine to Φ together with the belief that I can Φ by Ψ-ing motivate my Ψ-ing and constitute my motivating reason to Ψ. The belief that it would be desirable for me to Φ *causes* my desire to Φ and features in a *rationalising explanation* of that desire but does not motivate that desire (Lenman 2009: online).

Now, how does Smith's account translate to the case of semantic normativity? Well, there would first have to be some normative truth about meaning which, to use Lenman's phrase, 'spoke to the issue of whether some action was justified'. So, for example, it might be that a normative truth about the word 'green' would be that it is desirable or required to apply 'green' only to green objects. This essentially gives a statement of the correctness-conditions of 'green'. So, we may say that the proposition '"green" means *green*' speaks to which uses of 'green' are correct/desirable/required.

Next, someone who made the judgement that 'green' means *green*, where this judgement is seen as expressing a belief, would, insofar as they were rational, come to have certain desires to use 'green'; specifically, to use 'green' only when they were justified in doing so. This desire along with other beliefs, such as, say, beliefs about which things are green, would then motivate such a person to speak in certain ways. Consider:

> Here I am believing 'green' means *green* and thus applies to all and only green things. In virtue of having this belief, I am brought to desire to apply green (only) to things in its extension. This desire of mine to apply 'green' correctly, together with a belief that applying 'green' (only) to green objects is a necessary means to applying green as I ought then *motivate* me to apply 'green' (only) to green objects. The desire of mine to apply green correctly together with the belief that I can do so by applying 'green' (only) to green objects motivate my applying 'green' in such ways, and constitute my motivating reason to do so.

---

[17] Interestingly, Smith claims that there 'may be normative reasons of rationality, prudence, morality, and perhaps even normative reasons of other kinds as well' (1994: 95-6) ; thus it would seem that his account, if correct, should generalise to the semantic case (if meaning is indeed justificatory or normative in this sense).

Now, to turn to our central concerns: is meaning normative in this justificatory sense? And does this account of justifying reasons in the semantic case give the sceptic what he needs to argue against semantic realism?

My own view is that, no, clearly not. For one thing, justifying reasons here are said to be truths or propositions. Meaning-constituting facts are presumably not propositions; they are facts about speakers and thus entities of the wrong *type* to be normative reasons. Interestingly, we might be able to avoid this objection if we adopt a Platonic view of language such that there *are* 'truths' about the meanings of expressions in the public language. This is of small use to the normativist-sceptic, however, since the very notion of a normative reason as a *proposition which expresses a truth* seems to require the very notion of meaning that he is arguing against.

Secondly, it seems fairly clear that my beliefs about what expressions mean do *not* cause me to desire to use them in certain ways. I've argued this point at some length already, but let me add some considerations in its favour here. I know a few words of German. However, I do not take part in any communication in German, and have no desires concerning my use of the German words I know. However, the relationship between my semantic beliefs and desires, on Smith's account, is *causal*. Thus if meaning were normative in this sense, it ought to be the case either that I am *caused* to desire to apply my German words correctly, or I am not fully rational. However, I just don't have the relevant desires and (I hope!) this does not show that I am not fully rational. It would seem remarkably odd for the sceptic to reply that my beliefs about normative matters such as meaning 'might' or 'sometimes' cause me to have the relevant desires: the connection is not supposed to be probabilistic. If a desire to act in such-and-such a way is always brought about by beliefs about normative reasons, then it just doesn't seem to be the case that facts about meaning are normative reasons.

Notice that the 'public language' picture is not much help here either. If I were to start participating in more German conversations, then I would desire to use my German words correctly. But this desire arises from my intention to participate in this practice: without that aim the desire dissipates. But if meaning is normative in the justificatory sense, this is all wrong: my desire can't just 'go

away' because of what practices I choose to engage in. On a causal account, any beliefs about the meanings of words I may have are going to cause me to desire to do certain things with those words whether I'm participating in certain communicative practices or not. Such a view, it seems to me, is plainly false.

Smith's account has ultimately proved to be untenable in the semantic case. Perhaps we should try a different account of normative reasons. Dancy (2000) adopts an *externalist* account of normative reasons. He is at pains to argue against any view where it is claimed that 'the reasons for which we act are psychological states of ourselves' (2000: 98). So far he may be seen as in accordance with Smith, who similarly argued that *normative* reasons are not psychological states. However, unlike Smith, he doesn't think that motivating and normative reasons are two distinct kinds of beast; nor does he think that propositions are the right kind of thing to be normative reasons. Rather, both normative and explanatory reasons are *states of affairs*, but *considered from different points of view* (Lenman 2009).

The trouble with building any kind of *scepticism* out of such a view is fairly obvious. The fact that, say, Jones means *green* by 'green' could perhaps be seen as a motivating or a normative reason depending on the perspective we take of it. In Dacny's terms:

> When I call a reason 'motivating', all that I am doing is issuing a reminder that the focus of our attention is on matters of motivation, for the moment. When I call it 'normative' again all that I am doing is stressing that we are currently thinking about whether it is a good reason, one that favours acting in the way proposed (Dancy, 2000: 2-3).

So perhaps we can explain why Jones keeps calling cows 'horses' by citing the fact that Jones means *cows* by 'horses'. However, given his non-standard interpretation of 'horse', we might find it hard to justify what he is doing on the basis of such a reason. In other circumstances, it may turn out that we can appeal to Jones' meaning to justify his use. In such cases we might say that his use is correct or appropriate due to the meanings of his expressions. In any event, however, on Dancy's view Jones' meaning *plus* by 'plus' is a *state-of-affairs*, which is to say, a *fact* about the world. If Jones' meaning *plus* by 'plus' justifies his use in this sense, well and good: it can only do so if it is a fact! Thus while on

Dancy's view meanings could be justifying reasons, this in no way lends support to semantic scepticism.

Let us turn now to Gampel's (1997) account of the normativity of meaning. This is based around what he calls the 'essentially justificatory role' of meaning (EJRM). This is given two formulations, one in terms of meaning-constituting facts and one in terms of meanings themselves:

> EJR: Being able to justify S's use of an expression *x* is *essential* to the fact constituting what S means by *x* (1997: 226).

> EJRM: Being able to justify a person's use of an expression is essential to meaning (1997: 229).

The two formulations are, of course, related. I will try to give a brief account of just *how* they are related. Recall that Kripke speaks of two kinds of correctness with regards to his famous plus/quus example. The answer '125' is the correct answer to '68+57=?' first in the arithmetical sense (125 really is the sum of 68 and 57), and second in the 'metalinguistic' sense that it is the answer that corresponds to the *meaning* of the '+' sign, since that sign denotes the addition function. We may say that the answer '125' is *metalinguistically justified* since it accords with the meaning of '+'. Thus meaning is justificatory in the sense that certain linguistic usage is *in accord with* meaning. The idea is that the meaning of the '+' sign 'justifies' or even 'compels' the answer '125' to the query '68+57=?' (Kripke 1982: 11)

Now, according to Kripke, a meaning-constituting fact must have this property: it must show how some linguistic uses are in accord with the meaning that it constitutes, and thus that these uses are metalinguistically justified. So, the thought is that since the putative fact that I mean *plus* by '+' metalinguistically justifies my answer of '125' to '68+57=?', the fact that *constitutes* my meaning *plus* by '+' must itself show that this answer is metalinguistically justified. How can it do this? Here is where a familiar version of the normativity thesis enters play: a meaning-constituting fact for '+' would show that I was metalinguistically justified in returning '125' to '58+67=?' were it to demonstrate that '125' is the

answer I 'ought' to give in order to be in accordance with the meaning of '+'. In less prescriptive language, a meaning-constituting fact must be able to account for certain linguistic usage being *correct* in the sense of being in accordance with what is meant.

I now turn to the *essentialist* component of Gampel's view. In order to provide an *a priori* argument against some form of meaning factualism, EJRM cannot simply amount to the thesis that we *can* appeal to meanings in order to justify our linguistic usage. In other words, something more is needed beyond appeal to the fact that one might justify one's application of the expression 'horse' to a horse via the observation that 'horse' means *horse*. The reason for this is simple: Gampel adopts a view of justifying reasons similar to Dancy's such that anything whatsoever *can* justify a course of action in appropriate circumstances (Gampel 1997: 226). Thus on such a reading, the claim that meaning justifies use is an utterly mundane observation, true of every fact. So, Gampel needs something stronger. What he claims, in essence is this: meaning plays not just a justificatory role, but rather an *essentially* justificatory role. What does this mean? The thought is perhaps best illustrated by contrast with a different case. The fact that it is raining today can justify my taking an umbrella to work. However, it is not *essential* to *x*'s being a fact about today's weather that *x* serve to justify my umbrella related actions. Gampel's claim is thus that it *is* essential to being a meaning-fact that something serves to justify linguistic usage. He ties this claim up with a notion of *guidance*: meanings essentially serve to guide action (229).

Two qualifications are in order here. Firstly, Gampel is not claiming that meanings are categorically normative. That is, whether or not one ought to use an expression in such-and-such a way *is* contingent, according to Gampel, on a whole host of communicative aims and intentions that a speaker may have. In this, he is in agreement with the anti-normativists. Secondly, the kind of guidance Gampel is referring to seems to be distinct from the sort of guidance-cum-motivation discussed above. Indeed, Gampel stresses that he is not claiming that there is anything intrinsically motivating about meaning (229), nor does he claim that meanings are categorically prescriptive (228). Thus Gampel is a very interesting interlocutor in this connection: he claims that there *is* something essentially normative about meaning, but distances himself both from the

Mackiean argument and the argument from motivational internalism that we discussed above.

Gampel's argument from the normativity of meaning targets naturalistic reductions of meaning. The argument is roughly this: meanings are *essentially* justificatory of use: the meaning of 'green' essentially guides speakers in their use of 'green' and thereby justifies certain uses of 'green'. However, the kinds of naturalistic facts to which one might try to reduce meaning[18] do not have this essential property. Consider dispositions to use expressions under ideal conditions. We may *take* these as a guide to when uses of 'green' are justified, but they do not guide us in and of themselves (231). Compare the metre rod: we make *take* it as a guide to measuring distance, but it is not *essentially* any such thing (Ibid). Or consider: we could *take* someone's age as determining whether or not they could legally buy alcohol; but it is not an essential part of being an age to determine whether or not such alcohol-directed activity is justified or not. Thus Gampel's distinction appears to be between those things whose role is essentially to provide standards of 'correct' action and those things that we merely *take* to provide standards.

In order to assess Gampel's account of the normativity of meaning, it is necessary to flesh out the notions of guidance and justification that he appeals to. As noted above, clearly 'guidance' for Gampel is not the notion of guidance that we have discussed which is bound up with motivation. Rather, 'guidance' here seems to mean something like 'instructional': it is the essential role of meanings to *instruct* me as to which uses of expressions are 'correct' and hence justified, relative to various desires/intentions that I may have. Gampel's claim has to be that no naturalistic fact to which we might wish to reduce meaning could ever serve to be an essential guide to what kinds of linguistic use are correct and hence justified.[19] However, to say this is to beg precisely the question at issue in the

---

[18] In fact, Gampel challenges the association of 'naturalism' with 'reductionism' and suggests that a non-reductive naturalism can escape the difficulties posed by EJRM. Since I'm not convinced EJRM really poses a challenge to the reductive theories, I will set this proposal aside.

[19] Pagin (2002) has argued that it is unclear whether 'justifying' in Gampel's sense means 'correctness-determining', or 'citable as a reason'. I think the idea is that since meaning-facts essentially serve to determine correctness-conditions, they can always be cited as a justifying reason for a particular use of an expression that a speaker may make.

sceptical argument, and in the construction of any reductive naturalistic theory of meaning.

Let me explain. Here is one way in which meanings might be said to guide me in my linguistic use: meanings tell me what words apply to and what they don't. That is, if I know the meaning of, say, 'green', then, provided I have some further linguistic abilities, I can tell people the truth: 'grass is green', lie to them 'the sky is green' and so on. Thus facts about what expressions do and do not apply to seem to be the relevant ones when I consider how to act, be this to tell the truth, lie, mislead, or whatever. Thus Gampel's proposal seems to be that no naturalistic fact could serve to essentially demarcate what a word applies to from what it doesn't; that is, to demarcate correct from incorrect use.

However, it seems to me perfectly open to the reductive naturalist to argue as follows: coming up with a fact that determines what expressions apply to and what they don't is the whole *point* of a naturalistic theory of meaning. Supposing that we cite such a fact, wouldn't that be sufficient to account for meaning's essentially justificatory role? That is, to echo Fodor's thought, why wouldn't naturalistically grounding concepts such as extension be enough to account for the 'normativity' of meaning? It seems to me that Gampel is covertly smuggling in normative associations to the notion of a correctness-condition (i.e. to the notion of an extension) that need not be there. The meaning of 'cat' justifies uses of 'cat' to pick out cats (presuming that this is what a speaker desires to do) just because the term 'cat' applies to/denotes/is true of cats, and is false of non-cats. Similarly, if a speaker wishes to lie and blame a small household disaster on the cat, the fact that 'cat' applies to cats again justifies him in saying 'the cat knocked the ornament off the mantelpiece'.

Thus any fact that rules in the Xs and rules out all of the non-Xs as part of the extension of a term is going to be sufficient to play the envisaged justificatory role. So long as we can demarcate application to things within the extension of an expression and things *without* that extension, we're home and dry: if demarcating things within the extension from things without it is something that the found naturalistic fact *just does* (i.e. it does this without us having to *take it* as doing so), then EJR/EJRM seem not to tell against such a naturalistic account of meaning.

Of course, if it turns out that *on inspection*, no naturalistic account can do this, then that's another matter. For now, I wish to make the following two points: first, one cannot simply *pronounce* that no naturalistic fact will turn out to be able to do so; and second, that accounting for the essential justificatory role of meaning does not seem to be an *additional* task to naturalistically grounding concepts such as extension.

Finally, I turn to Zalabado's ([1997] 2002) account of JN. The first feature of his account is a different view to Gampel concerning exactly *what* is justified. For Gampel, it is quite obviously linguistic usage that is justified. For Zalabardo, it is the *procedures* we use for deciding whether or not to apply an expression to something. Suppose I go through some procedure for deciding whether to call some fire engine 'red'. Zalabardo's reconstruction of the normativity argument has it that, in order for my procedure to be justified, I must *consciously engage* with the fact that determines whether or not the predicate 'red' applies here. (Actually there are two facts that determine whether or not the predicate applies: the fact that constitutes what I mean by 'red' and the fact that the fire engine is, indeed, red. It is the former that Zalabardo is interested in here.)

Now the argument runs as follows. Were I to mean *red* by 'red' in virtue of some specific meaning-constituting fact *M*, then in order for my procedure of applying 'red' to the fire engine to be justified, I would have to consciously engage with *M*. However, for a good number of the accounts Kripke considers (and especially with regard to the dispositional account), it simply isn't the case that we so engage: one does not, as a rule, reflect on one's dispositions to apply a predicate – and, *a fortiori*, to one's dispositions to do so under some idealised circumstances – when deciding to apply it in a novel case. Thus on the dispositional account, our procedures for applying predicates turn out to be arbitrary. Similar remarks go for more sophisticated naturalistic responses. For instance, Fodor has it that (mentalese) predicates apply to properties because of a certain type of nomological covariance between instantiations of the property and tokenings of the predicate. However, when deciding whether or not to call the fire engine 'red' I do not consciously engage with nomological covariance between the expression 'red' and the property of redness. (Strictly speaking, for Fodor 'red' applies not to the fire engine *qua* object, but rather to its property of being

red.) Thus, were this nomological covariance the fact in virtue of which 'red' applied here, my procedure would not be justified, it would be arbitrary. Teleological accounts of meaning, which have it that a mental representation applies to a property in virtue of its being selected for representing that property, similarly fall prey to this kind of argument. No one reflects on the selectional history of a mental representation to determine whether or not a predicate applies in a given case. So Zalabardo's argument certainly seems to put pressure on various reductive/naturalistic accounts of meaning. What may we say in response?

I think it is fairly clear that the argument has a false premise: that I must consciously engage with the fact *in virtue of which* a predicate means what it does for my procedure to be non-arbitrary. Why must that be so? Why not simply engage with the fact *that* a predicate means what it does? To give an example: in the fire engine case, why would it not be enough to engage with the facts that:

1. 'Red' means *red*.
2. I perceive that the fire engine has the property of being red.

In other words, engagement with the fact *in virtue of which* (1) is true seems unnecessary here – neither my procedure for deciding to apply 'red' to the fire engine, nor my actual application of 'red' to the fire engine is arbitrary in light of the facts (1) and (2).

Notice further that if we need conscious engagement with the facts in virtue of which (1) obtains, then it seems that we should also require conscious engagement with the facts in virtue of which (2) does. For, surely, I perceive that the fire engine is red because of facts about the environment, the lighting conditions, and the human visual system. Do I need to consciously engage with all of these for my application of 'red' to the fire engine to be non-arbitrary? Surely not. Or compare:

3. Jones *knew* the car parts were stolen.
4. Jones bought them anyway.

On the basis of (3) and (4), a jury's deliberations in deciding that Jones was guilty of a crime would be justified; that is, the procedure they used in deciding that Jones was guilty would not be arbitrary. They wouldn't need to consider the *further* question of in virtue of what does (3) obtain (that's a question for philosophers!) Or take the case of recognising a grammatical sentence. Surely a

sentence is grammatical because of some fairly complex rules or principles of sentence construction in human languages – rules which linguists are working extremely hard to uncover. If my task is to recognise a grammatical sentence, my procedure for doing so does *not* require me to reflect on the facts *in virtue of which* the sentence is grammatical to be non-arbitrary. To put it another way: we surely *can* recognise (un)grammatical sentences, and we do not seem to be making arbitrary judgments about this. However, since most humans are not linguists, it seems many of us couldn't engage with the facts that *determine* whether the sentences are grammatical even if we wanted to! Thus the premise of Zalabardo's normativity argument seems to make unreasonable demands on a candidate meaning-constituting fact.[20]

3.7. Conclusion

In this chapter we have considered a number of ways that the claim that meaning is normative could be cashed out. I have argued that meaning is not normative in any of these senses, and that, further, no knock down argument from the claim that meaning is normative to non-factualism about meaning is in the offing. However, we have also seen that it is widely held that meaningful expressions possess conditions of correct application.[21] Any successful naturalistic/reductive theory of meaning, it would seem, must be able to give the intuitively correct extensions for our expressions. If a theory of meaning fails to clearly do so, then it will be open to the sceptic to interject and claim that sceptical hypotheses concerning deviant interpretations of our expressions are consistent with the candidate theory.

In the next chapter, we will see that several promising naturalistic/reductive accounts of meaning fall prey to exactly this line of argument. Sadly, no extant naturalistic account of meaning can specify the correct

---

[20] There is the epistemological consideration that, however the meaning of an expression is constituted, I must in some sense be immediately aware of the fact *that* the expression has the meaning that it does: I may not know what my meaning, say, *red* by 'red' consists in, but I do know *that* I mean *red* by 'red'. This further constraint on a candidate meaning-constituting fact is sadly beyond the scope of this work.

[21] Note, however, that we did raise a complication regarding the correctness-conditions for utterances of sentences. It seems that these cannot be straightforwardly 'built' out of the correctness-conditions of the component expressions of a sentence without recourse to pragmatic concerns. We will return to this matter in chapter 6.

extensions of our expressions. Because of this, it is open to the sceptic to object that none provides an answer to his challenge. However, I hope in this chapter to have removed a substantial block to the project of making sense of meaning more generally: the false claim that meaning is normative in any interesting sense, and that any candidate theory must make room for such normativity. Once such an *a priori* requirement is removed, it may seem that the problem of correctly capturing extensions is an interesting technical challenge that has not yet been met, rather than a pressing need to ward off the most extreme kind of meaning-scepticism.

Norm Relativity, Extension and the Problem of Error

4.1 Introduction

In the previous chapter, we examined the most prominent suggestions in the literature as to how the claim that meaning is normative may be fleshed out, and I argued that none of these provides the sceptic with a decisive argument from normativity to non-factualism. Furthermore, I argued that meaning is not normative in any of these senses. In this chapter, we turn to the second major focus of this thesis: naturalism in the theory of meaning. We will be concerned with how the most prominent naturalistic responses to the sceptic fare once the normativity constraint has been dispensed with.

While the claim that meaning is normative has turned out to be false, naturalistic theories of meaning still have to capture the idea that expressions have correctness-conditions, where these are not construed as prescription-giving. Thus the most pressing problem for such a theory is the so-called 'problem of error'. The problem, roughly, is this: if expressions have conditions of correct application, and such conditions must in some way be determined by meaning-constituting facts, then meaning-constituting facts must be such as to determine that errors are indeed *errors*. If they fail to do so, then the sceptic can insist that what we consider errors *could*, in fact, be legitimate members of the extensions of our expressions. And we would be left with no factual basis for distinguishing between the two claims. The task of accounting for error, it turns out, is an incredibly tall order.

In order to see whether the problem of error can be solved within a naturalistic framework[1], we will need to look at the extant naturalistic responses to the sceptic in much more detail than we have so far. So, this is how I propose to proceed. In §4.2, I will outline the problem of error in more detail. I will use the dispositional account of meaning as a basis for discussion and consider, in general

---

[1] See Hattiangadi 2007 for arguments to the effect that *non*-naturalistic/-reductive accounts of meaning fare no better in solving the problem of error than the naturalistic ones.

terms, how a dispositional theory of meaning may be modified in order to account for error. This will set the stage for §4.3 where I will examine some of the more influential naturalistic attempts to solve the problem in some detail. I will argue that while some of these attempts clearly fail, none is clearly successful. I will conclude in §4.4 by pointing towards a way in which the meaning-factualist can concede that the naturalistic responses to the sceptic fail, but maintain that there may be reasons for this other than the truth of semantic non-factualism. This response will be developed and defended in chapters 5 and 6.

4.2 The Problem of Error (or the Disjunction Problem)

Let's outline the problem of error in some more detail. Suppose Jones means *horse* by 'horse'. Given that Jones means this, it seems platitudinous to many philosophers to say that Jones' term 'horse' applies to horses, and that Jones himself applies 'horse' to a thing correctly iff that thing is a horse. The problem comes when Jones, suitably inebriated, stumbles across a particularly horsey looking cow, and utters 'Lo! A horse'. This utterance needn't be spoken out loud; in fact, I think the whole problem is far more compelling when set up in relation to tokenings of thoughts or mental representations (speakers not being disposed, on the whole, to go around remarking on the fact that they are faced with certain things they know the names of). So, Jones sees a cow, and tokens the mental representation 'horse'. If his thought/utterance is a labelling one (i.e. if Jones *applies* the token/term to the thing), then we naturally want to say that Jones has made an *error*. Now, here is the problem: how can our theory of meaning entail that horses are in the extension of 'horse' and cows aren't, so that we are able to say what we naturally want to say? (This is, of course, just the familiar problem of accounting for how an expression has the particular extension that it does.)

Recall the crude dispositional theory of meaning. This has it that what Jones means by, say, 'horse' is determined by the uses of 'horse' that Jones is disposed to make. Among these uses are applications, or labelling thoughts. (Recall that 'use' is a broader term than 'application'.) Now, if an inebriated Jones is disposed to apply 'horse' to horsey-looking cows in certain conditions, then the crude dispositional theory cannot do better than to say that what Jones means by

'horse' is the disjunctive property *horse or cow*. (For this reason the problem of error is often referred to as the *disjunction problem*.)

Let us begin with Kripke's discussion of the Crude Dispositional Theory (CDT) of meaning. It is worth emphasising from the outset that the crude theory is *really* crude for two separate reasons. First, it equates what a speaker means by an expression with *everything* that she is disposed to apply that expression to. Thus it is crude with respect to *how* dispositions to use expressions relate to meaning. Second, it equates what one is disposed to do with *everything that one in fact does*. Thus it is crude with respect to the individuation of dispositions themselves. So, the CDT has it that what Jones means by, say, 'green' is everything that Jones is disposed to apply 'green' (and here this includes everything Jones *in fact* applies 'green') to.

Now, Kripke raises a number of objections to this crude theory. The first of these is, of course, the putative normativity of meaning that we discussed at length in the previous chapter. That is, knowing how I am disposed to act will tell me merely what I would do with an expression in such-and-such circumstances, whereas what is required is a fact that tells me what I *should* do in those circumstances. Even though we have argued that the stronger normativity theses are false, there is still an objection to the CDT in the weakest reading of the normativity requirement. For even anti-normativists accept that meaningful expressions have correctness-conditions. If the CDT is true, then Jones means by 'green' *whatever* properties he applies 'green' to – even in error, under sub-optimal conditions, and so on. Thus if Jones applies 'green' to something blue in poor lighting, then on the CDT he means *green or blue in poor lighting* by 'green'. As his misapplications increase, the problem only multiplies. This is, of course, the disjunction problem with which we started, and it seems that the CDT has no resources for solving it.

While the disjunction problem is fatal to the CDT, this is not the only difficulty that Kripke raises. He also argues that one's dispositions are *finite*, whereas the potential application of an expression is not (1982 [2003]: 26). Thus, for example, the range of the addition function is infinite, whereas a finite brain can only cope with numbers that are so big. As we will see, the Sophisticated

Dispositional Theory (SDT) attempts to solve *both* the finitude problem *and* the disjunction problem via the use of *ceteris paribus* clauses. However, it sadly incurs several severe difficulties of its own.

The CDT clearly falls prey to the objection that it cannot solve the disjunction problem as it equates what is correct for me to do with what I (am disposed to) do. To avoid this problem, any theory of representation must make room for *mis*representation. There are two moves that one could make at this point. First, one could attempt to distinguish the things we are *disposed* to do from other things we do that are out of character. Thus, for example, we could argue that Jones is *disposed* to apply 'green' only to green things, and he just *happens* to sometimes apply 'green' to non-green ones.

Sadly, this won't, by itself, be enough. First, we would certainly owe the sceptic an account of what having a disposition consists in. Otherwise it is open to the sceptic to argue that making 'mistakes' – applying, say, 'green' to the non-green things – is something we are indeed disposed to do. Perhaps this difficulty is not insurmountable (see the discussion of Dispositional Realism below.) However, secondly, and more worrying, we may be disposed to make *systematic* mistakes (Kripke [1982] 2003: 30). Consider Jones who systematically forgets to carry when he performs additions. We may be forced to say that Jones is *disposed* to forget to carry. However, it is easy to envisage cases where we still *want* to say that Jones means *plus* by '+' all the same.

Thus even accounting for the difference between things we are disposed to do and things we just *do*, we still do not want to view *all* of our dispositions to use words as *meaning-constituting*: some of our dispositions we would rightly treat as *error-producing* (Boghossian 1989a). This brings us to the second option for the dispositionalist: to give some way of demarcating the meaning-constituting dispositions from the error-producing ones such as to give the intuitively correct extensions for our expressions. The CDT fails precisely because it allows into the extension of a term anything that I apply the term to. However, the sophisticated dispositionalist insists, we can do better than the crude theory. We can refine the theory first by challenging the assumption that *all* of our dispositions to use words are meaning-constituting, and second, by giving some account of what the

meaning-constituting dispositions are (or why the meaning-constituting dispositions *are* meaning-constituting).

So we are after an analysis of what it is to mean something by an expression of something like the following form (where F is a meaning-constituting fact):

Jones means *horse* by 'horse' iff F is true of Jones

So far, so good. Now here's the issue: the sceptic wants to know *on what grounds* I view F as the meaning constituting fact, and not, say G. In terms of our example, he wants to know why I insist on treating the disposition to apply 'horse' to horses as meaning-constituting, and, say, the disposition to apply 'horse' to cows (presuming I do systematically make such a mistake) as error-producing. The charge is one of circularity: the sceptic doubts that it is possible to give an account of why the former disposition is meaning-constituting *without* assuming my meaning (Kripke [1982] 2003: 28). This adds an important constraint to any response to the sceptic: it must be couched in non-semantic, non-intentional terms. Thus no assumption of what it is that any speaker means is allowed.

So here is the challenge for the sophisticated dispositionalist:

(EC)   In order to successfully respond to the sceptic, the dispositionalist
         must give a non-semantic, non-intentional account of what it is for
         a disposition to be meaning-constituting *without* making any
         assumptions as to what a speaker means by their words and thereby
         violating the terms of discussion with the sceptic.

4.3 Attempts to Solve the Problem of Error

4.3.1 Sophisticated Dispositionalism and Ceteris Paribus Idealisations

I may have several dispositions with regards to any of my expressions. Some of these will be dispositions to use the expressions correctly, others to make mistakes. The sophisticated dispositionalist hopes to make sense of the distinction

by adding a *ceteris paribus* clause to the crude dispositional account. The crude dispositional account has it that:

> CDT: Jones means *horse* by "horse" ↔ Jones is disposed to apply 'horse' only to horses.

The sophisticated dispositionalist maintains rather that:

> SDT: Jones means *horse* by 'horse' ↔ *Ceteris paribus*, Jones is disposed to apply 'horse' only to horses.

The *ceteris paribus* clause performs two functions. First, it allows the dispositional account to meet Kripke's objection from the finitude of one's dispositions. Second, it separates one's meaning-constituting dispositions from the error-producing ones: the meaning-constituting dispositions are those that are manifest under the *ceteris paribus* idealisation.

Now, the difficulty for the sophisticated dispositionalist is to cash out the *ceteris paribus* clause without either making recourse to semantic/intentional notions (to avoid begging the sceptic's question) or tacitly *assuming* what a speaker means (to avoid arguing in a circle). Following Fodor, we may think of the sophisticated dispositionalist as positing two types of situation: Type-1 and Type-2. Type-1 situations are the idealised conditions which the sophisticated dispositionalist desires to encapsulate with the *ceteris paribus* clause. In these terms, what the sophisticated dispositionalist needs is a naturalised (i.e. non-semantic, non-intentional) account of Type 1 situations.

Kripke raises two arguments against the possibility of providing such an account. The first is that there is no way we can satisfactorily flesh out the *ceteris paribus* clause. With regards to the finitude objection, Kripke writes:

> How should we flesh out the *ceteris paribus* clause? Perhaps as something like: if my brain had been stuffed with sufficient extra matter to grasp large enough numbers, and if it were given enough capacity to perform such a large addition, and if my life (in a healthy state) were prolonged enough, then given an addition problem involving two large numbers, *m* and *n*, I would respond with their sum…But how can I have any confidence of this? How in the world can I tell what

would happen if my brain were stuffed with extra brain matter, or if my life were prolonged by some magical elixir?...We have no idea what the results of such experiments would be. They might lead me to go insane, even to behave according to a quus-like rule (Kripke [1982] 2003: 27).

Kripke's second argument is really a challenge concerning circularity. Recall that, according to the sophisticated dispositionalist, ascriptions of meaning can be analysed as follows (where C means that the *ceteris paribus* idealisation holds):

Jones means *plus* by '+' ↔ (C → Jones is disposed to respond with sums to addition queries)

Assume that the left hand side of the above biconditional is true. We now need an account of C such that the whole biconditional comes out true. However, on pain of circularity, this must be couched in non-semantic non-intentional terms. Kripke's specific target here is any specification of C that makes reference to the fulfilment of intentions: in the context of the present challenge, 'if I were to fulfil my intentions, I would respond with the sum' cannot be favoured over 'if I were to fulfil my intentions, I would respond with the quum' without circularity, as it is assumed in the specification of C which function is meant (Kripke [1982] 2003: 28).

In his 'Theory of Content', Fodor has responded to the first argument. [2] He writes:

Apparently Kripke assumes that we can't have reason to accept that a generalisation defined for idealised conditions is lawful unless we can specify the counterfactuals which would be true if the idealised conditions were to obtain. It is, however, hard to see why one should take this methodology seriously. For example: God only knows what would happen if molecules and containers actually met the conditions specified by the ideal gas laws…for all I know…, the world would come to an end. After all, the satisfaction of these conditions is, presumably, *physically impossible* and who knows what would happen in physically impossible worlds? (1990: 94).

The idea is this: in the sciences it is, in fact, common practice to appeal to *ceteris paribus* clauses and idealisations where no-one has any idea what would happen (i.e. what counterfactuals would be true) if the idealisations obtained. And if this

---

[2] For criticism of Fodor's view of the role of idealisations in projects of naturalising semantics see Kusch 2006. For a defence of Fodor see Cheng 2009.

practice is legitimate in the sciences, why should it not be in the philosophy of language and mind?

> [Just as] it's not required, in order that the ideal gas laws should be in scientific good repute, that we know anything like all of what would happen if there were ideal gases…similarly, if there are psychological laws that idealize to unbounded working memory, it is not required in order for *them* to be in good scientific repute that we know all of what would happen if working memory really were unbounded. All we need to know is that, if we did have unbounded memory, then, *ceteris paribus*, we would be able to compute the value of *m* + *n* for arbitrary *m* and *n* (Fodor, 1990: 95).

In considering how successful Fodor's response is, we have to be careful to distinguish two things. On the one hand, there are the counterfactuals that would obtain were a certain idealisation to hold. Fodor is, I think, correct to say that we needn't be able to specify all of these for an idealisation to be in good repute. However, on the other hand, we have the idealisation itself: what does the idealisation consist in?

With regards to this second question, it seems that the idealisations involved in the sophisticated dispositional response to the sceptic are *not* clearly analogous to those used in the sciences. Consideration of the ideal gas laws will help to bring out this difference quite clearly.[3] One of the ideal gas laws states the relationship between pressure (*P*), volume (*V*), temperature in degrees Kelvin (*T*), number of moles (*n*) and a constant (*R*) as follows:

Ceteris paribus, $PV = nRT$

Now, the important thing to note for our purposes is that we can specify *precisely* what the *ceteris paribus* clause involves. It essentially subsumes five idealisations as follows:

(i) Molecules are perfectly elastic;

(ii) Containers are infinitely impermeable;

(iii) Molecules have zero volume;

(iv) There are no attractive forces between gas molecules or between gas molecules and the side of the container;

(v) Heat is the only source of energy present in a gas.

---

[3] In my presentation of the ideal gas laws, I follow Kusch 2006: 100-1.

Now, here is the first problem: can we provide anything like this fleshing out of the *ceteris paribus* clause with regards to sophisticated dispositionalism? It is certainly not obvious that we can. (In a moment we will consider an argument from Boghossian to the effect that it is impossible *in principle* for the sophisticated dispositionalist to adequately spell out the clause.)

The second problem with Fodor's response to Kripke is that it seems to be somewhat beside the point. The thrust of Kripke's first argument is not: how can we tell *which* counterfactuals would hold under a certain idealisation; but rather: how can we be sure that, under a certain idealisation, the *relevant* counterfactual would hold? In other words, how can we be sure that, if we idealise my mental powers in such-and-such a way, I will respond to addition queries with sums in the realm of enormous numbers? We may contrast the question of how we can be sure that, were the five-part *ceteris paribus* clause to hold, that *PV* would equal *nRT*. In this case, it seems that we *can* be fairly sure, due to experiment, observation, the experimental approximation of the idealisation, and subsequent attempts to move away from the idealisation in order to attain greater predictive accuracy (Kusch 2006: 101). Nothing like this seems to be the case with regards to the *ceteris paribus* clause in sophisticated dispositionalism (Ibid: 102). Thus it seems that Kripke's first argument escapes Fodor's objections.

Let us move on to some further problems for sophisticated dispositionalism. Boghossian raises arguments against the sophisticated dispositionalist which stem from concerns about meaning holism. As Miller has noted, Boghossian is not in the business of trying to show that some *specific* dispositional proposal is incorrect, but rather that 'there are principled reasons for thinking that reductive dispositionalism *could not* be true' (Miller 2003b: 73). The basic point is this:

> Under normal circumstances, belief fixation is typically mediated by background theory—what contents a thinker is prepared to judge will depend upon what other contents he is prepared to judge. And this dependence is, again typically, arbitrarily robust: just about any stimulus can cause just about any belief, given a suitably mediating set of background assumptions (Boghossian, 1989a: 539).

Thus, for example, I could judge that the cow in front of me is a horse if I believe that cows are horses, or that all the horses in the area have been cunningly disguised as cows, or that whatever my friend Phil tells me is true (given that I believe that Phil has told me that the animal in front of me is a really a horse), and so on and on. How does this cause a difficulty for the sophisticated dispositionalist? Well, recall that she is in the business of specifying a set of idealised conditions under which I will be disposed to apply my expressions only to things in their extensions. However (using magpies instead of horses as an example):

> The observation that beliefs are fixed holistically implies that a thinker will be disposed to think *Lo, a magpie* in respect of an indefinite number of *non-magpies,* provided only that the appropriate background beliefs are present. Specifying an optimality condition for 'magpie', therefore, will involve, at a minimum, specifying a situation characterized by the absence of *all* the beliefs which could potentially mediate the transition from non-magpies to *magpie* beliefs. Since, however, there looks to be a potential infinity of such mediating background clusters of belief, a non-semantically, non-intentionally specified optimality situation is a non-semantically, non-intentionally specified situation in which it is guaranteed that none of this potential infinity of background clusters of belief is present (Ibid: 540).

How could the sophisticated dispositionalist hope to specify such a situation? Well, they would need to provide 'a set of necessary and sufficient conditions for being a belief with a certain content' but this was 'precisely what a dispositional theory was supposed to provide'! In other words, for a dispositional theory of meaning to work, we require a pre-existing theory of meaning; 'but if we had *that*' Boghossian writes, 'we would not need a dispositional theory!' (Ibid)

In his 'Naturalising Content', Boghossian makes essentially the same point in a different way. What the sophisticated dispositionalist needs to show is that, for any given belief B, none of the 'background beliefs' which would 'frustrate the connection between symbol tokenings and their referents' (magpie beliefs and magpies, as it may be) is present. Otherwise the idealised conditions they specify could be consistent with the presence of such background beliefs. Now, as Boghossian puts it:

> Putting this worry to rest requires showing that [the optimal situation cited by the sophisticated dispositionalist] is enough to ensure the truth of the following open conjunction:
>
> (*) $-Bel_1$ & $-Bel_2$ & $-Bel_3$ & …

where the Bel$_i$ stand for the various clusters of background beliefs which could potentially frustrate the connection between being a magpie and the tokenings of an expression which refers to it (Boghossian 1991: 80).

So the sophisticated dispositionalist must concede that (*) forms part of the idealised conditions under which a token of their mental representation 'magpie' will covary with its referent. However, wonders Boghossian, 'could [such conditions] ever be recognised as sufficient for the truth of (*)?' (Ibid) Because (*) is not 'finitely stable', he writes, 'there is no finite way to state what beliefs the [sophisticated dispositionalist] must rule out'. Thus we are unable to *certify* that the conditions specified would amount to a Type 1 situation (Ibid: 80-81). But, surely, if our theory is to be successful, it *must* certify such conditions.

If, as I think they are, these considerations against theories that make recourse to Type 1 situations are compelling, then we need to try and rescue the idea that meaning can be reduced to 'naturalistic' (non-semantic, non-intentional) facts by trying to find a different type of theory. We will look at some examples of these in the next two sections. I should say at the outset, however, that I think the first three theories (Blackburn's pragmatist story, Martin and Heil's realism about dispositions, and Millikan/Papineau's teleological accounts) fail, and that the fourth (Fodor's Asymmetric Dependency Theory) is not entirely satisfactory as a theory of intentionality, even if it is able to escape some of the criticisms that have been levelled at it.

4.3.2 More Dispositions

Simon Blackburn (1984) has advocated a dispositional response to the sceptic which consists essentially in considering more dispositions when determining what one means than the sceptic does. Take Kripke's example of plus and quus. True enough, I may have the disposition to add numbers. I may also have the disposition to quadd them, or to return answers to addition problems that do not give the sum of the numbers. However, says Blackburn, among my dispositions is also the disposition to *retain* certain answers after an independent process of checking. So if I am disposed to retain *only* those answers which do in fact accord with the addition function, then I mean *plus* by 'plus'. Blackburn writes:

> But at least it is true that [we] can have, in addition to dispositions to give answers, dispositions to withdraw them and substitute others. And it is possible that putting the errant disposition [say, to return answers that are not the sum of an addition query] into a context of general dispositions of this sort supplies the criterion for which function is meant. The equation would be: By '+' I mean that function Φ that accords with my extended dispositions. An answer $z = \Phi(x, y)$ accords with my extended dispositions if and only if (i) it is the answer I am disposed to give and retain after investigation, or (ii) it is the answer I would accept if I repeated a number of times procedures I am disposed to use, this being independent of whether I am disposed to repeat those procedures that number of times (Blackburn 1984: 290).

As Hattiangadi points out, conditions (i) and (ii) constitute Blackburn's response to the problem of error. Those responses that I am disposed to give but that do *not* meet condition (i) are mistakes (Hattiangadi 2007: 117). Condition (ii) allows for the possibility of someone who gets sums wrong, and yet still means *plus* by 'plus' *even though* they don't actually go through the process of checking enough times to correct their answer.

We may say in favour of Blackburn's account that it is able to solve the problem of *systematic* mistakes which was raised as an objection to the sophisticated dispositionalist. Recall that, even *ceteris paribus*, someone who systematically makes mistakes will return intuitively incorrect answers to addition problems. However, such a person may have a disposition to retract those answers after investigation – they may even recognise that they are hopelessly incompetent at completing sums properly, and thus be disposed to put their independent error checking process into effect. However, there still seems to be a problem. This is just that even after a huge amount of checking of a difficult addition problem where I have made a mistake I may retain my mistaken answer. Perhaps sometimes I will retain such answers, and sometimes I won't. But, on the current account, this leaves it 'indeterminate [as to] whether I mean *addition* or some other bizarre function by "plus"' (Hattiangadi 2007: 117). Perhaps worse, someone might be disposed to revise their *correct* applications of the expression – that is, to give correct responses to addition queries, but then make errors in the checking procedure. Would such a person not mean *plus* by 'plus'?

Blackburn has a response to this worry. The general idea is that we have dispositions to act that are *indirectly* connected to the meanings of our

expressions, and that if we take *these* into account, we will be able to specify what is meant by those expressions (Hattiangadi, Ibid). Thus if Jones really does attach bizarre meanings to his expressions, this will come out *before* we get to an occasion where Jones makes an application that does not accord with the usual meaning of the expression. Blackburn writes:

> The concealed bent-rule follower is often thought of as though nothing about him is different until the occasion of bent application arises. But this is wrong. Someone who has genuinely misunderstood a functor is different, and the difference can be displayed quite apart from occasions of application (1984: 291).

What does this mean exactly? Well suppose Jones really *does* mean *quus* by 'plus'. And suppose he works as a restaurant manager. He takes reservations for 68 people in the morning, and 57 more in the afternoon. For dinner that evening, he sets up tables – for 5 people! If Jones were to insist that what he did was correct, we may well view him as a *bona fide* quadder. Now contrast Jones with Smith, who does the same thing, but is an adder who has made a bizarre mistake. Smith may well revise his judgement that '68+57=5' after the horrendous mess that results at his restaurant! Such practical projects as setting dinner tables for the correct number of guests 'provide us with a measure of success or failure'. The point is that I will revise my incorrect and *not* revise my correct responses in the light of these broader measures of success (Hattiangadi 2007: 118).

   Sadly, however, even this revision cannot completely avoid the sceptical problem. Suppose that, due to unforeseen circumstances, 120 of Smith's guests fail to show up for dinner. Then Smith's mistaken judgement still leads to a successful evening. Thus, even by the external standard of correctness, Smith is not led to revise his judgement. Suppose that Smith means *plus* by 'plus' and that he sets up tables for 125 people – but 130 turn up (5 unannounced). It is possible that Smith could blame the failure of his project to set up dinner plates for enough people on his addition. That is, he could think that he had made a mistake, and alter his belief that '57 + 68 = 125'. So the success/failure of a project will not guarantee that 'correct' dispositions won't be changed, or that 'incorrect' dispositions will be. So even taking these extended dispositions into account, we cannot rule out the indeterminacy that the sceptical argument exploits. Of course, this example takes just an individual case: it may be that if we take into account more of Smith's corrective behaviours, he will modify his dispositions again.

There still remain two problems, however. First, there is no guarantee that circumstances will eventually lead Smith to correct his behaviours in the 'right' way. Second, what are we to say about Smith when he has 'corrected' his use of addition such that '57+68' equals 130 *before* the further corrective behaviours kick in as a result of future pragmatic projects? Is he an adder all along? Does he fluctuate between being an adder and a quadder? It seems that we cannot make a firm decision about this. Perhaps we could say, were Smith to engage in such and such a project, then *ceteris paribus* he would correct his behaviour thus and so. But then the response lapses back into sophisticated dispositionalism, and we have already surveyed the difficulties here. The chief worry would be that we could not say with confidence that Smith would revise his responses in the correct way without *presuming* what he means by '+'. So we cannot escape the problem of error this way.

4.3.3 Dispositional Realism

Another attempt to solve the problem of error, and to thereby rescue the dispositional theory from the sceptic, has been to insist that Kripke gives the wrong account of dispositions. This is the response advocated by Marin and Heil. They write:

> [The problem of error arises for the dispositionalist] only given a particular perspective of dipositionality. If we imagine that a disposition is analysable in terms of counterfactual or subjunctive conditionals holding true of agents or objects, then perhaps we should be unable to distinguish [meaning-constituting from error producing dispositions]. We submit, however, that this is not the only way to understand dispositions, and it is not the right way (Martin and Heil 1998: 289).

Martin and Heil advocate instead a *realist* conception of dispositions. According to this conception, dispositions cannot be analysed in terms of counterfactual conditionals. This is because we can correctly ascribe a dispositional property to something and yet that thing may fail to *manifest* its dispositions due to interfering or blocking factors (Ibid: 290). They distinguish cases where dispositions are *lost* and cases where dispositions are either *blocked* by the presence of inhibiting factors, or rendered unable to manifest by the absence of 'reciprocal' factors (Ibid: 290-291).

Given all of this, the following answer to the sceptic may be advanced: the fact that Jones means *plus* by '+' and not *quus* is that Jones has a disposition regarding the '+' sign to return sums to addition queries. The fact that Jones is not able to *manifest* this disposition in certain circumstances (i.e. when he is inebriated, or when the numbers are simply too enormous to manage) is beside the point: this disposition he has, and *plus* is his meaning.

What, then, does having a disposition consist in on a realist view? Martin and Heil's proposal, insofar as I understand it, is that dispositions are *aspects* of properties. On their view 'every property has both a qualitative...and a dispositional side or aspect', these are 'correlative and inseparable'; and the notion that any property is purely dispositional or purely qualitative is a 'myth'. Thus 'the dispositionality of an intrinsic property is as basic and irreducible as is its qualitative character. Neither is more basic, neither "supervenes" on the other—insofar as supervenience implies a "direction" of dependence' (Martin & Heil 1998: 289-90).

Notice that Martin and Heil's response offers a way of making the distinction between those things we have a disposition to do and those that we do more or less accidentally. It simply isn't the case, on such a view, that the fact that I would do action *A* in circumstances *C* amounts to my having a disposition to perform *A* in *C*. Rather, I would need to have some property whose dispositional aspect would be my *A*-ing, and whose qualitative aspect could perhaps be detected. Compare the case of salt. Salt has the property of having a certain crystalline structure whose dispositional aspect manifests in salt dissolving in water. Or, to put it the other way around, salt's disposition to dissolve in water is the dispositional aspect of its property whose qualitative aspect is salt's particular crystalline structure.

Now, Hattiangadi raises a problem for this line of response. She argues that it cannot distinguish between meaning-constituting dispositions and error-producing ones. So it isn't going to make the problem of error go away. She gives an example like the following one (Hattiangadi 2007: 108-109):

Suppose I have two dispositions regarding the '+' sign:

D1: I am disposed to return answers to queries of the form *m + n = ?* which are the sum of *m* and *n*.

D2: I am disposed to return answers to queries of the form *m + n = ?* which are *not* the sum of *m* and *n*.

Now consider the following scenario:

D1 is my meaning-constituting disposition, D2 my error-producing one.

What is to stop the sceptic from claiming that:

D2 is my meaning-constituting disposition, and D1 is my error-producing one?

Hattiangadi argues that there is nothing in the dispositional realist's story which allows us to account for this. However, this may be too quick. Martin and Heil may well respond that I have D1 in lieu of the fact that I have some property P where D1 is the dispositional aspect of P. Yet they may deny that I have some property P* such that D2 is the dispositional aspect of P*. Thus they may argue that this *isn't* a case of my having two dispositions.

However, at this point it is open to the sceptic to ask *on what grounds* do we attribute the meaning-constituting (realist) disposition D1 to a speaker? Presumably not on the basis of having discovered what P is, but rather on the basis of such instances when I do, in fact, respond with the sum of *m* and *n*. But then, given that there will almost certainly be cases when I do not respond with their sum, so the sceptic would be free to attribute D2 to me as well, and raise the challenge that Hattiangadi runs against Martin and Heil.

This is not the only option, however. Perhaps future science will one day identify the qualitative aspects of P and P*. This would certainly provide the dispositional realist with the grounds to argue that some things that we do we are not disposed to do. If one determines that a speaker does not have property P* by showing that the qualitative aspect of P* is absent in that speaker, then such a speaker *ipso facto does not have D2*. Sadly, until such time as the qualitative aspects of such properties are identified, we seem to be at a stand off. The realist could argue that the sceptic's positing of D2 is glib and without basis. However, the sceptic could point out that the *realist's* positing of D1 is equally glib and

without basis! We *cannot tell* which of a speaker's linguistic acts actually consist in the manifestation of her dispositions and which do not *simply by looking at such acts*. In other words, from the point of view of the linguistic evidence, D1 and D2 are each equally well (or, perhaps, equally poorly) supported!

Perhaps, then, we should regard the jury as being out on the realist picture pending further evidence – evidence of the qualitative aspects of those properties whose dispositional aspects manifest as certain kinds of linguistic behaviour. However, it seems to me that the realist response faces another difficulty: it seems that it *cannot* deal with the cases of systematic mistakes. Suppose Jones *really does* have the realist disposition to make systematic errors in calculation, and thus he *fails* to return sums to addition queries. In such a case, it seems that even the dispositional realist would have to concede that Jones does not mean *addition* by 'plus' as meaning *addition* by 'plus' *consists in* having the realist disposition to return sums to addition queries. However, this seems wrong: someone can be utterly hopeless at *doing* addition, yet still grasp the concept of addition. That is, one can mean *plus* by 'plus', and stink at doing sums! So it seems that dispositional realism doesn't offer us a compelling solution to the problem of error.

4.3.4 Cause Theories and Effect Theories

I have certain dispositions to use the word 'horse' or, perhaps better for present purposes, to token the thought 'horse'. The problem is to account for what makes it the case that my utterances of 'horse' mean *horse* and that my tokenings of the thought 'horse' have the content *horse*. We now move on to look at two opposing approaches to this problem which I shall characterise as Cause Theories and Effect Theories. According to Cause Theories, what makes it the case that my tokenings of 'horse' have the content *horse* is what would, perhaps *ceteris paribus, cause* such thoughts to be tokened. Because we can formulate such views in terms of counter-factual conditionals, we may view them as fitting more or less into the family of dispositional views. According to Effect Theories, by contrast, what fixes the content of my thoughts is the behaviour they generate (Papineau 2006: 180). Let us turn to these first.

According to one influential theory of this latter sort, success semantics, such behaviour is to be made sense of in terms of the satisfaction of my desires. As Papineau succinctly puts it, according to success semantics, the 'truth-condition of any belief is that circumstance which will ensure the satisfaction of whichever desire combines with the belief to cause action' (Ibid). Let us illustrate this with an example. Suppose I desire some beer, and that I believe that there is beer in the fridge. Together these mental states may cause me to go to the fridge to get some beer, thus satisfying my desire for beer. The circumstance that ensured my desire was satisfied was just that there was beer in the fridge. Thus *there being beer in the fridge* is the truth-condition of my belief that there was beer in the fridge. So goes the success semanticist's story about the *content* of my belief. They also have a story about what *constitutes* my having such a belief: what makes it the case that I, say, believe that *p* 'is that [I] behave in a way that will satisfy [my] desires *if p*' (Ibid). In other words, what makes it the case that I believe that there is beer in the fridge is that I go to the fridge when I want a beer. Just to tidy all of this up, we have it that:

> *Content*
> A belief *B* has the truth-condition *T* just in case the obtaining of *T* will ensure that any desires which cause action in conjunction with *B* will be satisfied

> *Constitution*
> (Jones behaves in ways that will satisfy his desires if the truth condition of a belief *B* obtains) → (Jones has belief *B*)

Now one place in which success semantics may be thought to be very appealing is in its ability to deal with *mis*representation, and thus the problem of error. As Papineau explains, because beliefs are individuated with respect to their effects, as long as they *have* those effects, they will have the same truth-conditions, even if those truth-conditions do not obtain. For example, if I believe there's beer in the

fridge, and I desire beer, then, *ceteris paribus*[4], I will go to the fridge to get it. Thus the content of my belief is that *there's beer in the fridge* – even if there isn't any there.

However, as Papineau points out, in order to count as truly naturalistic (and thus avoid begging the sceptical question by assuming the intentional contents of mental states), success semantics also owes us an account of desire satisfaction which makes sense of the representational properties of desires in a naturalistic way (i.e. without presupposing representational/intentional properties.). To illustrate this problem with an example, we need a story about what makes it the case that my desire for chocolate is a desire about chocolate, and not, say, about ice cream which makes no reference to the content of my desire. The natural response here is to 'find some independent account of desire satisfaction' (Ibid: 181 cf. Whyte, 1991).[5]

Suppose I desire some chocolate. What makes this desire the desire that it is? In other words, what makes it the desire *for chocolate*? Here are a couple of suggestions. First, it may be that the content of my desire is determined by circumstances that would *extinguish* the desire – that is, which make the desire go away. Second, it may be that the satisfaction-conditions for my desire are those which make my desire go away and which 'make it more likely that the behaviour prompted by the desire will be repeated' next time I have the desire (Ibid). Let's take the first suggestion first. The immediate problem here is that many things may extinguish my desire for chocolate other than chocolate. A nice punch to the stomach (as Hattiangadi suggests), or some ice cream (having sated my sweet tooth, I may no longer desire chocolate), or, indeed, being anaesthetised. Surely the content of my desire for chocolate isn't a desire to be punched in the stomach, or to eat ice cream, or to start feeling physically sick, or be given a large shot of morphine…

---

[4] This clause needs to be there to eliminate cases where other factors are present and relevant to how I act. For instance, I may have *conflicting* desires: I desire beer, but I need to drive later, and I desire to *not* drive while mildly inebriated. My thanks to Bob Hale for pointing this out.

[5] The teleosemanticist claims to have a way to do this. We shall look at her theory in the next section. For now, let's have a look at a couple of unsuccessful attempts so as to get a better grip on exactly what the problem is.

What about the second suggestion? On this view what makes my desire for chocolate a desire for chocolate is that getting chocolate will make my desire go away *and* prompt me to get chocolate the next time I have the desire. There are two problems here. The first is that such a view 'fails to yield determinate satisfaction conditions' (Hattiangadi 2007: 125). Suppose I desire chocolate but decide to go out for a run before going to the shop to get some. After the run I find that the desire has gone away and, being a health-conscious sort of chap, the next time I desire some chocolate I am prompted to go for a run to make the desire go away and feel healthy. The trouble is, on our current account what I desired all along was to go for a run. The second problem is that such an account would seem to 'rule out the possibility of desire content when there is no reinforcement learning' (Papineau 2006: 181).

Perhaps we need to look elsewhere for an account of desire satisfaction, and I will go on in the next section to outline such an account. However, before doing so I wish to mention a further problem for success semantics raised by Brandom (1994). First recall the success semanticist's main thesis (the following is Brandom's formulation):

(SS)   A belief's truth condition is that condition which guarantees the fulfilment of any desire by the action which, combined with that desire, it would cause.

Remember also that this seemed to give us a neat account of misrepresentation: if I undertake the action, but was unsuccessful (I go to the fridge for beer but there is none), then my representation was false. Now, the problem Brandom foresees for this account is this: 'I can be thwarted as much by what I don't know as by what I falsely believe' (1994: 175).

Consider a case where I go to the fridge to get beer, but, unbeknownst to be, someone has glued the door shut, or moved the fridge, or encased it in iron, or whatever. At this point, it seems that the success semanticist faces a dilemma: either they must accept (SS) is violated (there being beer in the fridge does not *guarantee* my success in sating my desire for beer in these circumstances, so this

cannot be the truth-condition of my belief), *or* they have to assimilate ignorance to error (Brandom 1994: 175-6, cf. Whyte 1990; 1991).

Brandom neatly points out the problems with the second strategy. For one thing, it would seem that I would have to have a (potentially) infinite number of beliefs about the *absence* of frustrating factors (1994: 177). That is, I would have to believe that the fridge was not glued shut, that it was not encased in iron, and so on. This is far from plausible. What is more plausible, as Brandom points out, is that I *lack* the positive beliefs that the fridge is glued shut, etc. However, while we may infer from this, for example, that *I did not believe that the fridge was glued shut*, what the success semanticist requires is that *I believed that the fridge was not glued shut* (Brandom 1994: 176). Thus even if the success semanticist can give an account of desire satisfaction, the prospects for their account do not look good.

Teleosemantics

One option that might be explored is to opt for a teleological account of success. This would be to construe teleosemantics as essentially a continuation or development of success semantics. (It is arguable whether all teleosemanticists would see their semantics this way. See the discussion of Millikan below.) Here the notion of desire satisfaction is cashed out in terms of the notion of proper function. Proper function is then analysed along etiological Darwinian lines. Let us attend to the latter account first. The proper function of a biological feature is, according to one influential line of thought, 'determined by its evolutionary history' (Hattiangadi 2007: 126); thus its proper function is *what it was selected for*. Let us take the example of the kidneys. To say that the function of the kidneys is to filter blood is just to say that creatures whose kidney's were good at filtering blood out-competed (i.e. reproduced more successfully than) creatures whose kidneys were not so good at filtering blood. Another influential line of thought about the proper functions of biological features has it that it is not just evolutionary *history* that is important; rather, it is quite proper to view subsequent effects of a feature, which are unrelated to its *origins*, as proper functions of that feature, insofar as such effects 'explain why the trait of having [that feature] is preserved in subsequent generations' (Ibid: 127. re. Millikan 1990: 86). Take the example of birds' feathers. It has been suggested that feathers may have evolved

116

originally as an alternative method of insulation to, say, fur, or alternatively for display purposes, such as attracting a mate.[6] However, one would not wish to deny now that the proper function of the wing feathers of birds is to aid flight. Fortunately, we are able to say this, precisely because birds which are good at flying tend to out-compete their less good competitors. Thus it is quite defensible to view aiding flight as a proper function of wing feathers.

So, proper functions are those effects, original or subsequent, of biological features which are selected for by natural selection in the sense that those organisms in which the effect is more pronounced out-compete other organisms in which the effect is less pronounced. Let us now examine how functions are to replace desire satisfaction in the teleosemanticist's account of mental content. Recall that on a (non-teleological) success semantics account, we had it that:

(SS) The truth-condition (content) of a belief B is that condition which, if it obtains, will guarantee the satisfaction of any desire D which causes action in conjunction with B.

Things become a little more complicated on the teleological picture. First, note that there are three kinds of functions which we may wish to talk about in connection with representational mental states. First of all, there is the function of the biological mechanisms which *give rise to* the mental state (for example, the perceptual systems). Secondly, one might wish to talk about the functions *of the representational states themselves*. Finally, one might wish to talk about the functions of those biological mechanisms which *consume* the representations.

Millikan (1986) has it that *all three* have proper functions (We can diagram Millikan's view as follows (from Papineau 2006: 182):

---

[6] See, for instance, Carl Zimmer's article 'The Evolution of Feathers' in the National Geographic at http://ngm.nationalgeographic.com/2011/02/feathers/zimmer-text/1.

```
┌─────────────────────────────────────────────────────────────┐
│         Representation-Consuming Mechanisms (RCMs)          │
│ Function: To use representations in order to direct behaviour│
│                     towards some end.                        │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│                       Representations                        │
│ Function: To enable the RCMs to fulfil their function(s) by  │
│             gearing behaviour to circumstances.              │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│        Representation-Producing Mechanisms (RPMs)           │
│ Function: To produce representations which are consumed by   │
│          representation-consuming mechanisms.                │
└─────────────────────────────────────────────────────────────┘
```
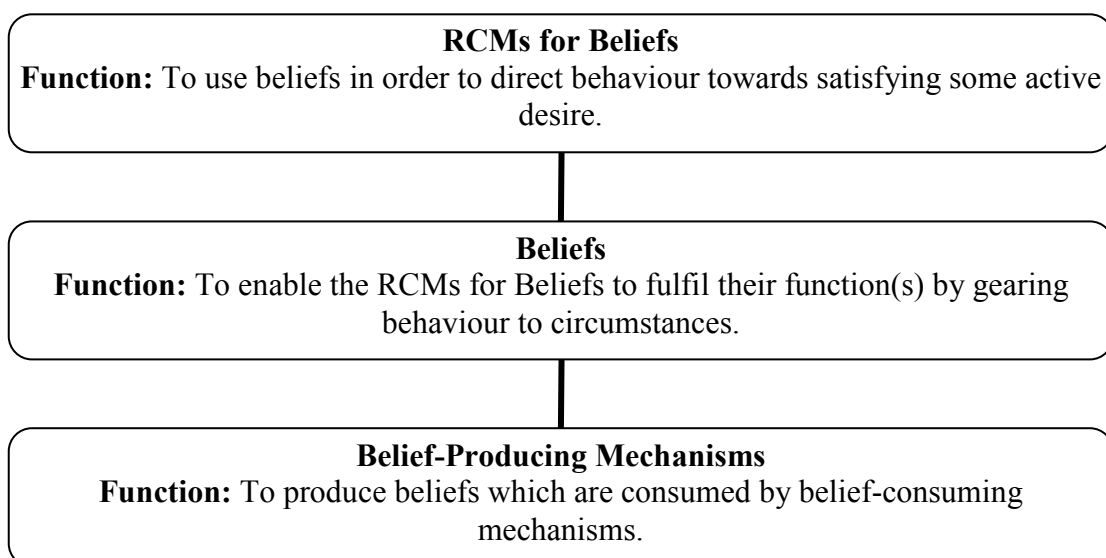
Now we have said that the teleosemanticist wants to replace the notion of desire satisfaction with the notion of biological function. We are now in a position to provide an accurate answer to the question: exactly *which* functions are supposed to replace the desires? Papineau writes:

> Millikan's version [of teleosemantics] coincides with the version that builds on success semantics if we equate the consumer mechanism for a *belief* with the decision-making process that uses that belief to select behaviour that will satisfy currently active *desires* (Papineau 2006: 182).

So, specifically, we have it that:

```
┌─────────────────────────────────────────────────────────────┐
│                       RCMs for Beliefs                       │
│ Function: To use beliefs in order to direct behaviour towards│
│              satisfying some active desire.                  │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│                          Beliefs                             │
│ Function: To enable the RCMs for Beliefs to fulfil their     │
│    function(s) by gearing behaviour to circumstances.        │
└─────────────────────────────────────────────────────────────┘
                              │
┌─────────────────────────────────────────────────────────────┐
│                 Belief-Producing Mechanisms                  │
│ Function: To produce beliefs which are consumed by belief-   │
│                  consuming mechanisms.                        │
└─────────────────────────────────────────────────────────────┘
```
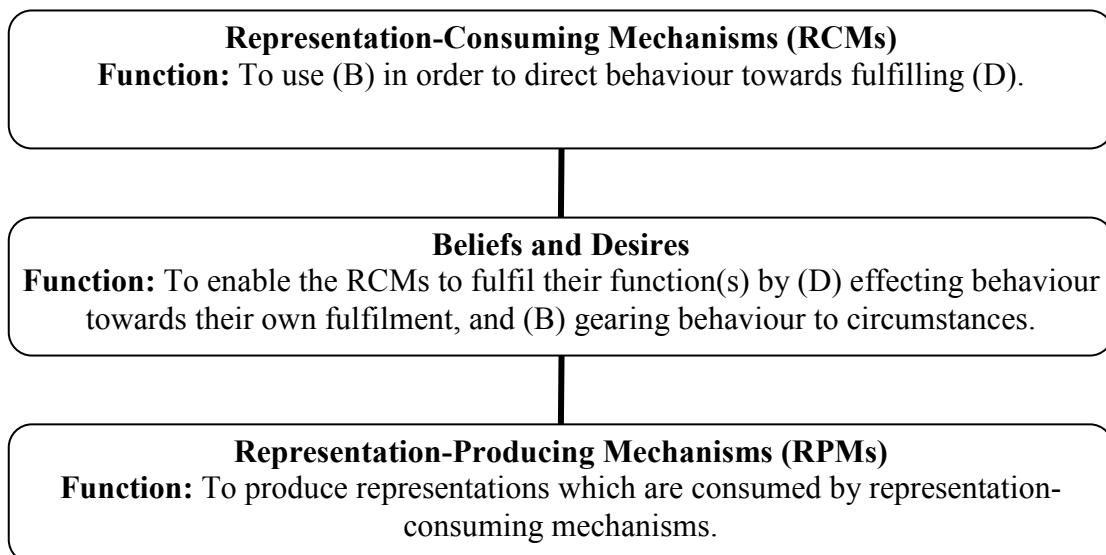
The trouble with this is that desires have entered the picture again. The thought was that biological functions were going to *take over* the work done by desires.

Accordingly, one way worry at this point that the teleosemanticist has not really analysed desires away at all. However, desires can be given a teleosemantic analysis as follows: the proper function of a desire is to be explained in the effects the desire was selected for (Papineau 2006: 181). According to Millikan, what desires effect is action in an organism to fulfil the desire (Millikan 1986: 63). So, the proper function of a desire is to effect action which will result in its own fulfilment. This would seem to imply that we can look to the *action* to figure out what the content of the desire was. (More on this in §4.3.5.) To give an example of the function of a desire, the desire for food may well, in conjunction with an appropriate belief, cause an organism to go and get some food. The desire would have had to have been selected for, on this account, precisely because it caused such action in historic organisms of that type.

Let's take stock. According to the teleosemanticist, the truth-conditions for beliefs and the satisfaction-conditions for desires are to be fixed in terms of their biological functions. Since, in Papineau's words, 'biological functions are in the first instance *always a matter of effects*' (2006: 182, my emphasis), we need to look at the effects of beliefs/desires in order to work out what their truth-conditions may be. The effect of a belief is to enable the RCMs for beliefs to direct the behaviour of an organism towards the fulfilment of some need/desire. The effect of a desire is to effect action towards its own fulfilment (presumably this can only be achieved in conjunction with an appropriate belief).

Let's diagram this once more for clarity, and then do it again with examples:

```
┌─────────────────────────────────────────────────────────────┐
│      Representation-Consuming Mechanisms (RCMs)              │
│  Function: To use (B) in order to direct behaviour towards   │
│                    fulfilling (D).                          │
└─────────────────────────────────────────────────────────────┘
                            │
┌─────────────────────────────────────────────────────────────┐
│                   Beliefs and Desires                       │
│  Function: To enable the RCMs to fulfil their function(s)    │
│  by (D) effecting behaviour towards their own fulfilment,    │
│  and (B) gearing behaviour to circumstances.                │
└─────────────────────────────────────────────────────────────┘
                            │
┌─────────────────────────────────────────────────────────────┐
│      Representation-Producing Mechanisms (RPMs)             │
│  Function: To produce representations which are consumed     │
│  by representation-consuming mechanisms.                    │
└─────────────────────────────────────────────────────────────┘
```

Now let's run over that again with an example. Suppose we are challenged to account for why my belief that *there's beer in the fridge* has the truth-condition that there's beef in the fridge. Well, according to the teleosemanticist, if has this truth-condition iff the effect of my belief is to gear my behaviour (via my RCMs) to going to the fridge when I desire a beer. What would make my desire for beer the desire *for beer*? Well, just that my desire would *cause* me (via my RCMs, and provided I has some relevant beliefs) to go out and get beer. Specifically, provided I had the belief in question, my desire would cause me (via my RCMs) to go the fridge and get a beer.

Finally, to complete the Darwinian story, we need to say a little bit about concept acquisition. Now, unless one wants to adopt a kind of *extreme* nativism, one might not want to say that concepts are the kind of things that can be selected for, where this is understood as a matter of genetic inheritance. In other words, concepts just don't seem to be the kind of thing that one can inherit from one's parents (unlike other traits such as hearts, kidneys and eye colour). Rather, concepts have to be *learned*. However, says Millikan, an *analogous* process of selection to the one which operates on inherited biological traits acts on learned things such as concepts. In other words, learning is a kind of artificial selection which operates in an analogous way to natural selection. Thus the Darwinian account of function – or something like it – can be extended to learned things like concepts. So, on our teleological story, we have it that:

(SS) The truth-condition TC of a belief B is that condition which 1. B represented to some organism O so that O was able to gear its behaviour (via its RCMs) to fulfilling its ends, and 2. B was selected for *because* it represented that condition. The satisfaction-condition SC of a desire D is that condition which 1. D represented to some organism O and 2. D was selected for *because* it caused historical organisms of that type, via their RCMs, to bring SC about.

Teleosemantics clearly has the edge over success semantics in that it gives some kind of account of desire satisfaction which does not take the content of the desire for granted. To go back to the example I used as a problem for success semantics: what makes it the case that my desire for chocolate is the desire *for chocolate* and not, say, *for ice cream* is that the desire was selected for because it caused me to go out and fulfil my need for chocolate, not ice cream.

In the next section, I will consider Fodor's critique of teleosemantics. Before doing so, however, I wish to register a couple of worries about the teleosemanticist's account of the content of desires. First, as I mentioned above, Millikan's account suggests that we can look to action to determine what a desire was a desire for. However, this may not deliver contents that are fine grained enough. Suppose I have a desire for pears, but there are none in the fruit bowl, so I eat an apple instead. Thus looking at my action, one would conclude that my desire was a desire for an apple, not a pear. Secondly, Millikan's proposal may well fall foul of the kinds of considerations concerning the extinguishing of desires that I raised above. In other words, if I have a desire for chocolate, and the function of the desire is to effect its own fulfilment, but this can be achieved by *extinguishing* the desire, then one might expect the desire for chocolate to lead one to seek out morphine, or punches to the stomach, or whatever.

4.3.5 Fodor's Critique of Teleosemantics

According to the teleosemanticist, then, the contents of my beliefs are to be fixed by their functions and the functions of my desires. The content of a belief becomes whatever condition guarantees that my desires will be satisfied, where satisfaction is a matter of my desires fulfilling their proper function – that is, when my desires have their normal effect. The trouble with this sort of view, according

to Fodor, is twofold. First, he charges that the notion that mental representations have functions is fundamentally in error. Secondly, he charges that even if such an assumption is removed, a teleological account of semantics cannot avoid indeterminacy – in fact, a *huge* amount of indeterminacy – when it comes to fixing the contents of beliefs.

So Fodor begins his critique by disagreeing with the teleosemanticist about whether beliefs/desires can *themselves* be properly said to have functions. According to his construal, organisms have mechanisms with evolutionary histories which were selected for 'their role in mediating the tokenings of mental states'. Thus the function of what we have been calling RPMs is to 'produce instances of [mental representations] upon environmentally appropriate occasions' (Fodor 1990: 65). Similarly, RCMs have functions – these will be many and various, but presumably unified in the sense that they 'have worked historically to aid our survival' – and, hence, are candidates for natural selection (Millikan 1986: 55).

Now, we may regard a situation where RPMs are doing what they were selected for as a Normal (i.e a Type-1) situation (Ibid: 64).[7] To solve the disjunction problem, Fodor writes, we do not require a distinction between Normally/abNormally functioning *beliefs*. All we require is a distinction between 'Normal and abNormal circumstances for *having* a belief' (Ibid: 65). With such a distinction, we can say that the content of a mental representation is constituted by the Normal circumstances for its tokenings. The teleosemanticist may be able to provide such a distinction via reference to the proper functioning of RPMs.[8]

I certainly agree with Fodor that it is far more intuitive to think that RPMs/RCMs have functions than that the representations themselves do. However, given what we have said above about the teleosemanticist's account, it must be borne in mind that they may not accept the way Fodor construes their

---

[7] 'Normal' with a capital 'N' is supposed to indicate a normative normality, rather than a statistical one. So, something may be 'Normal' and not normally happen if cases of malfunction are particularly abundant.

[8] It is also worth noting that the teleosemanticist has a claim here to have naturally grounded the normativity of meaning. As Hattiangadi puts it: 'something with the function of Φing is *supposed* to Φ, it *ought* to Φ, and if it fails to Φ, there must be…some kind of mistake' (Hattiangadi 2007: 126).

theory, and that there may be arguments to the effect that the theory as they construe it escapes his criticisms. (Of course, the teleosemanticist would also owe us an argument for why we should accept the claim that representations have functions.) Accordingly some modifications of the following criticisms may be required if they are to have real bite.

As a matter of fact, however, Fodor has some direct criticisms of the view that mental representations are to be individuated in terms of their functions. First, and most general, is the objection that it just doesn't follow from the fact that a certain type of thing is functional that individual things of that type have functions of their own. Fodor writes:

> The assumption that it's useful to have cognitive states does not entail that you can distinguish among cognitive states by reference to their uses. It's a sort of distributive fallacy to argue that, if having beliefs is functional, then there must be something that is the distinguishing function of each belief (Fodor 1990: 66).

To take a variant on one of Fodor's examples: having *fingers* is useful – fingers have functions. They allow us to manipulate objects. However, Fodor thinks it is an error to ask: 'What is the function of *this* finger?' since it doesn't follow from the fact that fingers have functions that 'each [finger] has its distinguishing function' (Ibid). So we *cannot* conclude from the fact that having beliefs is functional that *this particular belief* (say the belief that 5 is a prime number) is functional.

Fodor has more to say on the topic of individuating representational states by their functions. He observes, correctly, that functions are a 'species of Normal effects'. Thus we can individuate the heart by the fact that it functions to pump blood. But what about desires? Fodor takes this example: the desire to be rich and famous. What does this desire typically effect? Well, as a matter of fact, sometimes nothing at all, and sometimes *trying* to be rich and famous. Fodor points out with regard to the former possibility that:

> [The fact] that it is possible to have wants that are *arbitrarily causally inert with respect to their own satisfaction* is, indeed, one of the respects in which wants are intentional; it's what makes wanting so frightfully nonfactive (Fodor 1990: 69, my emphasis).

The point is just that I can 'want like stink' for something to be the case, but this is not only not 'Normally sufficient to bring it about', furthermore 'my wanting it doesn't alter in the slightest scintilla the likelihood that it will happen' (Ibid). With regards to the second possibility, Fodor points out that 'trying is no good for the job at hand since it is itself an intentional state' (Ibid: 67). The point here seems to be that it is no good to try and individuate a representational state via reference to *other* representational states. I presume this is no good in general for reasons mentioned above: first because circularity looms large, and secondly because belief holism may show that we'll never get anywhere this way.[9]

Actually, the first point possibly needs to be stressed more. If what desires effect are, in fact, other intentional states, then the teleosemanticist hasn't really explained how the contents of those desires are constituted. If desires cause me to *intend* to bring certain things about, or to act *intentionally* to bring them about, then content does seem to be sneaking into the individuation of content. This would seem to be the case – I doubt that the teleosemanticist would like to say that desires effect, as it were, *brute* behaviour towards their own satisfaction. Now, it may well be correct to say that desires, when accompanied by appropriate beliefs, sometimes effect intentional action which is directed towards their own satisfaction. But surely we cannot claim to individuate them this way. For one thing, as has been mentioned, it seems possible to have the relevant belief and desire and do *absolutely nothing* towards achieving the object of the desire. Maybe the teleosemanticist can respond by arguing that such a case is abNormal. However, it seems that we have no problem in attributing – to ourselves and others – desires that we do not act upon. If desires are to be individuated in terms of their Normal effects, how are we able to do this in the (perpetual) absence of such effects?

Finally, what about a third possibility – *actually* becoming rich and famous? The problem with this suggestion, according to Fodor, is that it is 'wildly implausible that it is, in any non-question-begging sense, a Normal effect of *wanting* to become it' (Ibid). That is, it is extremely rare that people who desire to

---

[9] On this last point, and specifically for Fodor, trying to fix the content of a representational state in terms of other ones flies in the face of his beliefs about the *atomic* nature of intentional states (i.e. that it's possible to have any one without having any others).

be rich and famous actually become so. So the only way to conclude that this is the Normal effect of the desire is to, in Fodor's words, 'sneak a look at the intentional object of the want' (1990: 67). This objection would seem to apply to cases where people desire something, but do nothing about it. The teleosemanticist can only conclude that such cases are abNormal on the basis of *knowing what the intentional object of the want is*. But how can they claim to do so – especially in cases of what we might call pervasive and systematic abNormality? It does look as if the account tacitly supposes what it is supposed to explain.

On the basis of all this, Fodor concludes that the teleosemanticist would do better to give up the idea that mental representations themselves have functions. He then turns to a teleological account that drops this assumption. Such an account would still have it that there are RPMs, that these have functions, and that the function of the RPMs is to 'mediate…causal relations between environmental states on one hand, and mental states on the other' (Fodor 1990: 69). The thought then is that possession of the RPMs bestows selectional advantages on the organisms that have them in virtue of their causing token representations of certain types in certain Normal, naturalistically specifiable circumstances. Let's illustrate this with an example. Suppose a frog has an RPM which causes the representation 'fly' to be tokened whenever certain circumstances obtain. This would presumably play a role in the selective fitness on the frog *viz.* allowing it to tell when there is food in the vicinity. So, we might say, the RPM was selected for because it produces representations which are germane to the survival of frogs, specifically with their ability to locate food.

Putting this in more general terms, we may say that Normal situations are as follows:

In circumstances C, if mental representations of type S are caused by instantiations of property P, then tokens of type S mean property P.

Now, Fodor argues, it is only in Normal circumstances that the causation of S tokens by the RPM is taken to be constitutive of content. So in not-C (i.e. non-Normal circumstances), S tokens are free to be caused by whatever may happen to

cause them, without the threat of the disjunction problem (1990: 69-70). Putting all this together we get:

i.   S is reliably caused by flies in C
ii. S is the Normal cause of an ecologically appropriate, fly-directed response
iii. S was selected because of (i) and (ii)
iv. S means *fly*

<div align="right">(Ibid: 70)</div>

Sadly, however, Fodor thinks that such an account does not give us a successful solution to the disjunction problem. I shall now try to explain why not.

The problem with this kind of account is, according to Fodor, that it cannot avoid massive indeterminacy when it comes to specifying the contents of (in this case the frog's) representational mental states. Consider the story we have just told. We said that the frog has a mechanism which causes state S in the presence of a fly. However, we could (and, indeed, a semantic sceptic might) choose to tell the story this way: the frog has a mechanism which causes S to be tokened in the presence of a 'little ambient black thing' (Ibid: 71). Crucially, whatever the content of S is (*fly* or *little ambient black thing*), the contribution of S to the fly-directed response of the frog *in an environment where little ambient black things are flies* will be exactly the same. Thus either would do just as well if we are to judge the content of the representation by the effect it has on the frog's behaviour. So, we can claim with just as must justification by an analogous argument to the one above that S means *little ambient black thing*.

In more general terms, suppose that the following holds in nature: Fs and Gs are correlated such that $F \leftrightarrow G$. Now suppose that an organism requires Gs to survive, so that we might expect there to be a selectional advantage to an organism that was, in some way or other, able to represent the presence of Gs to itself. As Hattiangadi points out, in general terms there are two options. Either the organism can represent Gs, or it can represent Fs and let the correlation do the rest of the work (Hattiangadi, 2007: 130). One proposal at this point is to insist that it is what the organism *actually needs* that is the content of the representation. Millikan (1995: 93) comes down heavily in favour of this option. However,

Dretske has a compelling response. I may need vitamin C to survive, but surely we would not want to say that the meanings of my mental representations 'lemon' and 'orange' is actually *vitamin C*. As Dretske puts it, when it comes to satisfying my need for vitamin C, 'representing things as oranges and lemons will do quite nicely' (Dretske 1986: 32).

Millikan has pushed this line of response elsewhere. Arguing against Fodor she insists that:

> The systems that use, that respond to, the frog's fly detector's signals, don't care at all whether theses correspond to anything black or ambient or specklike, but only whether they correspond to frog food…[So the firing of such signals] means frog food (Millikan, 1991: 163).

Now, as Fodor rightly points out, this is utterly question begging. We can happily describe an environment where we can view the frog's RCMs for fly detection as 'caring *precisely* about whether the signals mean' something ambient and specklike: an environment 'where the black specklike things are largely food' (Fodor 1991: 265). To presuppose what the RCMs 'care about' is to presuppose exactly what's at issue, *viz.* the contents of the representations.

Finally, consider again the semantic sceptic. Suppose that we are trying to defend the teleological account of the frog's meaning *fly* by the token 'fly'. Crucially, we need to make reference to Normal circumstances – i.e., the presence of flies, proper functioning of the frog's RPMs and so on. Could not the sceptic argue that the frog in fact means *little ambient black thing* by 'fly'? What could the teleosemanticist appeal to in order to rule out this possibility? Certainly not the representational state of the frog – that's exactly what's at issue. It seems that the teleological account of semantics has neither solved the disjunction problem, nor undermined the sceptical argument.

4.3.6 Cause Theories

I propose to leave the discussion of effect theories at that. Since they seem to face highly troubling difficulties of bizarre attribution of function, and huge indeterminacy about content, I think it best to look elsewhere for a response to the

semantic sceptic. In this section we turn to some members of the family of theories that attempt to fix content the other way round: cause theories.

Information Semantics

Informational views of meaning offer a causal theory of how the contents of mental representations are to be fixed. A general question one may ask at the outset is this: in virtue of what are mental representations *representations*? After all, there would seem to be many mental states which are *non-representational* – pains or emotions, for instance. What does the difference consist in? Or, to put it a different way – what exactly *are* mental representations (or what are representational states, if that's the same thing). The first plank in an informational theory of meaning is the thought that an entity is representational partially in virtue of the fact that it *carries information about* something. For x to carry information about y is for x to indicate that y is the case. (For this reason informational semantics is also referred to as indicator semantics.) This is usually cashed out in causal terms. So, if x carries information about, or indicates y, then y causes x.

However, carrying information is clearly not sufficient for the kind of meaning in which we are interested. First of all, instances of what Grice called *natural meaning* are clearly cases where we have information in the above sense. Fire causes smoke, so smoke carries the information that there's a fire. Thus we may say that smoke *naturally* means fire. But we would want to stop short of saying that smoke *non-naturally* means *fire*, or that smoke *represents* fire. The general point for a semantics of mental representations here is that all sorts of things may cause a mental representation to occur. Take a thought about cats. Cats are one possible cause of a thought about cats (perhaps, really originally, the thought that 'there's a cat on the mat'), but so *inter alia* are dogs, tigers, mice and other thoughts (for example, the thought that dogs are pack animals might well prompt the thought that cats aren't). Conversely, cats may actually fail to cause tokenings of 'cats'. For all that, all and only cats are in the extension of 'cat'.

One suggestion for solving this problem would be to opt for a *ceteris paribus* type of account which made use of the idea of idealised conditions. The

idea is that in conditions of Type 1, only cats will cause my cat thoughts, and so it is only cats that are the extension of 'cats'. We have seen how such an account faces some pretty severe objections both in the discussion of sophisticated dispositionalism, and in Fodor's reconstruction of the teleosemanticist's account (Fodor, you recall, views teleosemantics as a kind of Type 1 account). So we need another way of solving the problem. This Fodor attempts to provide in his Asymmetric Dependency account.

Asymmetric Dependence

To begin with, as we have seen, Fodor thinks that an approach that tries to distinguish *correct* from *incorrect* tokenings of a mental representation by appeal to *ideal conditions* is hopeless. So instead, he tries to make sense of the distinction by appealing to the *causal relationships* between extension-caused (correct) tokenings of a representation and non-extension-caused (incorrect) tokenings of it. Thus Fodor's account is an informational account: it tries to make sense out of meaning in terms of the *causes* of mental representations. Recall that the notion of information is cashed out in this way: if x causes y, then y carries information about x. So Fodor is concerned to build an account of meaning out of the concept of carrying information. There is one important caveat to mention here: Fodor hopes to provide an account of a *sufficient* condition for representation's having a certain content by providing an account of a *difference* in such causal relationships (so not a necessary and sufficient one – more on this later).

So, here again is the disjunction problem applied to a causal account. Lots of things can cause a mental representation to be tokened. Some of these, intuitively, will be in the extension of the mental representation, some not. How, then, can we draw a principled distinction between those causes *within* the extension, and those *without* it? Fodor's idea is this: the latter tokenings (those caused by things *outside* the extension of a representation) *asymmetrically depend* on the former (those caused by things *within* the extension of the representation). What does this mean exactly? The thought, to put it roughly, is this: were it not for the fact that things in the extension of a representation caused tokenings of it, nor would things that were *not* in the extension of that representation. However, the converse is not true: things that are in the extension of the representation

129

would *still* cause tokenings of that representation, even if the things that are not in its extension didn't.

Let's take horses as an example. Horses cause 'horse' tokens, and so do (let's say) cows. But were it not for the fact that 'horse's carry information about horses, 'horse's '*wouldn't carry information about anything*' (Fodor 1990: 91). Fodor brings out another important point here. Since many non-horses may cause tokenings of 'horse', we cannot *identify* meaning with information. Meaning, as Fodor puts it, is *robust*: although 'horse' tokens get caused in all kinds of ways, they all still mean *horse* (Ibid: 90-91). Thus 'the trick' in constructing a successful theory of meaning is to explain how meaning is 'insensitive to the heterogeneity of the (actual and possible) causes' of tokens of mental representations while, on the one hand, 'meaning is supposed somehow to reduce to information' and, on the other, 'information varies with etiology' (Ibid: 90).

The final important feature of Fodor's theory to bring out is that these causal relationships between things and tokenings of mental representations are not meant to be simply matters of 'brute fact'. Rather, the connections are *nomological* or *law like*. Note that above, we expressed asymmetric dependence in terms of counterfactuals: that is, in terms of what things *would* cause tokenings of mental representations under certain conditions. For this reason Fodor's view may be broadly construed as a dispositional view (though obviously not of the same type as either the crude dispositional theory, or sophisticated dispositionalism). Now, Fodor remarks:

> I assume that if the generalisation that Xs cause Ys is counterfactual supporting, then there is a "covering" law that relates the property of being X to the property of being a cause of Ys: counterfactual supporting causal generalisations are (either identical to or) backed by causal laws, and causal laws are relations among properties. So, what the story about asymmetric dependence comes down to is that 'cow' means *cow* if (i) there is a nomic relation between the property of being a cow and the property of being a cause of 'cow' tokens; and (ii) if there are nomic relations between other properties and the property of being a cause of 'cow' tokens, then the latter nomic relations depend asymmetrically upon the former (Fodor, 1990: 93).

There are two important points to bring out here. First, the way we put Fodor's account at first is not entirely accurate: for Fodor it's not a causal relation between cows and 'cow's (i.e. 'cow' tokens) that all other such causal relations

asymmetrically depend on, but rather a causal relation between *the property of being a cow* and 'cow's. Secondly, we need to investigate the relationship between laws, dispositions and counterfactuals. First, though, let's review the big picture. Meaning is to be built out of information – the causal relationships between things (or, more accurately, properties) and the mental representations. However, for any given representation, many things (properties) that are *not* part of the intuitive extension of that representation nevertheless cause precisely that mental representation to be tokened. So we need a way of distinguishing between the extension-causes of the representation and the non-extension causes. Fodor's idea is that the latter asymmetrically depend on the former: if the former didn't cause the representation to be tokened, neither would the latter; but, conversely, if the latter didn't cause the representation to be tokened, the former still would.

Laws, Dispositions and Counterfactuals

We noted above that the relations Fodor is trying to capture between properties and tokenings of mental representations are supposed to be not only causal, but also nomological (law-like). There is a general problem for Fodor lurking here, and it is this: one might think that nomological relations can be rephrased in terms of *dispositional* properties of those things that the laws govern. If this analysis is correct, then it seems that Fodor's account will be beset by the same problems that scuppered the sophisticated dispositional account.

To see how we might analyse nomological relations dispositionally, let us take for example the law that if one places sugar in water, it will dissolve (Miller, 2007: 219). To say that 'it is a law that if one places sugar in water, it will dissolve' may well be to say just the same thing as 'sugar is disposed to dissolve in water'. Now we are assuming here a counterfactual analysis of dispositions (as opposed to the realist account of Martin and Heil, see discussion above). How does all this impact on Fodor's account?

Well, Fodor has it that there is a law-like relation between the property of being a cow and the property of being a cause of 'cow' tokens. We may be tempted to analyse this, given what we have said above, in terms of counterfactual about individuals. To give an example (adapted from Miller, 2007: 219):

It is a law that cows cause Jones to token 'cow'

may be rephrased as:

Jones is disposed to token 'cow' in the presence of cows

which, given the counterfactual analysis of dispositions, we may write as:

If Jones were confronted by a cow, Jones would token 'cow'

Now, we have seen that this last formulation is fairly hopeless as an attempt to give an analysis of meaning, and that the most promising (though unsuccessful) attempt to salvage it was to add a *ceteris paribus* clause, where this clause was understood as an idealisation of the circumstances under which Jones would token 'cow':

If Jones were confronted by a cow, then *ceteris paribus* Jones would token 'cow'

Equivalently:

It is a law that, *ceteris paribus* cows cause Jones to token 'cow'

If Fodor's account is committed to something like this formulation, then perhaps he has to put a *ceteris paribus* clause into his account. However, he then faces the familiar and formidable problems encountered above of specifying the idealisation expressed by the *ceteris paribus* clause.

Further Problems for Asymmetric Dependence

There are, however, several further pressing difficulties that Fodor's account faces. The first we might call the 'no clear priority' objection. The second is that there is a charge from Boghossian that the account is *really* a Type 1 account in disguise. This would be disastrous, since asymmetric dependence

would again fall prey to the objections that beset sophisticated dispositionalism. The third difficulty is this: there still remains a nagging worry (for myself and others) that we don't really have a fully satisfactory account of representation here.

The first problem – the 'no clear priority' objection is attributed to Putnam by Hale and Wright ([1997] 2006). The thought is this: there seems to be no clear priority between, say, 'if horse didn't cause "horse"s, then pictures of horses wouldn't cause "horse"s' and 'if pictures of horses didn't cause "horse"s, then horses wouldn't cause "horse"s'. Hale and Wright write:

> Rather what seems to be true is that it is because 'horse' refers to horses that *both* horses *and* pictures of horses – and thoughts of horses, and cows in a darkened field [and so on] – elicit, *ceteris paribus*, tokenings of 'horse' ([1997] 2006: 444).

Hale and Wright go on to cash this out in terms of possible worlds. The objection is this: consider the closest worlds in which pictures of horses don't cause 'horse's. Well, it is hard to see how this could be the case if 'horse' referred to horses. So, if pictures of horses at world *w* don't cause 'horse's to be tokened, that can only be because 'horse' *doesn't refer to horses* at *w*. Thus 'the closest worlds in which pictures of horses do not cause tokenings of 'horse' are worlds in which horses don't either' (Ibid).

Let us consider the second difficulty. On Fodor's theory, a mental representation R means that P just in case the following conditions hold:

(i)   It is a law that Ps cause R tokenings

(ii)   For all other properties that cause R tokenings, the fact that they do so asymmetrically depends on (i).

The problem, as Boghossian sees it, is that if the representation R 'possesses asymmetric dependence base P, then there exists a world in which *only* P can serve as a cause of tokenings of [R]' (Boghossian, 1991: 71). In other words, there *is* a Type 1 situation implicit in Fodor's theory.[10] In fact, Boghossian argues that there *has* to be. For suppose that there were *no* such world in which only Ps caused Rs. Well then, on what basis could we conclude that P is R's asymmetric

---

[10] Perhaps this is a bit quick: Fodor could hold that it is possible that only P should cause tokenings of R without committing to the idea that one could specify under which conditions this would be the case. My thanks to Bob Hale for pointing this out.

dependence base? Moreover, suppose we have a set up of possible worlds like this:

W1: Cows cause R; horses cause R; cats do not cause R

W2: Cows cause R; horses don't cause R; cats cause R

Let us suppose that we want to say that R means *cow*. Now, Boghossian wonders, how are we to rule out the possibility that the asymmetric dependence base for R at W1 is not cows, but rather ((cows and horses) or (cows and cats))?

Fodor (1991) has a response to this. First of all, note that cats do not cause Rs in W1. Suppose W1 is the actual world. Well, cats *can't* be part of the asymmetric dependency base for R since for R to mean *cats* (or any formula containing *cats*) it *must* be a law that R carries information about cats *in the actual world*. Secondly, he remarks that *ceteris paribus* laws are notorious for not satisfying the pattern of inference that Boghossian is working with here. Boghossian has it that:

((X causes R) & (Y causes R)) → ((X&Y) cause R)

For a start, X and Y may not be 'nomically compossible'. Secondly, even if they are, their 'respective side effects may interact to cancel [R]'. Finally, since Fodor's theory is only intended to provide a *sufficient* condition for R to mean/have the content *horse,* then since the fact that we can find some *other* asymmetric dependence base really shows us that we may not have a *necessary* condition for it to so, this doesn't really tell against the theory (Fodor 1991: 272-3). I'm not sure how convincing all this is against Boghossian. I'm prepared to leave the jury out on this one, since I think there is a deeper worry for the asymmetric dependency theory, which I turn to now.

It is possible to see the sceptical argument as turning several deep seated intuitions about meaning in on themselves, as it were, and arguing that there is not fact to answer to all of them. Now, Fodor may well be better placed to answer to some of these intuitions than the sophisticated dispositionalist. For example, the sceptic raises a question with regards to my *knowledge* of my meaning: it is

immediate and non-inferential, whereas my knowledge of my dispositions is the opposite. Since Fodor can appeal to mental representations, this problem seems to be averted – they are *my* mental representations, after all. However, the problem now is that, as Miller puts it, it is hard 'not to sympathise' with the following criticism raised against Fodor by Crane:

> [It is difficult] to see how asymmetric dependence goes any way towards *explaining* mental representation. I think that the conditions Fodor describes probably are true of mental representations. But I do not see how this helps us to understand how mental representation actually works (Crane 2003: 180-181).

I think this may well be right. What the Asymmetric Dependency gives us is a story about *why* it is that a mental representation (say 'cat') represents what it does (*viz.* cats); but what it doesn't really seem to give us is a story of what it is for a mental representation to represent *anything at all*. That is, I am not entirely sure of *how* a mental representation represents – what are mental representations? What are they like? What is it that stands in the asymmetric dependence relation to properties in the world? It seems to me that in order to satisfactorily answer the sceptic, these questions must be addressed. Thus while I think a causal theory such as Fodor's is the most promising, we still do not have a fully satisfactory answer.

4.4 Conclusion

In his overview of naturalistic semantic theories, Loewer ([1997] 2006) claims that all attempts to naturalise semantics fail. This is because they either fail to provide determinate extensions for expressions (that is, they fail to solve the problem of error), or that they can only do so by smuggling in semantic/intentional notions, and thus fail to be truly naturalistic. In terms of Kripke's sceptical argument, we may say that such theories fail to rule out alternative sceptical interpretations of expressions, or else violate the terms of discussion with the sceptic. Our investigations in this chapter can only lead us to concur with Loewer's conclusion.

We have looked at a number of naturalistic/reductive responses to the sceptic which are intended to solve the problem of error. I have argued that first, some of the theories on offer *clearly* fail to solve the problem; and, second, that none of the theories on offer *clearly* succeed to solve it. At this point, it is

tempting to draw two pessimistic conclusions. The first is that the strong normativity requirement is not the be all and end all of the sceptical argument, and that the argument has plenty of powerful resources even once this requirement is relaxed. The second is that it certainly seems that the sceptic has won at this point, at least insofar as we are concerned with naturalistic theories of meaning.

However, as I will argue in the next two chapters, there are good reasons to think that a naturalistic reduction of meaning is likely to fail *other than* there simply being no facts of the matter about meaning. These have to do with the nature of scientific inquiry, and of semantic phenomena. I will argue that we are not in a position to know what the resources open to a naturalised semantics would ultimately be, but, further, that semantic phenomena may be such that they will resist reduction to scientific categories in any case. So the optimistic conclusion at this point is that although we have not been able to meet the sceptical challenge within a naturalistic framework, we needn't necessarily be too perturbed by this, as we *may* be able to show that there are reasons for this other than the truth of semantic non-factualism. So, anyway, will I try to convince you.

Semantics and Naturalism (I): The Reduction Base

5.1 Introduction

Straight responses to the sceptic come in two broad categories: reductive and non-reductive. A satisfying solution, it used to seem to me, would have to be of the former kind. Taking myself to be committed to some form of philosophical naturalism, I believed that meaning facts could not be *sui generis* and that we ought to be able to account for them in terms of more 'naturalistically respectable' facts about speakers. Indeed, it is hard to see why one would view semantic facts as inherently suspect unless one had such a metaphysical commitment – that is, a commitment to a view about which facts are 'respectable'. I have come to view this line of thinking as fundamentally misguided, and it is my aim in this chapter and the next to explain why. Rather than attempting a reductive solution to Kripke's paradox of the kind I originally thought to be the only satisfactory option, I now wish to partially align myself with those who think that the reductionism in Kripke's argument is unwarranted.[1] My claim is that there are good reasons to think that a reduction of semantic/intentional facts to non-semantic/non-intentional facts is likely to fail *other than* there simply being no facts of the matter about semantics/intentionality. Thus we can accept the failure of all extant reductive responses to the sceptic *without* conceding that there are no facts of the matter about meaning. (This does not rule out the possibility of a future reduction of the semantic/intentional to the non-semantic/non-intentional, though I am somewhat sceptical of such an eventuality for reasons that will become clear.)

My strategy will be to consider more generally the assumptions underlying the 'naturalistic' approaches to the problems of semantics and intentionality that we encountered in the last chapter. These approaches are fundamentally reductionist, and at least claim to have a clear conception of the 'naturalistically respectable' facts to which one might appeal in a reductive response to meaning-scepticism. I hope to provide an account of why meaning-constituting facts cannot

---

[1] See, for example, McGinn, 1984; Wright, 1984.

be characterised in terms of the entities of the natural sciences which, nevertheless, is consistent with meaning-factualism. Thus, even if we concede that the programme of 'naturalising semantics' fails, the sceptic's conclusion that there simply are no facts of the matter about meaning can still be resisted.

Part of the reason for focusing on naturalistic theories of meaning and content in arguing that there is no pressing need for a reductive account of those concepts is that naturalism and reductionism have been taken to run together in some of the most recent writing on the sceptical argument. For example, Hattiangadi writes:

> A semantic realist could also be a semantic naturalist, who claims that what makes it true that I grasp the meaning of a word are ordinary 'natural' facts, which are ultimately physical, causal, or functional. Alternatively, a semantic realist could be an anti-reductionist about semantic facts, holding that semantic facts are *sui generis* and irreducible (Hattiangadi 2007: 5).

> Anti-reductionists maintain that intentional, semantic, or normative facts are irreducible to natural facts, and therefore must be assumed as primitive (Ibid: 151).

In the passages above, anti-reductionism is set up in opposition to the view that semantics may be reduced to *natural facts*. Thus, it seems, naturalism and reductionism are taken to go hand in hand. Notice also the assumption that 'natural' facts are distinct from semantic or intentional ones. Whether or not there is a principled characterisation of natural facts which excludes semantic, intentional, or even normative facts is one of the main focuses of this chapter.

5.2 What is 'Naturalism'?

In the last chapter we looked at several attempts to naturalise content, without going into too much detail about exactly what philosophical naturalism involves. If naturalism is to provide the 'facts' with which one can respond to the sceptic, while simultaneously outlawing unreduced semantic or intentional 'facts', then it seems that we'd better have a clear idea of exactly what it is. Thus I think we may ignore recent calls to eschew any definition of the term (cf. Papineau 2007). One position we might call *metaphysical naturalism*. This is the view, expounded by, for example, Dennett (1984) that 'philosophical accounts of our minds, our knowledge, our language must in the end be continuous with, and

harmonious with, the natural sciences' (1984: ix). The main contention of the kind of naturalism at issue here is that 'everything there is, every individual, property, law, causal relation, and so on is ontologically dependent on natural individuals, properties, and so forth' (Loewer [1997] 2006: 108). The goal of a (metaphysically) naturalised semantics is thus to demonstrate that all semantic properties and relations can be shown to be 'ontologically dependent' on natural properties and relations (Ibid). Loewer understands the concept of ontological dependence as involving the following claim:

> For each instantiation of property M there are instantiations of natural properties and relations, P, P*,…,that together with natural laws and causal relations among the P instantiations *metaphysically* entail M's instantiation (Ibid).

While this relation is perhaps not exactly the same relation as that of *constitution* that Kripke challenges us to account for in the sceptical argument, it seems that a successful naturalisation of semantics would be sufficient to meet his challenge. That is, if we could show that Jones' meaning *plus* by 'plus' ontologically depended on well understood natural properties, relations, laws, and so on, then we could claim to have a fairly full story about what *constitutes* Jones' meaning *plus* by 'plus'.

However, as we saw in the previous chapter, the prospects for the specific proposals to naturalise semantics that are on the table look pretty bleak. Teleosemantic accounts of meaning and mental content seem to face insurmountable problems of indeterminacy in the contents they assign to mental representations, while Fodor's nomological account of mental content seems to say far too little about exactly what being a mental representation consists in; that is, in virtue of what do mental representations *represent*? What is at stake here if such projects fail? Loewer ([1997] 2006: 110) notes that we face a kind of trilemma. Either semantic naturalism is true; or there are no semantic properties (Churchland 1981); or semantic properties are 'metaphysically independent' of other properties (Davidson 1984; McDowell 1994). Loewer notes that the second horn of the trilemma – semantic eliminativism – has been taken to be self-refuting (Boghossian 1990), or obviously false in light of the success of 'folk-psychological' explanations (Fodor 1987). In any case, as a realist about meaning, I think it is something to be avoided if at all possible! Loewer points out further

that if the third horn is true, then, assuming the causal completeness of the sciences, semantic properties are unable to play any causal role in behaviour (Loewer [1997] 2006: 110). 'Neither option' he remarks 'is particularly pretty'.

Let us consider metaphysical naturalism further. We may learn a little more about the entities that metaphysical naturalism licenses by considering what have been taken to be the two main challenges to the position. Loewer outlines these as follows. The first challenge stems from the putative normativity of meaning. The thought here is that naturalistic properties cannot instantiate semantic properties, because semantic properties are inherently normative, whereas naturalistic properties are not. It is unclear whether this argument requires strong (prescription-giving) or weak (correctness-conditions) normativity. Fodor, for one, thinks that a naturalisation of semantic relations such as extension is enough to account for the latter (1990: 135-6). The second challenge is indeterminacy. The problem is that the totality of naturalistic facts about a speaker is supposedly not sufficient to rule out 'alternative assignments of truth conditions and references to that [speaker's] sentences and terms' (Ibid: 109). Thus naturalism cannot avoid massive indeterminacy in the meanings/contents it assigns to utterances. On the assumption that, notwithstanding the claims of the sceptic, there *are* determinate facts about the meanings of utterances, a naturalistic metaphysics seems at best incomplete as it fails to give an account of such facts. Note that if Fodor is correct this argument is related to weak normativity: the hope is that a naturalisation of semantic relations will serve both to set correctness conditions (extensions, referents, and so on) *and thereby* solve the indeterminacy problem.

I hope this provides a fairly clear idea of the hopes and fears for one way of understanding naturalism about meaning and content. I will now try to show in general terms why such projects of naturalising semantics fail. This ultimately boils down to an instability in the very characterisation of the 'natural' facts that such naturalists are working with, as well, perhaps, as a lack of clarity concerning the entities to be naturalised. We will focus on the former problem in this chapter, and the latter problem in the next.

5.3 Why Projects of Naturalising Semantics Fail (I): The 'Natural' Facts

Proponents of metaphysical naturalism claim that everything there is ontologically depends on natural entities, properties, and so on for its instantiation. A natural question at this point is this: do we have a principled way of determining what the natural entities, properties, and so on *are*? If we are to make the claim that all things depend on natural things for their instantiation, it seems that we ought to have at least a clear conception of what the natural things are.

5.3.1 What are the Naturalistic Facts?

From the two challenges to the naturalisation of semantics, we know certain things about the supposed *character* of naturalistic facts. First, since the normativity of meaning is taken to be something that naturalistic facts are unable to account for (even on the weaker reading where the normativity of meaning boils down to the fact that meaningful expressions possess conditions of correct application), it seems that a feature of natural facts is that they do not have the kinds of (weakly) normative implications that are taken to be characteristic of meanings: roughly, that certain uses of expressions are (in)correct. One may presumably speculate that in general naturalistic facts do not imply that one course of action is (in)correct, or that some action ought/ought not to be carried out. So perhaps the domain of moral facts is another domain that needs to be 'naturalised' on such a view. However, we should not rest content with negative characterisations of naturalistic facts if we are serious about carrying out a successful reduction of non-natural facts to them. We need a positive construal which spells out exactly what naturalistic facts are. However, as Loewer wryly notes, '[metaphysical] naturalists are seldom explicit concerning exactly which properties are the natural ones'. He goes on to outline the current 'working account' employed by metaphysical naturalists as follows:

(MN) The natural properties are those expressed by predicates appropriately definable in terms of predicates that occur in true theories of the natural sciences (Loewer [1997] 2006: 108).

Note that this provides us with the following gloss on the metaphysical naturalist's position:

(MN*) All properties and relations ontologically depend on such properties that are expressed by predicates appropriately definable in terms of predicates that occur in true theories of the natural sciences.

5.3.2 True and complete theories of the natural sciences

An initial worry at this point, in an instrumentalist vein, would be this: do we in fact have any true theories in the natural sciences? We may have good explanatory theories that organise, economise and explain a wide range of phenomena, and that mesh with other scientific beliefs, but this does not give us license to claim that these theories are *true* – unless 'true' doesn't amount to anything more than playing a role in one of our best explanatory theories. We *can* claim with confidence that they are the best models of the way things are that are currently available to us. Perhaps we can do no better (Chomsky 2000b). A second worry is this: even granted that we do have some true theories of the natural sciences, how do we know which theories are the true ones? This kind of scepticism may well derail the whole project of naturalising semantics before it has begun.

A possible reaction that one might have to the second worry is this: well, we don't know which theories are true, but we can still consider whether *if our currently accepted theories are true,* we can explain semantic phenomena on the basis of the facts as they are according to those theories.[2] I think this point is fair enough, but it is important to be clear about what would follow if, say, semantic phenomena turned out to be *inexplicable* in terms of our current theories. If we adopt this approach one could argue that there are several possible reasons for the failure of an explanation of semantic phenomena in naturalistic terms. For instance, it could be that there is something wrong with our conceptions of semantic phenomena; or it could be that our best theories are in fact *false*; it it could be that our theories, while true, are incomplete precisely because they fail to

---

[2] My thanks to Bob Hale for alerting me to this kind of response.

allow us to account for semantic phenomena. However, if we cannot certify that our current theories are true theories, the fact that semantic phenomena cannot be explained on the basis of those theories does *not* license the extreme scepticism that Kripke argues for. In other words, if we cannot be certain that our best theories are true theories, then we cannot conclude that there *are no* semantic phenomena on the basis that we have failed to account for semantic phenomena in terms of our best theories.

Perhaps the metaphysical naturalist can instead respond to this worry over whether our current theories are true by insisting that naturalistically respectable properties are expressed by predicates appropriately definable in terms of predicates that would occur in true theories of some idealised *complete* natural science. The problem with this idea is that it is clear that we are not at any such stage of science yet. Thus we would have no fixed conception of the naturalising base. Therefore it is futile to attempt the kind of reduction of semantic properties that the metaphysical naturalist has in mind at the present time.

If the metaphysical naturalist would rather make the claim that science as it currently stands can provide all of the requisite materials for a successful reduction of semantic phenomena, then they face several severe difficulties. First, they face the formidable problem of certifying that certain scientific theories are, in fact, true. Second, it would also be necessary to demonstrate that the theories that have been shown to be true are germane to a naturalisation of the 'non-natural' facts in question. In other words, even on the supposition that we have identified some *true* scientific theories, it may simply be that our scientific theories, while true, are incomplete, *precisely* because they cannot account for semantic/intentional facts. This is not to impugn the work of any particular scientist; or, indeed, the work of scientists in general. Come to think of it, it would be an odd thing indeed to impugn a physicist, say, for not accounting for semantic phenomena in his theories. The physicist is likely to reply (if he is feeling polite and extraordinarily patient) that his theories are not *intended* to account for such phenomena in any case. Thus not only would we need to be convinced that we had *some* true scientific theories, it would further be necessary to determine that we had *all* of the true theories we needed. Perhaps we have *some* of the requisite

concepts to naturalise semantics in the way envisaged, but perhaps others remain hidden for some further scientific inquiry to uncover.

Collins (2010) has remarked on these problems. On the idea that all of the *relevant* portions of scientific theory might well be complete[3], he writes:

> It is not clear that there is a particular (higher) portion of physical theory that is uniquely germane to the naturalisation of the mind. Penrose (1989) among numerous others, for example, has speculated on the relevance of quantum mechanical effects to the structure of mental properties. Surely a materialism/naturalism should not rule out such hypotheses, whatever their plausibility, in the name of naturalism (Collins, 2010: 43).

The problem is that *we just don't know* which scientific theories will turn out to be germane to the naturalising project.

Let us suppose, however, that we do have all the true theories of the natural sciences which are germane to the naturalisation of semantic properties and which, furthermore, are complete in the sense that they are 'empirically adequate over [certain] phenomena'. Collins points out that 'the very idea of [such] portions of science being complete *in the relevant sense* [for a successful reduction of semantic properties to "naturalistic" ones][4] is somewhat dubious' (2010: 43, my emphasis). Let us consider what Collins means by complete 'in the relevant sense'. First, he points out that while some portion of high-level theory may be 'complete' *viz.* 'empirically adequate over [certain] phenomena', our *understanding* of that theory may well vary at different stages of inquiry. Thus the higher-level theory may not be 'complete' in the relevant sense (i.e. sufficient for a metaphysical naturalisation of troublesome phenomena) until we have undergone changes in our 'conceptual understanding' of that theory. Such changes in our understanding may be brought about, not by learning more about the primitive concepts of our higher-level theories, but by changes in lower-level theories.

Let me explain this a bit further. Imagine that we have a theory; call it theory A, which is complete in the sense that it is adequate over certain

---

[3] Cf. Smart, 1978; Melnyk, 1997; Lycan, 2003.
[4] Collins is concerned with mental phenomena more generally here. Nothing in his argument turns on this.

phenomena. However, theory A seems incompatible with phenomenon B in the sense that there is no way that one can make sense of phenomenon B in terms of theory A. As it stands, theory A and phenomenon B do not sit happily together: phenomenon B is isolated, not integrated into and explained by current theory. However, suppose now that a lower-level theory, call it theory C, which previously could account for theory A, but not phenomena B, undergoes some changes. This leads us to realise that we can account for both the generalisations of theory A *and* for phenomenon B in terms of a more general underlying process, now described by the revised theory C, which we had hitherto been unaware of. Thus we reconceptualise our understanding of theory A and phenomenon B, and see them both as, in some sense, special cases of the generalisations of the new version of theory C. We have, in a sense, learned nothing new about the entities invoked by, or the primitive concepts of, theory A, or about phenomenon B: no further investigation into their properties could have effected the unification. This could only come about when we reconceptualised our understanding of theory A and phenomena B in terms of the more general theory, theory C. Thus, as Collins writes, 'how we understand [some] theory at the end of inquiry, as it were, will be a function of our fundamental theory' (2010: 43).

As an example of this kind of re-conceptualisation, Collins cites the resolution of the conflict between electromagnetic theory (as characterized by Maxwell's equations) and inertial mechanics under special relativity (Ibid). The important point for our purposes is that while Maxwell's equations were preserved, their *interpretation* changed in light of the new mechanics, thus resolving the apparent conflict (Collins, personal communication). So, what may seem incompatible with a higher-level theory at an early stage may well turn out to be perfectly compatible with it 'once that theory is integrated with lower level theories'. So the naturalising project may require a fundamental change in our understanding of the relevant higher-level theory, which may only be attained by integration with lower-level theories (Ibid). Of course, this isn't to rule out the *possibility* of a successful naturalisation of semantics, only to cast doubt on the fact that we have a sufficient understanding of the naturalising base to carry it out as matters stand.

So unless metaphysical naturalists have, first, a solid proposal for determining which scientific theories are *true* theories, and second, some means of showing that the science relevant to a naturalisation of the domain in question is complete in the relevant sense, any project of naturalising that domain will fail to get off the ground. Since we simply don't know whether our best scientific theories are true, or which true scientific theories will prove germane to an understanding of semantic phenomena, or whether we need new knowledge, as opposed to changes in our 'conceptual understanding' of the relevant theories, it seems that the metaphysical semantic naturalist is in severe difficulty with respect to both of these conditions. It is worth emphasising again that even if we did know which theories were true and germane to our understanding, we would not know how integration of such theories with lower-level ones would affect our conceptual understanding of them. Thus a supposed incompatibility may well vanish at a later stage of inquiry. Thus not only is the naturalising project struggling to get started, it may even prove to be unmotivated once we reconceptualise our true scientific theories. That is, the supposed incompatibility between semantic phenomena and 'naturalistic' phenomena may well vanish at a later stage of inquiry.

A related point, and one that Chomsky is fond of stressing, is that the integration of theories and recalcitrant phenomena often does *not* proceed via reduction. Rather, 'quite different forms of accommodation' are necessary (2000a: 145). Take the mind/body problem. Chomsky writes:

> Suppose we identify the mind/body problem (or perhaps its core) as the problem of explaining how consciousness relates to neural structures. If so, it seems much like others that have arisen through the history of science, sometimes with no solution: the problem of explaining terrestrial and planetary motion in terms of the 'mechanical philosophy' and its contact mechanics, demonstrated to be irresolvable by Newton, and overcome by introducing what were understood to be 'immaterial' forces; the problem of reducing electricity and magnetism to mechanics, unsolvable and overcome by the even stranger assumption that fields are real physical things; the problem of reducing chemistry to the world of hard particles in motion, energy, and electromagnetic waves, only overcome with the introduction of even weirder hypotheses about the nature of the physical world. In each of these cases, unification was achieved and the problem resolved not by reduction, but by quite different forms of accommodation (2000a: 144-5).

The general moral is that the *integration* of scientific theories has *almost never* proceeded by reduction, at least if what we mean by 'reduction' involves the

presumption that it is always the lower-level theory that is in good order.[5] A better term is Chomsky's 'unification'. This implies no order of priority in terms of what has to be made to fit with what. The assumption underlying reductionism seems to be that our lower level theories are more likely to be in good order. History shows, however, that this is a very dubious assumption indeed.

5.3.3 Metaphysical Fixity

Another important question is this: why should we draw metaphysical conclusions about the 'reality' of entities in some domain based on the current state of natural science? Collins raises this problem for the metaphysical naturalist's thesis in this way: it assumes some sort of 'metaphysical fixity' both in terms of the facts which are to serve as the naturalising base, and in terms of those facts which are '"real" enough to be ripe for naturalisation"'(Collins 2010: 41). However, is serving in a current theory of 'natural science' good grounds for taking an entity to be 'real' and 'metaphysically fixed'?

I think it is fairly clear that the answer is 'no'. Of course, serving in a *true* theory would be good grounds for taking an entity to be in some way 'metaphysically fixed' and therefore real, but I've already surveyed the difficulties here. We aren't in a position to assess which are the true theories. Furthermore, a metaphysical naturalist who construes their position as the view that completed science will account for everything cannot take this as much more than a matter of faith, *unless* there is good reason to take the entities of current scientific theories as somehow 'real' in a sense beyond 'part of our best explanatory theories'.

In fact, further reflection on naturalistic inquiry may provide an argument that serving in some current scientific theory not good grounds to treat an entity as metaphysically fixed and 'real'. (It is important to bear in mind that something like the assumption that it is good grounds must surely underpin current attempts

---

[5] One might argue that, in some of the cases described, what happened was that the reduction base was *enlarged* to include various phenomena, and then integration indeed proceeded by reduction. This doesn't change the fact, however, that substantial changes in the ontology of the reduction base were needed for integration to occur, or that such changes were recognised to be necessary due to the need for integration with higher level phenomena. My thanks to Bob Hale for pointing out this line of argument.

to naturalise semantics.) To help understand this line of argument, it will be helpful to outline an alternative conception of naturalistic inquiry that eschews any metaphysical glosses of its core principles, and is rather couched in purely methodological terms. Following Chomsky (1993), let us call this *methodological naturalism*.

Unlike the metaphysical naturalist, the methodological naturalist makes no metaphysical claims whatsoever. Rather, they take it that mental and linguistic phenomena are part of the natural world, and thus investigation into them requires no departure from the methodology by which we usually investigate such 'natural' phenomena – that is, scientific methodology. Thus Chomsky (2000a: 76) writes:

> Let us…understand the term 'naturalism' without metaphysical connotations: a 'naturalistic approach' to the mind investigates mental aspects of the world as we do any others, seeking to construct intelligible explanatory theories, with the hope of eventual integration with the 'core' natural sciences.

The contrast between a naturalistic approach to questions about mind and language, say, and a non-naturalistic one thus consists in the methodology deemed to be appropriate in the study of such domains. Importantly, such naturalism may well have its limits:

> This approach…should be un-contentious, though its reach remains to be determined. Plainly, such an approach does not exclude other ways of trying to comprehend the world. Someone committed to it (as I am) can consistently believe (as I do) that we learn much more of human interest about how people think and feel and act by reading novels or studying history than from all of naturalistic psychology, and perhaps always will (Chomsky, 1993: 42).

Collins (2010: 44) commends methodological naturalism which he breaks down into down into eight features as follows:

i. Openness to scientific advance: naturalism involves no presumptions about the categorical resources available to kosher explanation. Our putative fundamental categories…may wax and wane.

ii. Possibility of closure: naturalism involves no presumption that every salient phenomenon submits to explanation.

iii. Conceptual invention/elimination: the concepts employed in successful scientific enquiry are typically invented for purpose and have only a loose correlation with commonsensical notions.

iv. Methodological monism: naturalism involves no presumption that distinct domains require distinct theoretical or evidential approaches. Any such differences there might be are to be arrived at, not *a priori* determined.

v. Reduction/explanation: science typically proceeds…by 'in principle reduction', where we can explain why a domain of phenomena is the way it is, without any predicate-to-predicate map or a practicable deduction.

vi. Domains/models: the domains of natural sciences are determined by on-going inquiry and invariably involve a *selecting* of relevant phenomena largely determined by the theories themselves so that a theory has first application to an idealised model.

vii. Integration: the ultimate goal of science is the explanatory integration of different domains (consilience); at a given moment in time, there is only a defeasible presumption that a putative lower level domain is explanatorily in order relative to higher domains.

viii. Principles in place of metaphysics: the hard sciences seek unifying principles (symmetries/invariances) in 'hidden parameters'…instead of laws that describe phenomena.

Now, Collins stresses that *methodologically* naturalistic inquiry, unlike metaphysically naturalistic reductive projects, requires no 'conception of a fixed naturalising base, or a fixed conception of what is to be naturalised' – where 'naturalised' here means, simply, 'submitted to methodologically naturalistic enquiry'. Thus it simply is not a *prerequisite* of methodologically naturalistic enquiry in some domain that we are able to show that it is reducible – either in a more classical sense or in principle – to some other domain, set of properties, entities or whatever. Of course, from theses v and vii, we can see that the *ultimate aim* is consilience which *sometimes* proceeds by 'in-principle' reduction; but this does not mean that in order to count as naturalistically 'real' in the first instance some phenomena must be shown to be so reducible at the outset of inquiry.

As we have seen, there are concrete examples which show that reduction-in-principle (hence integration and consilience) is sometimes only possible after revision of a *lower-level* theory. Take one of the most well supported scientific theories of the moment, which is having a huge influence on a wide range of human thought, including in philosophy: evolutionary theory. At the time at which Darwin proposed the theory, it was inconsistent with contemporary mainstream physics, as was the latest development in chemistry (chemical bonding); it was the physics, not the biology or the chemistry, which had to change (Collins 2010: 45).

Adopting the position of the metaphysical naturalist, such avenues of inquiry could have been prematurely closed off; as a matter of fact, claims to the effect that chemistry was not a factual discipline which concerned 'real' entities were advanced. The reason was that at the time there was no way to show that the entities, properties, and so on invoked in the biological and chemical theories could be shown to 'ontologically depend' on the entities invoked in the latest physical theories. This was for a good reason: the physics was wrong! As Chomsky puts it:

> Prior to the unification of chemistry and physics in the 1930s, it was commonly argued by distinguished scientists, including Nobel Prize winners in chemistry, that chemistry is just a calculating device, a way to organize results about chemical reactions, sometimes to predict them. Chemistry is not about anything real. The reason was that no one knew how to reduce it to physics. That failure was later understood: reduction was impossible, until physics underwent a radical revolution. It is now clear – or should be clear – that the debates about the reality of chemistry were based on fundamental misunderstanding. Chemistry was 'real' and 'about the world' in the only sense of these concepts that we have: it was part of the best conception of how the world works that human intelligence had been able to contrive. It is impossible to do better than that (Chomsky, 2000b: 18).

Given that such a gross mistake was made as to which properties could be considered 'real' on the metaphysical naturalist's thesis, we should proceed with extreme caution when considering the 'reality' of phenomena which are not currently well integrated with the natural sciences. Note further the somewhat instrumentalist conception of what it is to count as 'real' from a scientific perspective, as Chomsky sees it. We may draw two conclusions from this. First, that something does not fit with our current 'best conception' of how the world works at some level of theoretical understanding does not suffice to determine its

non-existence. Secondly, doing service in such a conception is no guarantee of metaphysical or ontological fixity – the notion of 'reality' in play here is much weaker. Thus we should not advance metaphysical theses on the (non)existence of entities on the basis of a lack of integration with the natural sciences: the natural sciences themselves offer no guarantee of the 'realness' of some phenomena, beyond the instrumentalist view outlined above.

Take the present day cases of cognitive science and linguistics. Do we need to show that the entities they investigate – the faculty of language, other mental systems, and so on – 'ontologically depend' on physics in order to treat statements made in them as factual? As Collins puts it:

> Formulations of [the] thought [that mental phenomena remain isolated from the kind of integrated explanation one finds elsewhere in the natural order] require no metaphysical gloss, which does nothing, it seems, other than encourage scepticism of the valuable work that is proceeding, *as if science in this domain should be uniquely held to the highest metaphysical standards that have never been adopted in the physical sciences* (Collins 2010: 44, my emphasis).

As the discussion above makes clear, the assumption of metaphysical fixity really is quite pernicious: much of the unification of scientific domains that Chomsky surveys would have been impossible had the presumption been made in physics or chemistry. It seems that the presumption primarily serves to prematurely close down valuable lines of inquiry.

The point about metaphysical fixity, in fact, cuts the *other* way as well. That is, not only can we not conclude that an entity that appears in a scientific theory is metaphysically fixed and real, neither can we conclude that such entities that are *not* mentioned in scientific theories are *not* metaphysically fixed and real. That is, it doesn't follow from the fact that there are no scientific theories of chairs and tables, or countries, communities and governments, or whatever, that these things aren't as real as things get (we will return to this point in the next chapter, see especially §6.3.1). Thus doing service in a scientific theory is neither sufficient nor necessary for existence or for metaphysical fixity. However, no more can we *presume* metaphysical fixity on the part of such non-scientific entities, at least if such a claim is understood in the sense that naturalistic inquiry must respect our

common sense categories. This latter point is especially true if we are committed to methodologically naturalistic inquiry.

5.4 Conclusion

This concludes my remarks on the 'naturalistic' facts. The metaphysical naturalist faces severe difficulties with respect to their characterisation of the 'naturalising base'. Firstly, there is no guarantee that any of our current scientific theories are true. In fact, the history of scientific theories provides excellent empirical evidence that many, if not all, of our current theories are false (or, perhaps, at best approximately true). Secondly, even if our best theories are adequate over the phenomena they explain, we may not yet have the correct conceptual resources to square such theories with recalcitrant phenomena, such as meaning and (perhaps) the human mind. Thirdly, there is the worry that the history of progress and unification in science shows us that we should presume no metaphysical fixity in our scientific categories. This point cuts both ways: serving in a scientific theory is neither necessary nor sufficient for being considered to be real or metaphysically fixed, and being a non-scientific entity neither rules out, nor guarantees, metaphysical fixity.

In the next chapter, I will move on to the other side of the equation: the characterisation of the 'non-natural' facts to be reduced. It seems to me that the metaphysical naturalist here faces a severe methodological problem: hitherto we have been concerned here with reduction *within* science (chemistry to physics, biology to chemistry, or whatever). The metaphysical naturalist's thesis, however, seems stronger than the claim that the entities of a higher-level *science* should reduce to a lower-level science. They seem to make the claim that *all entities*, be they of common sense, science, or perhaps even philosophical theories should reduce to the sciences. This seems to me to be highly questionable, to say the least. Thus the natural question at this point is whether metaphysical naturalists aiming at a *metaphysically* naturalistic account of semantics take this to be a reduction *within* science, or from *without* science to within. In other words: is semantics a science or not? In the next chapter, I will aim to show that it is not. However, this diminishes neither its value as a means to understanding the world, nor its claim to making factual claims about it. What it does cast doubt on is

whether it will reduce to the ontology of the natural sciences, now or at any stage of inquiry.

Semantics and Naturalism (II): The Reduction Target

6.1 Introduction

In the previous chapter, I argued that we are not in a position to know what the resources of a proposed reduction base for semantics would be. This is primarily because there is no metaphysical fixity in our scientific categories: science is an incomplete enterprise at this point, and our fundamental categories may be subject to revision. In fact, the history of revision of scientific theories provides excellent empirical evidence for the view that even our best current scientific theories may well be false. It seems a fool's errand to try and construct a theory of phenomenon B using only the resources made available by a *false* theory of phenomenon A.

In this chapter, we move on to examine the reduction target: semantics. I will argue for the view that there is no scope for a comprehensive *science* of semantics as philosophers have construed the topic. Thus there is no hope for a naturalisation/reduction of semantics to scientific categories. I think it might be necessary to make a couple of preliminary points about this, one to do with the sense of 'science' in play here, and the second on the nature of reduction. Let's deal with 'science' first. In the sense in which I use the term here, 'science' means rather more than systematic empirical inquiry. Taxonomy, for example, would meet such a criterion for counting as science. By 'science', I mean something deeper and more explanatorily powerful than mere description of phenomena, however systematic.

Following Chomksy (1993, 2000a), let us characterise scientific inquiry (sometimes called 'naturalistic' inquiry, since the methodology to be outlined here is intended to characterise scientific inquiry into the 'natural' world) to involve the following:

1. Naturalistic inquiry into some phenomena consists in 'attempting to construct intelligible explanatory theories that provide insight and understanding of [such natural] phenomena'.

2. The phenomena that we attempt to investigate in this way 'are selected to advance the search into deeper principles'.

3. Such an approach 'will look forward to eventual integration with the core natural sciences', although whether such integration is possible will be a 'question of fact, not dogma' (Chomsky 1993: 41).

Stainton (2006: 914) emphasises the following points about this definition of naturalistic/scientific inquiry: first, taxonomic/descriptive projects do not qualify as scientific on this view unless they also contain an explanatory component. Secondly, the 'search for explanation is taken to imply...underlying realities, "deeper principles" that give rise to surface phenomena'. Finally, the aim of *integration* as stated in (3) should not be confused with 'success in reduction', for reasons that should be familiar from the previous chapter.

In fact, (1)-(3) above, when conjoined with the further claim that 'inquiry into the mind is an inquiry into certain aspects of the natural world' which we should investigate 'as we would any others' (i.e. in accordance with (1)-(3)), form the basic commitments of the 'methodological naturalism' that we introduced in the last chapter (Chomsky 1993: 41). Thus a further point about scientific/naturalistic inquiry as we shall characterise it here is that any object of study 'is a naturally occurring object, not an artificial construct' (Stainton 2006: 914). Thus a methodological naturalist holds something like the following claims to be true:

1. Language and mind are naturally occurring phenomena, and thus apt to be subjected to naturalistic inquiry.

2. Naturalistic inquiry consists in holding (1)-(3) above as methodological commitments. That is, we investigate natural phenomena by positing underlying principles that allow us to explain surface phenomena. Thus this variety of naturalism is purely methodological; it involves no particular *metaphysical* commitments.

3. There is no reason to abandon this methodology when inquiring into the 'mental': if mental/linguistic phenomena are natural phenomena, then we should investigate them as we would any other natural phenomena.

Now for preliminary point two on the nature of reduction. First, we should note that reduction is a process that goes on *within* science. One theory is

155

successfully reduced to another when two conditions are met (Rosenberg, 2005: 80-83):

1. The *significant generalisations* of the target theory are subsumed by the generalisations of the base.
2. The *ontology* of the target theory is preserved.

Condition (1) requires that the laws of the target theory may be derived from, or seen as more specific instances of the more general laws of the base. Condition (2) requires that the entities of the target theory be definable in terms of the ontology of the base theory; if the two theories are not trading in the same entities, then we do not have reduction at all, we have revision/replacement of one theory by another. The first condition on successful reduction makes it clear why there should be some emphasis, when characterising scientific inquiry, on the search for deeper principles/underlying realities as we have it above. For if what reduction consists in is the subsuming of specific laws/principles by *deeper* or *more general* laws/principles, then only such areas of inquiry as actually uncover/postulate explanatory laws/principles can enter into a successful reduction. The implication for semantics is clear: if semantics is to be reduced to scientific entities, laws, and so on, then semantic relata had better be able to enter into law-like generalisations. As we will see, it is doubtful that they can.

When it comes to reducing meaning/semantics to naturalistic categories, we must be careful to distinguish between certain common sense conceptions of meaning and certain technical conceptions of it. Consider first the idea of meaning from a common sense point of view. Surely there is at least one, and probably more, common sense conceptions of meaning, and some of these may well include the notion of some sort of relation between words and the things they are employed to talk about. However, there are two principal reasons why the common sense conceptions of meaning are not apt targets for reduction. First, common sense includes no *laws* of meaning/semantics. Thus in an important sense, there is nothing to be reduced here: there are no laws to capture. Secondly, there is no particular reason to think that the natural sciences and common sense share their ontology; in fact the opposite is probably true. As Chomsky (1993: 39) points out, ever since the introduction of Newton's 'occult force', gravity, science has departed from the strictures of common sense wherever it saw fit to do so, positing entities and properties that in no way reflect common sense

understanding. Thus the notion that science might provide the resources to 'reduce' the ontology of common sense is flawed for these two reasons: first, there are no laws of common sense conceptions of phenomena to capture, and, second, there is no reason to think that the ontology of the natural sciences will allow us to 'preserve' the ontology of common sense. In other words, there is no particular reason to think that every object of common sense understanding will play a role in scientific theorising, or even be possible to so much as characterise from a scientific point of view. (We will return to these points at length later.)

I hope these remarks are sufficient to at least partially make the case that reduction is a process that goes on within science and not a relation that obtains between the objects of scientific and common sense thought. In any case, the relationship between the natural sciences and common sense will occupy us at length in what follows. For now, I will simply stress the point that one does not 'reduce' common sense to science. One might reduce one portion of theory to another – perhaps a specific piece of physical theory to more general principles, or, more ambitiously, portions of chemical theory to physical theory.

Now, as I have indicated, if semantics is to be an apt domain for naturalisation, then semantics had better be a scientific endeavour. In what follows, I will argue that this simply is not the case. I restrict this conclusion to what I will call *philosophical semantics*: the study of the word-world relations that hold between expressions and the mind-independent things for which they stand. This does not preclude the possibility of other semantic projects, such as lexical semantics (roughly the study of internalist word-word relations), holding genuine scientific status. Indeed, I think that they do. The claim is that we cannot conceive of a genuine explanatory science that trades in word-world relations, which, of course, is a severe problem for semantic naturalisers, because it's *exactly* these relations that they want to naturalise. To put this point slightly differently: before the semantic naturalisers can get started on their project of capturing semantic generalisations (laws) in the terms of other, lower-level sciences, they must first establish the science of semantics in its own right. This they cannot do, for the reasons I will provide in the rest of this chapter.

6.2 The Target: word-world relations in semantics

In this section, I will make the target of my remarks absolutely explicit. Someone who endorses the general project of philosophical semantics may be characterised as committed to the conjunction of the following claims (here I partially follow Stainton 2006: 916-7):

1. There are public languages, which are the shared property of communities of speakers; this is the 'fundamental' sense of 'language' (Dummett, 1986: 473).

2. In key respects, particularly with regards to their semantic properties, public languages are like the formal languages invented by logicians. Thus:

    a. Names refer to individuals.

    b. Predicates have extensions; these are sets of objects to which the predicate applies.

    c. Sentences are associated with truth-conditions/Fregean senses/propositions. Whichever route one wishes to go down, the meaning of a (declarative) sentence is taken to be something truth-evaluable (relative to a set of contextual factors Cf. Kaplan).

3. It is thus the job of a compositional truth-conditional semantics to describe and explain the relations that hold between expressions of a public language and the worldly objects and states of affairs that they stand for; specifically between names and their bearers/predicates and their extensions/sentences and their truth-conditions.

For the semantic naturaliser, we may add:

4. The relation between expressions and the things they stand for must be cashed out in terms of the categories/entities/relations that appear in our best theories of the natural sciences.

There are several strands to my argument, all of which help to establish the general conclusion that semantics as philosophers conceive it is not a scientific endeavour and thus not an apt domain for naturalisation. In this section, I will present these in turn. First, we will consider two arguments from ontology (from Stainton 2006) which both attempt to undermine confidence in the claim that the entities that semantics deals in are robust enough to enter into scientific

theorising. The first of these, the Radical Argument from Ontology, questions whether public languages and words therein so much as exist. The second, the Moderate Argument from Ontology, applies to both sides of semantic relations. It grants that both sets of relata exist, but charges that the conditions for their individuation rule them out as objects of naturalistic inquiry. As we will see, the distinction between the perspectives afforded by our scientific and common sense ways of seeing the world is crucial to the Moderate Argument. We will also here consider a closely related argument (from Chomsky 1993; 2000a and developed in Bilgrami and Rovane 2007) to the effect that the perspectives provided by common sense are so complex as to prevent scientific generalisations about meaning and reference from 'ranging over all speakers of a natural language', or even over 'a single speaker at different times' (Bilgrami & Rovane 2007: 182).

We will then consider whether the study of the truth-conditions of natural language declaratives is a theoretically tractable domain. Exactly what it is to be theoretically tractable in the intended sense is a question that will occupy us at length throughout this part of the chapter. Roughly speaking, a theoretically tractable domain would exhibit the kind of systematicity required for law-like generalisations. In other words, one would have to be able to discover the (law-like) principles that underlie the assignments of truth-conditions to natural language sentences. I will argue that, while aspects of linguistic meaning may be systematic, linguistic meaning itself under-determines truth-conditions, which depend on a whole host of other factors not themselves systematic enough to be susceptible to naturalistic inquiry.

6.3 Arguments against the possibility of a science of semantics

6.3.1 Two Arguments from Ontology

Stainton (2006) presents two arguments from ontology both of which target the claim that the entities that semantics trades in are apt to enter into scientific theorising. The first of these, the Radical Argument from Ontology (RAO), questions whether the relata that semantics trades in even exist. The second, the Moderate Argument from Ontology (MAO), grants that words and the things they are used to refer to are real enough, but argues that many of the things

we refer to are not of the right kind to be described or explained from a scientific point of view. As we will see below, this argument trades in a distinction between the perspectives provided by scientific and common sense thinking about the world. The MAO charges that many of the 'things' to which we refer when adopting the common sense perspective on the world cannot enter into a scientific theory. There are a myriad of reasons for this claim, to be taken up below. To mention just one here: one might worry that there is no hope of providing an appropriate theoretical definition for many of the things to which people refer (see Fodor 1980, for a similar claim). There is an added interest to considering these arguments from ontology, as in doing so we will be able to cast further light on the intuition that meaning is normative.

As we have defined it, the project of philosophical semantics is to describe and explain the relations that hold between public language expressions and worldly items. Thus, as Stainton (2006: 917-8) points out, it is obvious that the relata must be able to stand in the relevant relations that semantics describes; but, equally obviously, the relata cannot do this if they are not real. However, that they are not real is precisely the conclusion of the RAO. The RAO as Stainton presents it is naturally read as targeting the 'public expressions' side of the semantic relations; however, one could equally well run the argument against the 'worldly objects' side. Consider the following, a possible axiom of a semantic theory for English:

1. 'Green' (in English) applies to x iff x is green.

The radical argument from ontology in essence presents us with the following challenge: how are we to individuate the public language 'English', and, thereby, to individuate *words* in English? Let us consider the public language question first. As Stainton points out, 'the way we divide up languages in common sense, and in much philosophical theorising, does not actually correspond to any robust divide' (Ibid: 918). For example, people may speak of Chinese as a single language even though the two main 'dialects' of Chinese, Mandarin and Cantonese, 'are not mutually intelligible'. Conversely, one speaks of Swedish and Danish as distinct languages notwithstanding the facts that 'they are far closer to mutual intelligibility than the "dialects of Cantonese"' (Ibid), and that, as Chomsky points out, 'the language of southern Sweden was once Danish but became Swedish a few years later, without changing, as a result of military

conquest' (1993: 20). As Stainton points out, the only 'semi-robust divide here is mutual intelligibility' – but this is of no help since public languages are not individuated this way anyway (Ibid).

We cannot help matters much by considering dialects, since 'what counts as a dialect is equally peculiar' (Ibid): one might talk of, say, British English and American English as single dialects, thereby glossing over massive variation between the ways English and American speakers use language (think of speakers from Glasgow, Birmingham and Boston). How we, as a matter of fact, *do* individuate languages/dialects is a complex matter. As Chomsky puts it:

> What are called 'languages' or 'dialects' in ordinary usage are complex amalgams determined by colours on maps, oceans, political institutions, and so on, with obscure normative-teleological aspects (Chomsky, 1993: 19).

Thus we slice 'languages' and 'dialects' in the way we do:

> ...not because of any robust linguistic divide, but because of colonial history, similar writing systems, shared canonical works of literature, present military might, arbitrary national boundaries, religious differences, and so on. That, and not 'nature's joints', is what makes it the case that people 'speak the same language/dialect' (Stainton, 2006: 918).

What, then, *does* naturalistic inquiry discover when it is turned on linguistic phenomena? Stainton (2006: 918) claims that we discover three things:

(i) Individual idiolects;

(ii) Sets of idiolects 'that share some non-obvious underlying parametric feature'; and

(iii) The universal human language faculty.

As Stainton points out, none of these 'corresponds even remotely' to the notion of 'shared public language' (Ibid).

In fact, it is worth pausing for a moment over what scientific inquiry into language has actually targeted as it will be important later on. What current linguistic theory targets is, in fact, both the universal language faculty of human beings, and also what Chomsky has termed 'I-language'. I-language is certainly related to idiolects and the language faculty, but is not exactly either of these, so Stainton's list is not quite accurate. Current linguistic theory has it that all human

beings have a faculty of language which has a genetically determined 'initial state'. The initial state takes various kinds of linguistic experience as input and yield as output an I-language (Chomsky 2000a: 4; 73). I-language is characterised as a function-in-intension that *strongly generates* linguistic expressions. To say that I-language strongly generates expressions is to say that it generates *structural descriptions* of symbol strings, and not simply the strings themselves. The 'I' in 'I-language' is thus intended to bring to mind the idea that language is *intensional* in this sense:

> On this view a language is not a set of symbol strings, but the generative procedure (or function computed) that maps from lexical items to structures that might be taken to underlie our linguistic performances (Collins 2008: 140).

The 'I' is furthermore intended to bring to mind the facts that I-languages are *internal* and *individual*. Thus the study of I-language is the study of the mental states of individual speakers. In other words, I-languages are states of the mind/brain (Smith 2004: 35). I-languages are individual in the sense that everyone's ways of speaking and understanding are, to some extent, individual and idiosyncratic. Here we see the relationship of I-languages to idiolects: both are individual to any particular speaker. However, it is potentially misleading to think of current linguistic theory as being concerned with idiolects if these are conceived of as something *external* to the speaker. In other words, we might think of I-language as the mental structure that underlies an individual speaker's linguistic performance, and we may think of that performance as her idiolect.

Finally, we may reflect on the tasks that linguistics has set itself. One is to construct a theory of the initial state of the language faculty. This state is believed to be sufficiently general to be able to yield as outputs *all* of the potential human I-languages. Thus, to use a variant of one of Chomsky's favoured examples, if I were to have grown up in Japan, the I-language that would have been generated by my language faculty would have been very different to the I-language that I, in fact, have (2000a: 4-5). The linguist's theory of the initial state of the language faculty is known as *universal grammar (UG)*. The linguist also, however, attempts to construct theories of individual I-languages. A theory of an I-language is known as a *grammar* for that I-language. That brings me to the end of my short digression into I-languages. We will return to the notion later on.

Now, returning to our main line of argument, one might argue (if one held, say, a fairly uncompromising commitment to scientism): since public languages cannot be individuated solely by recourse to 'nature's joints', there simply cannot be any public languages in nature. Thus public languages are something like a useful fiction; we use the concept of a public language to generalise, in a fairly loose way, over speakers whose idiolects are similar in certain respects.

Let us move to our immediate concern: the ontological status of semantic relata. Now, clearly if there aren't any such things as public languages, we are in real difficulty when we try to give any substance to the notion of a *word* in a public language. Consider the following example (similar to one by Stainton that we will consider below): take the words 'promise' as uttered in London, and 'promesse' as uttered in Paris. Now, what, if anything, makes it the case that we have two distinct words here (the English one and the French one), rather than slightly different pronunciations of the *same* word? As Stainton points out, one cannot say there are two words here on the grounds that we have an English word and a French word, if neither French nor English is 'objectively real' (2006: 919).

Let us take a different example to make a broader point. Stainton offers the following case: consider the following: 'fotoGRAFer' (said in Bombay), foTAHgrafer' (said in Toronto) and 'fotOgrafo' (said in Buenos Aires). Here again we want to say that there are two words, the English one and the Spanish one, but again appeals to the English and Spanish languages won't allow us to make this distinction if they aren't objectively real. Stainton makes the further point that any appeal to dialects is no help either since 'within a single country, or single part of a country, there can be "many different pronunciations of the same word"' (2006: 919). Thus one cannot say with much confidence that 'foTAHgrafer' is the same word as 'fotoGRAFer' in different dialects of English; after all, different pronunciations of the word 'photographer' may be found in Bombay and Toronto, and the pronunciations given here may appear in certain regions of Spain. Furthermore, Stainton notes, there is considerable variation between the ways that men/women/children/adolescents pronounce the 'same' words. Given all of this variation, 'there seems to be no good reason to count public words the way common sense wishes to: we can't put aside the differences on reasonable grounds' (Ibid).

Stainton also points out that we cannot individuate words by appeal to communities. Suppose, for example, that we tried to make the case that 'foTAHgrafer' and 'fotoGRAFer' were different pronunciations of the same word as shared by linguistic community A. Well, how do we individuate community A? A specific worry here is that 'it's not possible to individuate the right community except by appeal to shared language' (Ibid). Thus the community response is circular. Secondly, and more generally, one might object that 'communities are no more "robust" than languages turned out to be' (Ibid). That is, as it was with 'languages', the way we slice 'communities' doesn't reflect any objective natural facts; thus we cannot make appeal to such groups in scientific theorising – again, anyone who is strongly committed to scientism will claim that such things as communities simply fail to exist.

Stainton raises a further problem about the individuation of words which ties in nicely with our earlier discussions concerning the normativity of meaning. Even if we grant the notion of a linguistic community, we cannot simply identify, say, *English words* with *those words spoken by the community of English speakers*, since within that community there will be variation in meaning, error, over-generalisation of past tense rules, and so on. Consider:

1. He has an instinctive understanding of driving.
2. The response was odd hack.
3. I brang the book to school and we sang songs from it.

The use of 'instinctive' to mean something like *intuitive* can surely be found among the community of English speakers; slips of the tongue (such as 'odd hack' for 'ad hoc') are surely rife in the community, and it is typical of English children to over generalise linguistic rules and come up with regular past tense forms for irregular verbs (bring → brang, sing → sang). How do we exclude all of these utterances from constituting the English language? This is, I suppose, a restatement of the sceptical problem only at the community level.

There is another way of making the general point about public words/languages. Suppose that we treat public languages as essentially being generalisations over the particular idiolects of certain speakers, perhaps by making use of Putnam's idea of the division of linguistic labour. We then face this

general difficulty: which aspects of whose idiolects do we generalise from in order to characterise 'the language'? If there is no mind-independent 'thing' – the language – to which speakers stand in a relation, it looks as if our decisions on this point will be fundamentally arbitrary, at least on linguistic grounds. For example, the word 'theory' is notoriously understood in different ways by scientists and lay-people. As scientists use the term it means either 'an established sub-discipline in which there are widely accepted laws, methods, applications and foundations' or 'a body of explanatory hypotheses for which there is strong empirical support' (Rosenberg 2005: 69). We might treat the scientists' understanding of the word 'theory' as determining the semantic properties of the word 'theory'-in-English; however, this would have the consequences that (a) many speakers of English do not have the *English* word 'theory' as part of their idiolect, and (b) that philosophers in particular are not speaking English, or at least speaking it incorrectly, when they use the word 'theory'! Consider the phrase 'theory of meaning'. The scientist may complain against, say, Miller's (2007) glossing of this phrase ('an elucidation of our intuitive concept of meaning') that, by the definitions given above, this is simply not what 'theory' means-in-English. Physical theory is not 'an elucidation of out intuitive concept of the physical'. I'm merely being deliberately facetious here; what we would ordinarily say is that the word 'theory' has several senses. However, on the basis of what do we say this? In particular, on what grounds do we say that the lay-person's use of 'theory' to mean *hypothesis/uncorroborated speculation* is *incorrect, not part of English* or *not proper use of English*, while both the philosopher's and the scientist's use of 'theory' *are* proper uses of English?

In fact, at this point a version of the normativity of meaning re-enters the picture. Stainton argues that words such as 'instinctive' meaning *intuitive*, 'theory' meaning *speculation*, and 'brang' are not part of English not because of how people *do* speak (we may say, not because of what they *do mean* by their expressions); but because of certain norms that dictate how they *ought* to speak (2006: 920). Contrariwise, there are 'words-of-English' that are so seldom used as to be almost obsolete – there are no facts about what particular speakers do mean by them, but they *are* words-of-English nonetheless, even if 'almost no-one would recognise them as such' (Ibid). (Stainton gives the example of 'peavey': *an implement consisting of a wooden shaft with a metal point and a hinged hook near*

the end, used to handle logs: perhaps no-one now has 'peavey' as part of their idiolect, but it is, nonetheless, a word-of-English.) Applied to our example of 'theory' above, it might be claimed at this point that what rules in the philosopher's and scientist's use of 'theory', and rules out the laypersons, is that the former uses are correct – there are norms that determine that such uses are correct – and the latter uses are incorrect.

On closer inspection, I think it becomes clear that this version of the normativity thesis actually doesn't have a great deal to do with meaning itself. First of all, note that it is granted above that speakers may attach non-standard meanings to their expressions (*intuitive* to 'instinctive', for example). So there will be determinate facts-of-the-matter that, say, Jones means *intuitive* by 'instinctive' – but note now that *no normative implications concerning the 'correct' use of English follow from this whatsoever*. In other words, in no way does it follow from the fact that certain speakers (whom we would ordinarily class as English speakers) use the word in this way that English speakers 'ought' to do so. Furthermore, *ex hypothesi*, neither is there a pre-existing norm that makes Jones' meaning possible. That is, if we grant 'peavey' the status of word-in-English because there is some norm that rules it in, we rule *out* Jones' use of 'instinctive' on the grounds that there is *no* norm that rules it in. Thus no norms follow from Jones' meaning, nor is it necessary for there to be pre-existing norms for him to mean what he does.

Where, then, does the normativity under consideration issue from? I think the answer is reasonably clear: such norms issue from social- and authority relations among speakers. Essentially, experts in certain fields, pedagogues, lexicographers, and the like dictate that *these* are the proper or correct uses of certain words-in-English, and any other uses are 'incorrect'. While we may be able to individuate public languages such as English only in this way, note that, first, the norms in question do not arise on the basis of the meanings that arbitrary speakers attaches to their expressions. Rather, they only arise on the basis of the decrees of experts and authorities. Second, note that the phenomenon of a speaker meaning something by his or her expressions is *prior* to the norms. Thus the norms are not essentially to meaning itself: they neither issue from meaning, nor

make meaning possible. Rather, norms enter the picture when we try to make sense of the notion of 'public language'.

Returning to the RAO, it is quite possible to be somewhat sceptical of the existence of norm-individuated words-of-English; and certainly so if one is a naturalist in the methodological sense of inquiring into aspects of the natural world. As Stainton puts it:

> It's at least hard to see how there can really *be* such things [as norm-individuated words of public languages], to stand in objective relations with external objects, sets thereof, and so on (2006: 920).

If one is indeed a naturalist in the methodological sense, one might balk at the idea of uncovering empirical regularities or laws concerning words that are not a part of any speaker's idiolect but exist nonetheless because of norms, where the norms in question consist in essentially social relations among speakers.

So here, in a nutshell, is the gist of the RAO:

1. Semantics is the study of the relations that obtain between words-in-a-public-language and things-in-the-world, but:
2. On the basis that:
   a. There is no 'objective way to individuate/count words across or within a dialect'; and
   b. What 'makes something a word-in-a-public-language would have to appeal to "ought" rather than "is"' (Ibid: 920),

   There are no such things as words-in-a-public-language; *so,*
3. There can be no semantics if semantics is the study of word-world relations.

The criticism of the RAO that perhaps most immediately comes to mind is as follows: it is so obviously true that there are such things as public languages, and words therein, that we cannot discuss any argument to the effect that such things don't exist with a straight face. I don't think this criticism has any real bite. True enough, the conclusion of the RAO may offend common sense, but science is not bound by the shackles of common sense thinking about the world. As Chomsky (1993: 38-9) has noted, ever since Newton postulated gravitational

fields to account for the 'occult quality' of bodies of acting on one another at a distance without any direct contact, 'conformity to common sense understanding' has been 'put aside as a criterion for rational inquiry'. He goes on: 'if study of Newton's occult quality leads to postulation of curved space-time, so be it, however common sense may be offended thereby' (39).

Perhaps more pressing is the worry that certain phenomena, essentially centred on linguistic communication, would at least require a radically different explanation to what has been generally assumed in the philosophical literature if the RAO goes through. However, I have already given an indication as to how such a criticism may be met. In fact, Chomsky has argued explicitly that communication does not require or entail the existence of a shared public language:

> Successful communication between Peter and Mary does not entail the existence of shared meanings or shared pronunciations in a public language (or a common treasure of thoughts or articulations of them), and more than physical resemblance between Peter and Mary entails the existence of a public form that they share…It may be that when he listens to Mary speak, Peter proceeds by assuming that she is identical to him, modulo M, some array of modifications that he must work out. Sometimes the task is easy, sometimes hard, sometimes hopeless. To work out M, Peter will use any artifice available to him, though much of the process is doubtless automatic and unreflective. Having settled on M, Peter will, similarly, use any artifice available to him to construct a 'passing theory'[1] – even if M is null. Insofar as Peter succeeds in these tasks, he understands what Mary says as being what he means by his comparable expression (2000a: 30).

Collins (2008) has taken up this line of thought, arguing that the existence of independent public languages is neither necessary nor sufficient for communication. Such public languages are not necessary since 'communication clearly proceeds across "languages" or even in the absence of language'. They are not sufficient since, if they are independent of individual speakers, then 'a complex story must be told of how interlocutors orientate themselves in the same manner towards the language' (2008: 137).

It seems, then, that the various communicative phenomena that public languages are invoked to explain can be explained without recourse to such languages. Indeed, adopting the I-language picture explained above, we can see

---

[1] Cf. Davidson [1986] 2006.

those cases where communication is easiest as those where the I-languages of the interlocutors are most similar. Thus, in Chomsky's terms, the more similar the I-languages of the interlocutors, the more likely that M will tend towards being null. Furthermore, we can make sense of those cases where communication proceeds easily in the absence of a shared public language. Take the case of interlocutors on the Dutch/German border. While we characterise the languages that such interlocutors speak as 'Dutch' and 'German' respectively, the underlying reality is that the I-languages of such speakers are, in fact, much more similar to one another than they are to those of speakers from the Dutch and German capitals, say. Thus the challenge to account for communication given the conclusion of the RAO can perhaps be met.

A more telling criticism of the argument is this (here I follow Stainton 2006: 920-1): the argument argues that there is no such thing as a word-in-a-public-language *precisely* by raising questions to do with the pronunciation and individuation of words such as 'photographer' and 'intuitive'. But now, how can something that does not exist be said to so much as have one pronunciation, let alone several? Furthermore, the argument makes reference to obsolete or extremely rare words which, although seldom (maybe never) used still count as words-of-English because of norms (perhaps norms that issue from the pronouncements of lexicographers). Surely such things have to *exist* if the argument is going to exploit them – but if so the conclusion of the argument is paradoxical.

There are two ways of responding to this criticism. The first is to concede the point that public languages and words therein exist, but to claim that there could be no genuine '*science* that treated of them' (Ibid: 921). This is essentially the strategy of the MAO, to which we will turn presently. However, the radical line can also be pushed, and to this end Stainton (Ibid: 922) raises a couple of counter arguments that defend the RAO, one of which I will deal with briefly here. One might argue that just because the RAO makes reference to the words whose existence it denies, this does not concede the existence of the words. One can talk about all kinds of non-existent things – the present King of France, unicorns, dragons, and Harry Potter. As Smith (2004: 152) puts it, 'we can talk about [public] languages as we can talk about unicorns, but there are few treatises

of the anatomy of the unicorn, and equally few theories of [public] languages; and for the same underlying reason – neither exists'.

However, I'm not particularly interested in pursuing the RAO much further for several reasons. First, I think its conclusion is too strong. As Stainton points out, if semantic relata do not so much as exist, as the argument claims, then not only can there be no comprehensive science of philosophical semantics, there can be no truths *of any kind* about the meanings of public language words (2006: 923). On the contrary, I think there are perfectly good ways in which we can think and talk about the meanings of expressions in, say, French or English, and it is hard to imagine how we could pursue certain practical projects (learning to speak a second language being chief among them) without the notion of meaning-in-a-public-language. The second reason why I don't wish to further pursue the RAO is that it seems to me to set the bar for existence far too high (cf. Collins 2009: 56). It seems to me that one should not restrict one's ontology to that of the natural sciences: if one were to do so, it would not just be public languages that were in trouble, but a whole host of other phenomena that fall outside of the purview of such sciences. Furthermore, given its origin in Chomsky's writings, I think the RAO misrepresents the thrust of Chomsky's original arguments to the effect that semantics falls outside a naturalistic approach. These are offered not in a metaphysical, but rather in a *methodological* spirit.[2] Thirdly, the RAO targets *public languages*; thus even if its conclusion goes through, it remains open that we could have a scientific semantics for idiolects. (On a related note, I have stressed throughout this thesis that the sceptical challenge is to account for speaker- or idiolect-meaning. If this could be naturalised then it seems that the sceptical challenge would be met without our having to worry about the scientific status of public languages.) Finally, the conclusion of either argument from ontology is sufficient for my purposes, that is, to show that the outlook for a science of semantics is bleak. In fact, I think the MAO provides much stronger resources for securing such a conclusion. Let us turn to it now.

The MAO is probably best viewed not as a linear argument, but as a cluster of claims about the entities involved in semantic relata, and the nature of scientific theorising. Unlike the RAO, the MAO grants that both sides of the

---

[2] Cf. Collins 2008: 144; 2009: 56 f.; 2010: 47 f.

semantic relations do, in some sense, exist. However, it charges that such entities as one finds on either side of a semantic relation are simply not the kinds of entities that are apt to enter into scientific theorising. Thus the argument concerns the ontological *character* of the semantic relata. Although Stainton does not restrict the argument to one class of semantic relata, as he runs it, it is most naturally read as an attack on the scientific credentials of the 'things-in-the-world' side of the semantic relations (although, as we will see, there are ample grounds for running the argument against the 'words-in-public-language' side also). As the problems it identifies for a science of semantics are not restricted to the 'words-in-a-public-language' side of semantic relations, one advantage of the MAO over the RAO is that the MAO *does* tell against the possibility of a science of semantics for idiolects. I will present the argument as Stainton runs it, but I will also bring in further considerations that essentially point to the same conclusion. Let us turn to the 'things-in-the-world' side first.

In order to assess whether the semantic relata are apt to enter into scientific theorising, one must bear in mind the commitments of the methodological naturalist, the criteria for successful reduction, and the criteria for naturalistic inquiry as outlined in §6.1. Recall, therefore, that naturalistic inquiry is the search into deeper principles (laws) that underlie surface phenomena. Thus the entities that any particular science trades in must be capable of entering into law-like generalisations. To put the point another way, there must be laws concerning the entities of the domain of inquiry, and it is precisely the aim of the inquiry to unearth what these laws are. Reductionism, in turn, requires that the laws of the target domain be accountable for in terms of the laws of the base domain, and that the entities of the target domain must be preserved.

In Stainton's version of the argument, crucial to the MAO is the distinction between science and common sense. He argues that the methodological naturalist thinks that we can make a distinction between the world as it is revealed by science and as it is revealed by common sense *cognitively*, 'in terms of the kinds of concepts deployed' (2006: 923). Importantly for our purposes, the concepts of common sense are distinguished by their 'rich and complex internal structure' that 'eschews elegance in favour of day-to-day practicality', having 'inherently built in implicit references to human hierarchies,

rights/obligations, and our intentional states, rather than aiming for an objective description of the world independent from us' (Ibid).[3] (As we will see later on, the complex structure of the concepts of common sense is one reason why the prospects for a science of semantics are not good.) By contrast to common sense concepts, scientific concepts are 'hard-won achievements of many years of collective labour', austere in content, and deliberately decoupled from human interests and concerns – 'the whole idea of a scientific concept is to how us how things "really are" independently of us' (Ibid: 924).

Stainton points out that both science and common sense provide us with a *perspective* from which to view the world. Importantly, we may be unable to 'see' certain aspects of reality from each perspective. So, for example, the scientific perspective reveals to us such things as quarks, genomes, tectonic plates, and the like, as well as the mind/brain 'at various levels of abstraction' (2006: 924). However, there are plenty of things that cannot be seen from this perspective. Among them are 'normative categories', such as 'good wine, liveable cities, and well-prepared osso bucco' as well as 'non-normative yet mind dependent things like clouds, tea, desks, sunsets, breakfast cereal, and hockey scores' (Ibid). While blind to the concepts of science, the common sense perspective reveals to us those objects of our day-to-day existence, such as tables and chairs, as well as our general folk notions of psychology, for example. In general, the distinction between the two perspectives is that common sense reveals to us those entities 'whose individuation conditions inherently involve complex human interests and purposes', whereas the scientific perspective is blind to such entities (Ibid: 925).

Turning now to the project of philosophical semantics, it is important to bear in mind that the things-in-the-world side of semantic relations include the putative referents of concepts of both kinds (i.e. the referents of both scientific and common sense concepts). In fact, the domain of semantics involves *everything* that can be talked about. Thus the worry raised by the MAO is this: the kinds of entities seen by common sense are 'so highly dependent on human

---

[3] Added to these features are the claims that common sense concepts are 'part of our biological endowment', 'constructed from innately given semantic features'. I am not sure that these latter claims are necessary to effect the distinction, and doubtless they will be controversial, so I will gloss over them here.

perspectives and interests' that they 'cannot be seen by the peculiar instrument that is natural science' (Ibid).

We can take a favoured example of Chomsky's to make the point vivid. Consider the word 'London'. Suppose we have a semantic theory that includes as a principle something like the following axiom:

'London' refers to London.

Our question is: what is the thing London to which 'London' refers? Chomsky writes:

> We can regard London with or without regard to its population: from one point of view, it is the same city if its people desert it; from another we can say that London came to have a harsher feel to it through the Thatcher years, a comment on how people act and live. Referring to London, we can be talking about a location, people who sometimes live there, the air above (but not too high), buildings, institutions, etc., in various combinations. A single occurrence of the term can serve all these uses simultaneously, as when I say that London is so unhappy, ugly, and polluted that it should be destroyed and rebuilt 100 miles away (Chomsky, 1993: 23).

Chomsky goes on to conclude that 'no object in the world could have this collection of properties'. Note that this isn't a result of a commitment to any particular ontology. Rather, the point is that the properties we attribute to an entity such as London are too disparate to hold of a single entity. From one point of view, London is the same entity even if none of the same people live there (as in the case where London is burnt down and rebuilt), from another point of view we might individuate London by reference to its population. The point is that a single entity cannot at once be individuated by its population and *not* individuated by its population; nor can it at once be a certain population and a certain area of air above that population.

As Stainton points out, nor is what Chomsky says to be taken to suggest that London fails to exist. Rather, it is to say that what *makes it the case that* London exists isn't a simple matter of there being some object in the world (London) that can be identified and individuated scientifically – that is, in the absence of human concerns and interests. As a generalization of the point, Stainton considers such words as 'bargain', 'ownership', 'tenure', 'nearby' and

'polite': we seem to be at a lost to say what the external correlates of such expressions would be, and to provide an account of how natural science could hope to 'see' such things.

I pause here to ward off another potential misunderstanding. It is not that only, for example, physical stuff that can be individuated by reference to its spatio-temporal position that can be seen by the scientific perspective. As we have noted, the mind/brain at various levels of abstraction can be seen from such a perspective also. The issue really has to do with what can be seen in the absence of those peculiar human perspectives and concerns that allow us to talk about, say, 'London' in the way that we do. The point is that:

> In general, being objective and ignoring interest relative distinctions, the 'scientific perspective' cannot see entities whose individuation conditions inherently involve complex human interests and purposes (Stainton 2006: 923-4).

Returning to the main thrust of the MAO, a second concern is this: if certain entities cannot be seen from the scientific perspective, how can they enter into the kinds of generalisations that science aims to uncover? That is, if science is blind to such things as pencils, then there can be no laws about pencils *qua* pencils (cf. Fodor 1980: 71). In order for the naturalising project in semantics to work, we would need to find descriptions of everything within the domain of semantics (i.e. everything) that would be capable of entering into scientific (law-like) generalisations (Ibid). It looks like this project is doomed to failure due to the peculiar perspectives of science and common sense.

The implication of all of this for the project of philosophical semantics, considered as a science, is obvious: if the semantic relata cannot be 'seen' from the perspective of natural science, and therefore cannot enter into scientific generalisations, then there is no hope for a reduction of semantics to scientific categories. There are two principal reasons for this. Firstly, there will simply be no scientific semantic generalisations to reduce. Secondly, there can be no hope of preserving the ontology of semantics (that is, the ontology of both science and common sense), if science is essentially 'blind' to the ontology of common sense. Thus neither of the conditions that are a prerequisite for successful reduction can be met.

I pause here to deal with an important objection to the MAO. The MAO relies on a distinction between scientific and common sense concepts, but it is not clear that we have a difference in kind here. Rather, the difference may only be a difference of *degree*. This objection is important as it makes it difficult to see how the concepts of common sense thought and talk about the world could be somehow in principle unsuitable for scientific use.[4] However, Stainton (2006: 927) offers a three part rebuttal of this objection as follows. First, the MAO does not require an 'exclusive and exhaustive' distinction between what science and common sense can 'see' in order to secure the conclusion that there cannot be a *comprehensive* science of semantics. Such a conclusion can be secured so long as there are a lot of things that science cannot see at the extreme common sense end of our continuum. Secondly, Stainton notes that the science of semantics would have to be able to 'see' *every* object that can be talked about. Thus the MAO hinges, not on the claim that *science cannot see common sense objects*, but rather on the claim that *there is no* single *science that can see every common sense object*. A comprehensive science of semantics – that is, 'a comprehensive science of language that described world-word relations' – would need to be just such a science. Stainton's third reply is that the objection to the MAO gains much of its plausibility from loose terminology. Note again the special sense of 'science' in play here: science involves the search for explanatory insight and the postulation of deeper principles that give rise to surface phenomena. Systematic description is not enough. Thus while it may be plausible that 'some systematic empirical inquiry' can see each common sense object, it does not follow from this that a genuine explanatory *science* can do so (2006: 927-8).[5]

---

[4] My thanks to Bob Hale for this point.

[5] There is another point that one might raise against the above objection to the MAO which has to do with loose terminology. This is that we should also be sceptical of the idea that graphological or phonological similarities between scientific and common sense terms amount to the identity of the concept being deployed when those terms are used. For instance, when Chomsky uses the expression 'language', this sometimes means specifically *I-language*, a specifically human cognitive capacity. The fact that the concept he is deploying is associated with a graphological form that is shared with the common sense conception of language is neither here nor there. Similar remarks apply for terminology of other sciences. Thus, for example, in everyday parlance people use the term 'water' to refer to 'a vast range of solutions containing not only H_2O'; it is when the term is deployed in technical discourse about chemistry that it takes on the more restricted meaning H_2O (Isac & Reiss 2008: 40). (In fact Chomsky (2000a: 151) points out (quoting Jay Atlas) that this is a simplification: nuclear engineers now distinguish between *light water* and *heavy water*. Only the latter is H_2O.)

Why is it important to make the distinction between science and other forms of systematic empirical inquiry in the way that I have been making it? After all, my way of slicing things has the result that many areas of inquiry that we might naturally think of as scientific do not count as genuine science. The answer has to do with the avowed aims of naturalistic inquiry as I outlined them above. Recall that a genuine science, as I use the term here, involves 'the aim of integration with other "core natural sciences"' (Stainton 2006: 934). Clearly some areas of inquiry may be empirical and systematic, and yet not be amenable to integration with the core natural sciences – for instance, if they uncover no laws or principles. Now this is of crucial importance for the semantic naturaliser, for it is exactly this kind of integration that she envisages for semantics. However, the MAO is aimed at showing that semantics simply is not amenable to this kind of integration. (This is not to deny, of course, that it is a systematic and empirical discipline.)

There is a related Chomskyan argument for the conclusion that a scientific semantic theory is unlikely to be forthcoming to consider here.[6] This stresses the fact that reference is always mediated by a speaker's conceptions of the world: reference to things is 'embedded in the context of various conceptions' of objects to which speakers intend to refer (Bilgrami & Rovane 2007: 187). We noted above that the concepts of common sense are quite complex. As Chomsky notes, in general such concepts provide 'a certain range of perspectives for viewing what we take to be things in the world' (2000a: 36). Recall Chomsky's discussion of 'London': we can regard London with or without regard to its population or its physical location, or as a certain area, or even as the air above; or as some combination of these (see also Chomsky 2000a: 37). Chomsky sums this up by saying that lexical items such as city names provide 'intricate modes of reference' (2000a: 37). The problem we raised above was that there 'neither are, nor are there believed to be things-in-the-world' with all of the properties encapsulated by the modes of reference of such lexical items as city names (Ibid). However, here I want to raise a different problem.

The scientist of semantics seeks to uncover universal laws or generalisations about meaning and reference. However, as Bilgrami and Rovane

---

[6] See Chomsky 1993; 2000a and Bilgrami and Rovane 2007.

point out, since the 'extremely rich and diverse' conceptions we have of the things to which we refer (such as London) 'infect reference itself' making it such a 'mediated and contextual notion', the whole project of uncovering such 'scientific generalisations about reference' is thwarted. This is because such generalisations fail to range over all speakers at a time and even over a single speaker at different times (2007: 182).

Let me explain further. We have already seen that a single lexical item may provide a range of ways of thinking about the world. Thus, for example, there is a range of possibilities concerning exactly what a person is referring to when they use the expression 'London': it may be that what Jones refers to on Tuesday when using the expression 'London' is the population of London, as in 'London always seems so impersonal and unfriendly to me'. However, what Jones refers to on Friday when he says 'London is terribly polluted' could be the air above London, or perhaps even the water supply. Thus any principle such as '"London" refers to x' will fail to generalise over time-slices of Jones.

Let us consider some further cases. By using the term 'water', Jones may refer to H_2O, or to both heavy and light water, or to just heavy water (H_2O again), or to water qua H_2O in liquid state (thus excluding ice and vapour), or to H_2O with various impurities (such that, for instance, 'water' may be distinguished from 'de-ionised water'). If Jones says 'the water one gets out of the tap in Portsmouth tastes very different to the water that one gets out of the tap in Sheffield', by his term 'water', he means H_2O with two different sets of impurities (it would make no sense to say that $pure$ H_2O tastes different in two different places). To take different example (from Chomsky 2000a: 128), consider the case when a tea bag is dipped into a cup of hot water. We might refer to what is in the cup as 'tea', even though its molecular structure is indistinguishable from H_2O. Contrast the case where a shipment of tea bags is accidentally dumped into the local reservoir: we would describe what comes out of the tap in this case as 'contaminated water', even though it is physically indistinguishable from what we call 'tea'.

Another favoured example of Chomsky's is 'book' (1993: 23). The word 'book' may be used to refer to the *contents* of the book, such that Smith and Jones can take out the *same book* from the library, but later switch copies by accident; or

to the physical thing itself ('the book is two hundred pages long, but missing pages 93 and 94'); or to some physical properties of the book ('the book is red with white writing' referring only to the cover of the book[7]); or to the abstract idea of the book ('Jones wrote the book in his head, then forgot about it'); or to some combination of these. In Bilgrami and Rovane's phrase, it is a highly mediated and contextual matter as to which of these potential referents will be the actual referent of some use of the expression 'book' by Jones.

In fact, matters are even more complicated than this. As Bilgrami (1993) points out, one's conceptions of a thing are naturally affected by the beliefs and knowledge that one has about it. Thus, for example, the chemically knowledgeable will likely have different conceptions of water to the chemically ignorant. However, if reference is infected by the conceptions of things that one has, and these in turn are affected by one's knowledge and beliefs about the world, and no two people have exactly the same knowledge and beliefs, then it becomes hard to see how there could be any true scientific generalisations over what different speakers mean and refer to when they use expressions.[8] But now, if there are no such generalisations to be had, then there is no hope of reducing meaning and reference to the core natural sciences: given the lack of true generalisations about meaning and reference, in a very important sense there is nothing here to be reduced.

Let us now return to the MAO and the other side of semantic relations: expressions in public languages. We noted above that there are simply no laws that hold of pencils and other objects of common sense as such. Smith makes exactly the same point regarding public languages and words therein:

> Crucially, there are no laws or principles that hold of 'English' or 'Chinese' or 'Dutch' as social or national constructs, any more than there are principles that

---

[7] Cf. Carston 2002: 23. See the discussion of part-dependent adjectives below.

[8] Fodor (1987; 1990) has raised a similar concern about the possibility of there being true scientific generalisations over *narrow* mental content, where such content is cashed out in terms of functional/conceptual/inferential role. Fodor thinks that, absent a defensible analytic/synthetic distinction, conceptual role semantics leads inexorably to *holism* about meaning and content (1990: 52). The trouble is that, once such holism is accepted, we will be hard pushed to find two agents (or even two time slices of the same agent) whose mental representations share narrow contents. Thus generalisations over mental content will, in fact, fail to generalise, rendering a science of such contents impossible (Fodor 1987: 57). See Bilgrami 1998 and Bilgrami & Rovane 2007 for discussion.

hold of the Dutch visual system; but there are (grammatical) principles that hold of the linguistic knowledge of each individual (2004: 151).

In fact, Smith here is making explicit Chomsky's reasoning for focusing his linguistic inquiries on the minds of individuals – that is, on I-language – rather than on the external, social constructs we talk about in ordinary usage. If a domain is to be scientifically tractable or amenable to naturalistic inquiry, then the entities that the domain trades in *must* enter into law-like generalisations. Entities such as public languages and expressions therein, individuated according to various complex human concerns that have little to do with the actual linguistic competence of speakers, simply fail to do so. Thus even if such a construct as a public language exists, neither it nor the words within it are amenable to scientific inquiry.

Suppose one were to argue in response that the English language is indeed susceptible of scientific study as such study may proceed *via* the study of the speakers of that language. There are two problems with this view. First, for the reasons discussed above, it is remarkably difficult to identify the community of 'speakers of English'. As we have seen, the grounds on which the English Language is individuated are *not* purely linguistic; so it would be a mistake to include all speakers that would be informally classified as 'speakers of English' as constituting the 'linguistic community' at issue here. Secondly, we will find considerable *variation* among speakers – indeed, we may end up with contradictory principles. For example,

> 'Disinterested' means (in English) *unbiased*.
> 'Disinterested' means (in English) *uninterested*.

Since both meanings are used by 'speakers of English', we cannot rule one of these principles out on the grounds that only one or the other is used by the relevant linguistic community. The second problem is that the postulation of a public language in fact seems redundant in animating the inquiries that linguists engage in.[9]

---

[9] For reasons of space I cannot go into the arguments here. These are developed in Collins 2008: 141f.; 2009; 2010.

Of course, the semantic naturaliser may accept the conclusion of the MAO with regards to public languages, but press for a scientific semantics for I-languages or idiolects (or, perhaps, for mentalese (re. Fodor 1990)). However, the difficulty for such a project resides in the other side of the relation: the things-in-the-world for which the 'words' of idiolects or mentalese are said to stand. All of the problems surveyed for a science of such relations above recur at the level of idiolects. Thus while there may be laws that hold of the linguistic knowledge of each individual, these will *not* involve relations between the expressions of that individual's idiolect and things-in-the-world.

We may, however, be able to uncover laws that involve syntactic relations, or semantic relations *between expressions*. Importantly, as Bezuidenhout (2006) points out, it seems that adopting the 'public language' picture does not help us to account for these. Consider, for example, some cases of referential dependence:

John wants to meet Tom and like him.
John wants to meet Tom and like himself.

In the first sentence, we know that 'him' refers to Tom, whereas in the second sentence, we know that 'himself' *cannot* refer to Tom. Contrastingly, in 'John wondered who set him up', 'him' *may* refer to John, but equally it may refer to some other (unmentioned) man. How are we to account for such facts? Bezuidenhout (2006: 132) argues that there 'does not seem to be a purely social explanation' for such linguistic patterns in terms of 'social norms, customs or practices':

> Following such patterns in one's use of language is not like following the rules of the road or other such social conventions. Learning a language is not like learning the rules of the road, and we couldn't decide to change the way we speak in the same way that we could decide to start driving on the opposite side of the road…or decide to start using a metric system of measures (Ibid).

If one were to insist that it is just a matter of, say, social convention that we interpret the pronouns in the way that we do, then, as Chomsky (in Antony & Hornstein 2003: 311) points out, we deprive the notion of 'convention' of 'connotations of agreement, compact, intention to conform' and so on. In fact, it seems to me highly unlikely that many speakers are so much as *aware* of the

principles that govern the interpretation of pronouns in cases such as the above. Indeed, figuring out such principles is a major task for linguistics. It's not as if we *consciously agree* to interpret the pronouns in such cases in the way that we do. Furthermore, Chomsky argues, simply describing such linguistic patterns as 'conventional' seems only to serve to 'terminate inquiry where it ought to begin' (Ibid: 312). That is, if we want to know *why* we interpret pronouns in the peculiar ways that we do, we gain no explanatory purchase by labelling such patterns as 'conventional'.

6.3.2 Problems of Tractability – Linguistic Meaning and What is Said

There is a second cluster of arguments to consider here, all of which aim to demonstrate that the key notions of projects of philosophical semantics – essentially truth and communication – are altogether too unsystematic, complex and unwieldy to submit to scientific theorising. (Of course, one might object that complexity is not in and of itself a barrier to scientific theorising. So I owe an explanation of the kind of complexity at issue here. This will emerge as we proceed.) I think the picture that emerges out of this way of thinking is something like the following: linguistic expressions have certain semantic properties. The semantic properties of complex expressions are determined by those of the simpler expressions out of which they are composed plus syntax. So far, so familiar. However, the semantic properties referred to here fall short of correctness-conditions/truth and falsity. The word 'semantic' is being used more generally to mean something like the 'meaning-properties' of the expressions. What an expression is used to mean and refer to, and whether an utterance of a declarative sentence is true or false is not just a matter of its semantic properties, but a massive interaction effect involving the semantic properties of the expression and a whole host of other factors, themselves perhaps not capable of entering into a scientific theory.

In order for this line of argument to be successful, I need to show that there are semantic properties of expressions that are linguistically encoded and systematic yet fall short of truth-conditional content. This flies in the face of much philosophical opinion which has it that meanings simply *are* truth conditions. However, I think that there are compelling arguments that demonstrate that

linguistic meaning is distinct from truth-conditions. Implicit in this way of thinking is something like the following hierarchy of what we might call levels of meaning (Carston, 2002: 19):

1. The linguistic meaning of an expression/sentence.

2. 'What is said' by (an utterance of) a sentence. This is typically understood as propositional/truth-conditional, and thus this is level at which philosophical semantics is taken to operate.

3. 'What is meant' by a speaker who utters a sentence in a particular context – a matter of their communicative intentions.

Consider the following example as a way to motivate the distinction between (1) and (2). Suppose we are unpacking shopping and you turn to me with a bottle of ketchup in your hand and ask: 'Where should I put this?' Suppose that I answer: 'Put it on the top shelf in the cupboard', pointing at a specific cupboard. Clearly the indexical expression 'it' does not *linguistically mean* 'the bottle of ketchup that you have in your hand'. However, the proposition that I express is something like *put the bottle of ketchup that you have in your hand on the top shelf of the cupboard that I am indicating to you*. Cases involving indexical expressions thus serve to motivate the distinction between (1) and (2). I will argue shortly that the distinction holds in other non-indexical-involving cases as well.

Now there are some time worn examples which show that (2) underdetermines (3). Thus, for example, when Professor A who is asked about the philosophical acumen of student B by Professor C remarks 'Student B has very nice handwriting', Professor A *says* that *student B has very nice handwriting*, but *means* that *student B is not a very good philosopher*. There is nothing here in the truth-conditional content of 'what is said' that determines 'what is meant'.[10] The more interesting thesis for our purposes, and one that will be defended here, is that (1) underdetermines (2), and not just for the obvious cases such as those involving

---

[10] There is even a question as to whether these categories for analysing meaning are fine-grained enough. Suppose someone asks me at a party whether we need more beer, and I say 'Well, every bottle is empty' (the example is from Stanley and Szabó 2000: 219-20). Now what precisely do I *say* in such a case? Stanley and Szabó's own view is that quantifiers are associated with contextually-completed 'domain variables' such that domain-restricted truth-conditions for quantified sentences are delivered when a quantified utterance is made in context. (For critcism of this kind of move, see Piertroski 2003.) Alternatively, one might deny the existence of such variables and think that I *say* that every bottle of any substance anywhere is devoid of content. This opens a gap between what I say and the proposition I actually communicate in such a case (something like *every bottle of beer at this house party is devoid of drinkable content*). However, I also imply that *yes, we need more beer*. We will return to examples of this sort presently.

indexical expressions. I will argue for this view shortly. However, the fact that (1) underdetermines (2) would not, by itself, be sufficient to undermine confidence in a science of semantics. For it might be that the factors *other than* linguistic meaning that play a role in determining 'what is said' are systematic and uniform, and thus theoretically tractable. Thus the claim that semantics, understood as the study of such things as reference/truth-conditional content ('what is said'), is not likely to submit to scientific theorising may be understood as a commitment to the conjunction of the following two claims:

1. Linguistic meaning underdetermines what is said.
2. What *does* determine what is said is too complex, disparate and messy to be systematically explained by a scientific *theory*.

I now turn to arguing for these claims.

How might linguistic meaning undetermine what is said (the truth-conditional content of an utterance)? Carston (2002: 21-28) gives the following taxonomy:

1. Multiple encodings.
2. Indexical references.
3. Missing constituents.
4. Unspecified scope of elements.
5. Under-specificity of encoded conceptual content.
6. Over-specificity of encoded conceptual content.

Let us consider these in turn. *Multiple encodings* are cases where one linguistic form carries more than one meaning. To use some time worn examples, consider:

(1)   Visiting relatives can be boring.

(2)   The duck is ready to eat.

(3)   I went down to the bank this afternoon.

Each of these sentences is ambiguous, but the reasons why are not always the same. In (1) we have a syntactic ambiguity, the sentence could either be read as *it can be boring to visit relatives* or *relatives who come to visit can be boring*. In (2) and (3) we have lexical ambiguities: 'ready' and 'bank'. Due to the multiple propositions that *could* be encoded by forms such as these, it is not obvious that

the linguistic meanings of the expressions *by themselves* determine which proposition is intended in such cases.

Indexical expressions will surely be familiar to any philosopher of language. If only for completeness, then, note that the linguistic meanings of the following sentences fail to supply something truth-evaluable:

(4)   She went there.

(5)   Put it in here.

(6)   The doctor will see you now.

Assignments of place, person and time are necessary in such cases to determine a truth-evaluable proposition. Thus for an utterance involving an indexical expression to be truth evaluable, pragmatic processes that fix such assignments are necessary. *Missing constituent* cases are those where some element necessary for the determination of a full proposition is missing from the linguistically encoded meaning of an expression. As Carston notes, such expressions fail to determine a proposition *even after* disambiguation and indexical assignment has taken place (2002: 22). Consider (the examples are adapted from Carston (Ibid):

(7)   Ibuprofen is better.   [Than what? For what?]

(8)   It's different.   [To what?]

(9)   She's departing.   [From where?]

(10)   He's too old.   [For what?]

(11)   It's snowing.   [Where?]

Although (7)-(11) are all partial phrases, Carston notes that the phenomenon of missing constituents can be present in full sentences as well. Consider (examples from Carston 2002: 23):

(12)   Bob is well groomed.

(13)   This fruit is green.

(14)   That is difficult.

(15)   It is serviceable.

Surely, one might think, sentences (12) through (15) do manage determine a proposition. Carston (2002: 23-4) demonstrates that this is not, in fact, the case. Firstly, the adjective in (12) is, to use Gross' phrase, 'scalar'. In other words 'it allows for comparison between things with respect to the degree with which they have the property concerned' (Carston, 2002: 23). Thus Bob may be well groomed for Bob (i.e. compared to his usual appearance), but not for someone else – an impeccable city banker, say; or Bob may be well groomed for your average philosophy student but not for your average barrister (Ibid). In (13) the adjective is 'part dependent', 'on different occasions of use it may apply to different parts or aspects of the thing it is being predicated of' (Ibid). Thus 'green' may apply to the peel or the flesh of the fruit, depending on the context. Consider also 'the book is green'. I may be referring in such a usage to the *spine* of the book – but not the writing on it – (say if I've asked you to get it down from a shelf), or the cover of the book (say it's lying in a loose pile somewhere) or, more unusually, the pages of the book (Ibid: 54).

Turning now to *unspecified scope of elements*, we may note that many natural language sentences are ambiguous with respect to the scope of certain elements (typically quantifiers and negations) within them. For example, consider the following (adapted from Carston (2002: 24)):

(16)   Everyone isn't ready.
(17)   She didn't park the car outside Tesco at 10pm.

As Carston points out, the negations in (16) and (17) are ambiguous with regards to their scope. (16) could be understood as:

(16a) No-one is ready.
(16b) Not everyone is ready.

Which reading is 'correct' in a certain context of utterance depends on whether the 'negation takes scope over the universal quantifier or *vice versa*' (2002: 24). Similarly, (17) is ambiguous to exactly what is being denied: the act of parking the car, the location of the parking, the time of the parking, or some combination of these.

Similar phenomena may be observed in the following:

(18)   You can please some people all of the time.

(19)   Why do you think Lord Sugar doesn't like Jim?

(18) can be understood either as specifying a specific person/group of people who is/are such that, whatever the time, you can please them, or as stating that, at any time, one can please someone or other. This isn't, strictly speaking, a scope ambiguity, rather, it is unclear whether or not (18) involves a uniqueness claim regarding the existential quantifier. In fact, even this isn't quite right because (18) is open to it being a unique person, or a unique *group* of people, who can always be pleased. (19) is ambiguous with regard to exactly what is being questioned. (I'm unsure as to whether this counts, strictly speaking, as a scope ambiguity, but it seems natural to include it in the discussion here.) We could understand the questioner in (19) to be wondering what their interlocutor's reasons are for thinking that Lord Sugar doesn't like Jim, or what their interlocutor thinks Lord Sugar's reasons are for not liking Jim.

Under-specificity of encoded conceptual content refers to cases where some process of *adding to* or *enriching* the conceptual content provided by linguistic meaning seems necessary in order to arrive at a fair representation of 'what is said'. Interestingly, in such cases, even though linguistic meaning may provide a fully propositional content, it is unclear that this content is what the speaker would ordinarily be understood as having 'said'. Consider the following (adapted from Carston 2002: 26):

(20)   Learning to play the piano takes time.

(21)   London is some distance from Sheffield.

(22)   Something has happened.

(23)   I haven't eaten breakfast.

(24)   I haven't eaten scallops.

(25)   There's nothing in the fridge.

While (20), (21) and (22) do determine truth-evaluable contents, these are mind-numbingly dull: every activity takes some period of time to complete, everywhere is some distance from everywhere else, and *something* or other has always happened given some period of time. As Carston points out, these 'dull truisms' are 'virtually never what a speaker has intended to express' (2002: 26). She notes that therefore some pragmatic process of 'enriching or adding conceptual material' is necessary in such cases if we are to 'arrive at what the speaker intended to express'.

For example, (20) may understood, in a certain context of utterance, as expressing something like the proposition (resolved for indexicals) that *learning to play the piano will take a longer period of time than you seem to think.* Similarly, (21) could be understood as expressing the proposition that *London is some considerable distance from Sheffield* or that *the distance between London and Sheffield is greater than you seem to think.* Carston (Ibid) suggests *something bad has happened on the day of utterance [to x]* as a possible content for (22). She notes that all of these examples seem to have a *negative* flavour: the relevance of each of these seems to be 'alerting the hearer to a state of affairs that runs against his prevailing hopes or expectations' (Ibid).

As Carston points out, (23) and (24) require some assignment of time period to be fully propositional. Their interest lies in 'the difference that the object...makes in each case to the understanding of the identical verbs' (2002: 26). Thus (23) would typically be understood as *I haven't eaten breakfast TODAY*, whereas (24) would be understood as *I haven't eaten scallops IN MY LIFETIME* (Ibid: 27). However, note that (24) certainly *could* be used to express the proposition that *I haven't eaten scallops today* in an appropriate context. Suppose that Jones has eaten scallops everyday for the past week, but today does not eat any. When asked, 'Have you eaten today?' by a friend, he could well reply, 'Yes, and − shock horror − I haven't eaten scallops!' The point that we can express different propositions by such a sentence is an important point that I will return to presently. As Carston points out, on almost all occasions of utterance, sentences such as (25) are strictly false. There is almost always something in the fridge (especially if we count air, shelves, drawers and the like). Thus to arrive at the

proposition intended, one must 'restrict the domain over which "nothing" operates' – for instance, to edible food, or ingredients sufficient to make a meal in conjunction with what's in the cupboard and the vegetable rack, and so on (2002: 27).

The above cases of underspecificity of conceptual content involve some addition of conceptual constituents. However, as mentioned, there are other cases that involve some sort of conceptual enrichment. Consider (adapted from Carston 2002: 27):

(26)   I'm hungry.

(27)   Ann wants to meet a bachelor.

(28)   The garden is uneven.

As Carston points out, (26)-(28) require some narrowing or strengthening of their conceptual concept in order to deliver what is said. Thus the relevant degree of hunger in (26) may require a snack or a meal, to stop in the next hour or so or at the next service station. Ann probably doesn't want to meet just any old bachelor – she is understood as wanting to meet someone who belongs 'just to a particular subset of the set of bachelors': the youngish, handsome, eligible sort (Ibid). Finally, note that just about *any* garden will be uneven (strictly speaking); in uttering (28), a speaker calls attention to the fact that the garden is uneven enough to merit levelling, or that it does not serve as a good football pitch, or that its upper surface is such that one should mind their step.

Finally, *overspecificity of conceptual content* covers cases where the literal linguistic meaning of a sentence is too narrow to be a plausible contender for the proposition expressed by an utterance of that sentence. For example, consider:

(29)   Her face is heart-shaped.

(30)   The steak is raw.

(31)   The audience was silent

As Carston points out, we need to 'loosen' the concepts involved here if we are to arrive at a plausible account of what is said on any occasion of utterance of (29)-

(31). Thus, one does not mean that someone's face is *literally* heart shaped by uttering (29) – we need to 'relax' the concept of being heart-shaped. Similarly, we may describe a steak as raw when it has received insufficient cooking so that it is not literally raw (i.e. uncooked) (Ibid). Likewise, no audience is likely to be *utterly* silent – small rustling sounds of sweet packets, the odd cough, and, of course, breathing are all likely to ensue during a theatre performance.

What the above taxonomy shows is that the failure of linguistic meaning to provide something fully propositional (truth-evaluable) – that is, to specify 'what is said' – is *rife* in natural language. The under-determination of propositional content by linguistic meaning extends far beyond those cases involving indexical expressions and vague terms. We must now consider the implications of this result for various projects of semantics – both philosophical semantics, the study of word-world relations, and semantics in its more general sense as the study of linguistic meaning.

Recall that in philosophical semantics, the meaning of a declarative sentence is often identified with its truth-condition, and the meaning of a word with its correctness-condition. The above taxonomy shows, however, that this may be a mistaken approach. Correctness-conditions may not be 'strict' – they may be relaxed or tightened depending on the context. It is not immediately clear whether such conceptual tinkering should be considered as part of their linguistic meaning or not. When it comes to truth-conditions, we may separate out two distinct contributions: the contribution made by linguistic meaning, and the contribution made by pragmatics. As I have argued above, the contribution made by pragmatics is significant. The question now becomes *given its heavy reliance on the contribution made by pragmatics, is the study of semantics (understood as the study of truth-conditions) theoretically tractable*? In other words, given the pragmatic contribution to the determination of truth-conditions, are we likely to get a systematic scientific theory of truth-conditions or 'what is said'? I will now argue that the answer to this question is 'no'.

Interaction Effects: Data and Phenomena

Piertroski (2003, 2005) argues that the truth or falsity of some utterance of a sentence is a 'massive interaction effect', determined not just by linguistic meaning but also by a whole host of other factors perhaps not amenable to scientific investigation, such as 'facts about how "reasonable" speakers would use' certain sentences (2003: 218). Rather than specifying truth-*conditions*, he suggests that the contribution of linguistic meaning is weaker, providing something like truth-*indications*. These may be construed along essentially Strawsonian lines as something like general directions concerning what a sentence may be *used to say*. That is,

> The meaning of [a sentence] S is a compositionally determined intrinsic property of S that *constrains and guides without determining* how S can be *used to make* true or false assertions in various conversational situations (2005: 256).

Importantly, sentence *use* is a kind of human action; and thus perhaps not itself a theoretically tractable phenomenon (Ibid). In Chomsky's phrase, 'as soon as questions of will or decision or reason or choice of action arise, human science is at a loss'.

Both Davidson ([1986] 2006) and Chomsky (2000a) have made the point that interpretation and communication are topics that are beyond scientific study. By 'interpretation', what is meant is this: '[an] interpreter, presented with an utterance and a situation, assigns some interpretation to what is being said by a person in this situation' (Chomsky, 2000a: 69). In order to do so, they will draw on any clues provide by the linguistic meaning of the utterance, the linguistic faculties as well as 'all other capacities of the mind', and even 'non-linguistic inputs'. Thus what is potentially relevant to the assignment of the interpretation is 'everything people might know and are capable of doing' (Carston 2002: 1). Chomsky argues that since 'the interpreter…includes everything that people are capable of doing' it is 'not an object of empirical inquiry, and…nothing sensible can be said about it' (2000a: 69).

Smith (2004: 152-3) provides a nice way of capturing the distinct contributions of linguistic meaning, and the other factors that go into an assignment of an interpretation. On the one hand, we have linguistic meaning which we may think of as the partial *encoding* of a thought or proposition. Thus in

an act of communication, the contribution made by linguistic meaning has to be decoded. On the other hand, we have pragmatics – those contingencies of context, belief, and so on, on the basis of which we make *inferences* regarding the proposition that our interlocutor has in mind. Thus communication involves both decoding and inference. And – this is the crucial point – virtually *any* belief or piece of knowledge could be relevant to drawing the correct inferences, putting communication beyond the reach of a systematic theory.

Steven Pinker (1994: 227) gives an excellent example of just how complex the interpretation of even a very simple utterance can be. Consider the following dialogue:

WOMAN: I'm leaving you.
MAN: Who is he?

Now, Pinker invites us to reflect on how we interpret the pronoun 'he' in the man's utterance. On the face of it, the man's utterance is a complete non-sequitur. However, we naturally and effortlessly interpret it as *who is the man that you are leaving me for*. To understand it, we must draw on all kinds of knowledge and beliefs that we have about human interaction. How do we arrive at such an interpretation? First, we need to know something about psychology – that say, people may react with jealousy and suspicion in such a situation. We also need to know that, on occasions, people may leave their partners because they have met somebody else, and that this is a conclusion that someone in the man's position might jump to. Our question is whether we could have a systematic theory of how we get so effortlessly to the interpretation of the utterance.

Consider further these examples (from Stainton 2006: 931). We assign different meanings to all of the modifiers in the following list due to our *knowledge* about the world: 'Christmas cookie' (made to be consumed at), 'Girl Guide cookie' (sold by), 'oatmeal cookie' (made of), 'yellow cookie' (coloured), 'fortune cookie' (containing), 'doggie cookie' (made to be eaten by), and Sainsbury's cookie (sold at). Thus, Stainton argues, if someone is able to assign the right kind of truth-condition to, say, 'Jones ate a Girl Guide cookie' it is because they know the relevant facts about girl-guides, *viz*. that they sell cookies,

but do not comprise the ingredients of cookies, or reside inside cookies, and so on. It is the rich interaction between the language faculty and our systems of belief and knowledge about the world that allows us to arrive at the correct truth-condition in this case.

To take yet another example, Stainton points out the difference in the way we interpret sentences such as 'Jones may turn up later' and 'Jones may have a biscuit if he wishes to'. The interpretation assigned to 'may' in each case is dependent on the relevant bits of knowledge we have about the world, i.e. that turning up is not usually people are given *permission* to do, unlike having a biscuit. That is, interpreting the first sentence as expressing *possibility* and the second as expressing *permissibility* requires again the relevant knowledge about the world (Ibid).

So, to return to our question: can we have a systematic, scientific theory of just how expressions come to have the particular truth-conditions that they do have? Bogen and Woodward (1988) provide the grounds for an argument to the effect that we are not likely to get a semantic theory of the envisaged kind. The crucial distinction they make is that between *data* and *phenomena*. *Data* play the role of providing evidence for phenomena and typically can be observed; however 'data typically cannot be predicted or systematically explained by theory' (305-6). By contrast, scientific theories do 'predict and explain facts about phenomena' (306). Phenomena are 'detected through the use of data, but in most cases are not observable in any interesting sense of that term' (Ibid). Facts about phenomena may in turn serve as evidence, but they provide evidence for high-level general *theories* by which they are explained (Ibid). Thus 'facts about data and facts about phenomena differ in what they serve as evidence for (claims about phenomena versus general theories) (Ibid)'.

Now, Bogen and Woodward argue, it is a mistake to think that theories explain facts about data (Ibid). This is well illustrated by use of the following example (307-310): Suppose we set up an experiment to determine the melting point of lead (which, for the record, is taken to be 327 degrees C). To do so, one must conduct a series of measurements. Now, Bogen and Woodward point out:

Even when the equipment is in good working order, and sources of systematic error have been eliminated, the readings from repeated applications of the same thermometer to the same sample of lead will differ slightly from one another, providing a scatter of results. These constitute *data*. Given the absence of systematic error, a standard assumption is that the scatter of observed thermometer readings not only reflects the true melting point (the *phenomenon* in which we are interested), but also the operation of numerous other small causes of variation or 'error', causes which cannot be controlled for and the details of which remain unknown. If one can make certain general assumptions about the character of these causes of variation (for example, that they operate independently, are roughly equal in magnitude, are as likely to be positive as negative, and have a cumulative effect which is additive), then it will follow that the distribution of measurement results will be roughly normal and that the mean of this distribution will be a good estimate of the true melting point. Standard scientific practice is to report this estimate along with the associated standard error, which is directly calculable from the variance of the distribution of measurement results (1988: 308, my emphasis).

Now, the first important point to note here is this: the sentence 'lead melts at 327 degrees C' does *not* report what is observed. Rather, 'what we observe are the various particular thermometer readings – the scatter of individual data points' (309). The mean of these (327 degrees C) 'does not represent a property of any individual data point; indeed, there is no reason to think that any individual measurement will coincide with the mean value (Ibid). So,

[w]hile the true melting point is certainly inferred or estimated from observed data, on the basis of a theory of statistical inference and various other assumptions, the sentence 'lead melts at 327.5 + 0.1 degrees C' – the form that a report of an experimental determination of the melting point of lead might take – does not literally describe what is perceived or observed (309).

Now, the second crucial point is this: it is the job of a scientific theory of molecular structure to explain *the phenomenon that lead melts at 327 degrees C*. However, such a theory does not explain – and, indeed *could not* explain – why the various data points that were observed were, in fact, observed:

The outcome of any given application of a thermometer to a lead sample depends not only on the melting point of lead, but also on its purity, on the workings of the thermometer, on the way in which it was applied and read, on interactions between the initial temperature of the thermometer and that of the sample, and a variety of other background conditions. *No single theory could accurately predict or explain an outcome which depends upon the confluence of so many variable and transient factors* (309-10, my emphasis).

In other words, it is the job of a theory to account for *phenomena* and not for *data*; data cannot be explained by a single theory as they arise not just because of the

phenomena that that theory explains, but also because of a whole host of other factors.[11]

Now, let us return to the semantic case. Analogously, we here have the *data* that various utterances are interpreted as having such-and-such truth conditions in such-and-such circumstances. However, it is crazy to expect a single scientific theory to account for *why* this should be so. Scientific theories target and explain the stable *phenomena* that underlie the data we observe, and *not* the data itself. Linguistic meaning may be such a stable phenomenon, but, as we argued above, linguistic meaning does not, in and of itself, determine truth-conditions. Truth-conditions are rather the result of a massive interaction between linguistic meaning and matters of knowledge, belief and context; they are analogous not to the melting point of lead, but to the various meter-readings that we make when trying to determine what that melting point is.

Moreover, it is *this* kind of complexity that shows the search for a scientific semantic theory is hopeless: there are just too many disparate factors that go into determining the particular data that we have to go on to expect a single theory to account for them all. To ask why a sentence uttered by a speaker in such-and-such circumstances had the truth-condition that it did is analogous to asking why we observed precisely *those* meter-readings, or why a particular leaf followed precisely the trajectory that it did when falling from the tree (Piertroski 2003: 240). As Stainton puts it, 'genuine sciences are in the business of describing causal forces, not such highly complex particular effects' (2006: 933).[12] Thus, as Chomsky points out, such questions as why the leaf followed the particular trajectory it did are not topics for scientific inquiry (2000a: 49), even though it is 'quite true that scientific laws together contribute to how the leaf in fact fell' (Stainton 2006: 933). Or to take another case: it is not the part of grammatical theory to account for every facet of a speaker's actual performance, such as false

---

[11] For a recent restatement and defence of the data/phenomena distinction, see Woodward 2011.

[12] Stainton raises a related worry with regards to the so-called 'sciences of the complex', such as meteorology and economics. Although these 'no doubt…uncover statistical patterns in the weather…and the stock market, using sophisticated empirical methods', they still do not qualify as sciences in our sense as they do not involve 'postulating unobservables' or 'the aim of integration with other core natural sciences' (2006: 933-4). Thus even if semantics turned out to be a 'science of the complex', it is still unclear that it would be a good candidate for naturalisation in the metaphysical sense.

starts, hesitation, and so on. Linguists abstract away from such aspects of the linguistic data in order to get to the underlying grammatical phenomena. Note that this does not render linguistic data *inexplicable*: it may be that we can combine a systematic theory of linguistic meaning with some common sense observations about belief and context to give a more informal explanation of the truth-conditions of some utterance in context (Piertroski 2003: 24). However, this doesn't give us 'a *theory* that explains the particular facts, at least not in any interesting sense of "theory"' (Ibid).

6.4 Conclusion

Calls to naturalise meaning to other entities, properties and relations are not well motivated. In order for a successful reduction of some domain to a domain of scientific discourse, we need a clear theoretical conception of the reduction base *and* the reduction target. There are good reasons to think that the project of philosophical semantics – the study of the relations between expressions of public languages and things-in-the-world – is not theoretically tractable. This is because the factors that enter into the determination of truth-conditional content are too varied, complex and intricate to submit to (methodologically) naturalistic theoretical investigation. Furthermore, the notions of public language, and the concepts of common sense, are too dependent of human interests and concerns to enter into scientific theorising.

From this I think we can draw a further, more general moral concerning a certain conception of philosophy. Above we drew the distinction between the perspectives afforded by science and common sense. If philosophy consists in the investigation of the latter, and the above arguments are cogent, then the whole project of metaphysical naturalism seems to be based on a faulty foundation. In other words, the idea that our 'pre-theoretic' notions of phenomena must be 'naturalised' amounts to little more than the insistence that the concepts of common sense be cashed out in terms of the concepts of the sciences in order to be metaphysically respectable. It seems to me that they *cannot* be so naturalised, for the reasons outlined above. Against the metaphysical naturalist's claim, it can be said that our pre-theoretic ways of thinking about the world are in perfectly good order. The objects of common sense *exist*. However, we cannot account for

their existence in terms of scientific concepts: science is blind to such things. Accordingly, pre-theoretic claims about the world do not need to be reduced to scientific claims to be respectable, or to be taken as aiming to state truths about the world.

Where does all of this leave us with the meaning-sceptic? Firstly, I have certainly not provided anything close to a refutation of the sceptical argument. However, I have aimed to undermine claims about meaning that could be utilised by the semantic sceptic. The first of these was the claim that meaning is normative. This seemed to provide the sceptic with grounds for several *a priori* arguments against the possibility of factualism about meaning. Happily, we have seen that this claim is false. The second claim, undermined in this chapter and the last, was commitment to a kind of militant naturalism: the claim that, if meaning could not be shown to reduce to respectable prior ontological categories (as provided by the natural sciences), then it should be eliminated from out ontology.

However, there is still a formidable hurdle to face as a full refutation of the sceptic will require, as Wright rightly points out, a satisfactory account of both the epistemology and curious properties of meaning. In other words, we must account for how meaning-states can have the remarkable – and, seemingly, paradoxical – properties of being known immediately and with fair certainty, yet not locatable in consciousness, and speaking to how language relates to the world in a variety of unconsidered circumstances.

Kripke's sceptic argues for the outrageous conclusion that there are no facts of the matter about what we mean by our expressions. In order to secure this shocking result, the sceptic considers various contenders for a meaning-constituting fact and finds them all wanting. In the present work, I have aimed to undermine two of the sceptic's powerful lines of attack: arguments from the claim that meaning is normative to semantic non-factualism, and claims that meaning should be eliminated from our ontology if it cannot be shown to reduce to the categories of the natural sciences.

Hattiangadi (2007) suggests that, granted the assumption that meaning is normative the sceptic can strengthen his argument from a 'by elimination' argument to an *a priori* one. However, while the claim that meaning is normative seemed *prima facie* to be the basis of a powerful reason for rejecting *a priori* any naturalistic theory of meaning, we have seen that in the end it is unable to deliver. Either the claim that meaning is normative, when understood in a sense that would provide the sceptic with an argument to non-factualism, has turned out to be false, or the arguments themselves have involved highly implausible premises. Thus, *pace* Kripke, there is no requirement on a theory of meaning to make room for the normativity of meaning. However, even disregarding the arguments from strong semantic normativity, the fact remains that the sceptical argument has not yet submitted to a naturalistic solution.

The most pressing problem for a naturalistic account of meaning once strong normativity has been dispensed with is the problem of error, or the disjunction problem. This is to give an account of the extensions of terms such that they are determinate and fine grained enough to exclude sceptical alternatives to the intuitive meanings of our expressions. Our survey of the most prominent naturalistic accounts of meaning in the literature led to the conclusion that they all seemed to fail in this aim. Either the extensions that such responses determined were not fine grained enough, or they only became so once semantic and intentional notions were surreptitiously smuggled into the accounts. Thus they

either failed to rule out sceptical alternatives, or violated the terms of discussion with the sceptic.

This left us with some rather pessimistic results: all extant naturalistic responses to the sceptic fail, and the strong normativity thesis perhaps seemed less pivotal to the success of the sceptical argument than it appeared at first blush. One might even have concluded that the failure of all naturalistic theories of meaning perhaps showed that there were no meaning-constituting facts to characterise after all, and thus conceded that the sceptic had won.

However, as I argued in chapters 5 and 6, the failure of all extant naturalistic accounts of meaning can be pinned on other reasons. These stem from the fact that the naturalist's project of accounting for facts of all kinds in terms of the entities of the natural sciences faces several severe difficulties of its own. Firstly, we are not in a position to say that any of our current best scientific theories are *true* theories. In fact, the historical record of such theories provides excellent empirical evidence that many, perhaps all, of our current theories are in fact false. Thus the naturalist may be trying to make sense of recalcitrant phenomena using only the resources provided by *false* theories.

Secondly, there is the related point that one should not assume that the concepts of scientific theories are real and metaphysically fixed entities. As we have seen, science may treat a concept as some piece of instrumental apparatus at one time, but later commit to regarding it as a real thing-in-the-world. Often this comes about as a result of modification of a *lower level* theory. Thus it is a mistake to proceed methodologically by taking the concepts of the reduction base to be fixed, and trying to make sense of everything else in terms of them. Furthermore, phenomena that have seemed isolated from a given theory have proven to be perfectly consonant with it once that theory is reconceptualised because of changes in lower level theory. Thus while a theory may be complete in the sense of being empirically adequate over certain phenomena, our understanding of that theory may depend on modifications made to lower level theories rather than new knowledge about the primitive concepts of the higher level theory.

History teaches us that true reduction of one theory to another is extremely rare. More often what happens when we integrate various domains of inquiry is *unification*: this may involve modifications to both lower and higher level theories, or wholesale revision of either; there is no presumption that the concepts of the lower level theories are more firmly fixed. In fact, developments in higher level theories may even prompt revision of lower level ones. While the sciences of mind, language and brain are not yet as established as physics, or chemistry, this is no reason to regard them with increased suspicion, or to make *a priori* stipulations regarding the resources they can utilise.

Successful reduction – indeed, successful unification – proceeds on the basis of having clear and highly articulated theories of both the base and the target phenomena. In chapter 6, I argued that the required kind of theory for semantics is beyond our limits. As philosophers understand the term, semantics is the study of word-worlds relations, where the semantic relata are understood to be words in public languages and real things-in-the-world. I argued that the ontological status of both relata rendered them 'invisible' to the peculiar perspective of natural science. This is because their individuation depends crucially on subjective human concerns, and the particular perspectives provided by human common sense. Furthermore, public languages are simply not robust enough entities to enter into scientific inquiry. Crucially, there are no laws or principles that hold of languages *qua* sociological constructs. Similarly, there are no laws that hold of the entities that correspond to the concepts of common sense – that is, there are no laws of pencils *qua* pencils, say (Fodor 1980). Natural science may be unable to give identity conditions for many of our common sense concepts precisely because they are so dependent on human interests and concerns.

A further problem is that it is not clear that semantics, conceived of as involving the study of the truth-conditions of natural language declaratives, is a theoretically tractable domain. While aspects of meaning may be systematic enough to be captured in scientific generalisations, truth-conditions may not be. We surveyed numerous cases of the under-determination of truth-conditional content by linguistic meaning. Truth-conditions are not encoded by meaning; rather, they are inferred on the basis of the contributions of linguistic meaning (which is encoded) and pragmatics (which is not). Thus communication requires

both decoding and inference; however, the list of beliefs and pieces of knowledge required for drawing the right inferences is potentially *infinite*; thus the study of interpretation borders on the study of *everything* (Chomsky 2000a).

The above remarks clearly do not constitute a *solution* to the sceptical paradox. I have not been able to cite a fact that categorically refutes the scepticism. However, I hope that I have gone some way to effecting a partial *dissolution* of the paradox. There are facts of the matter about what we mean, and the propositions we express, but these cannot be captured naturalistically (in the metaphysical sense) as the truth-conditional content of our utterances is beyond the scope of any naturalistic scientific theory. This is not because there are no facts to capture; rather, such facts are too intractable to be scientifically captured. Therefore, any attempt to make sense of them using only the resources made available by the natural sciences is bound to fail. Perhaps some time in the future, natural science will be able to make sense of the contribution made to truth-conditional content by linguistic meaning itself; perhaps this will even be integrated with some higher portion of neuro-physiological theory, say. Until such a time, attempts to naturalise semantics seem doomed – but this doesn't amount to the success of meaning-scepticism.

Of course, to defeat the sceptic once and for all, one needs an account of the properties of meaning-states. That is, one must answer Wright's question of how such states can speak to what to do in a potential infinity of circumstances of use, yet be immediately and privately known even though we cannot seem to 'locate' them in consciousness. This is surely a deep and difficult problem. However, as such one may demur from drawing sceptical conclusions about meaning on the back of our failure to provide a fully satisfying solution (as yet). One thing is for certain: such a cause is not aided by *a priori* stipulation of the kinds of facts one may appeal to in an account of meaning. Happily, we have seen that such stipulations are unwarranted.

Bibliography

Almog, J., Perry, J., and Weinstein, H. (eds.) 1989. *Themes from Kaplan*. New York: Oxford University Press.

Antony, L and Hornstein, N. (eds.) 2003. *Chomsky and his Critics*. Oxford: Blackwell.

Ayer, A. J. 1936. *Language, Truth and Logic*. Oxford: Basil Blackwell.

Barber, A. (ed.) 2003. *Epistemology of Language*. Oxford: Oxford University Press.

Bezuidenhout, A. L. 2006. 'Language as Internal'. In Lepore & Smith 2006, pp. 127-139.

Bilgrami, A. 1993. 'Discussion'. In Chomsky 1993.

————— 1998. 'Why Holism is Harmless and Necessary'. *Philosophical Perspectives*, Vol. 12, pp. 105-126.

Bilgrami, A. & Rovane, C. 2007. 'Mind, language, and the limits of inquiry'. In McGilvray, J. (ed.), *The Cambridge Companion to Chomsky*. Cambridge: Cambridge University Press.

Blackburn, S. 1984. 'The Individual Strikes Back'. *Synthese*, Vol. 58, pp. 281-301.

Bogen, J. & Woodward, J. 1988. 'Saving the Phenomena'. *Philosophical Review* Vol. 97, pp. 303-52.

Boghossian, P. 1989a. 'The Rule-Following Considerations'. *Mind* Vol. 98, No. 392, pp. 507-549

——————— 1989b. 'Review of McGinn 1984', *Philosophical Review*. Vol. 98, pp. 83-92.

——————— 1990. 'The Status of Content', *Philosophical Review*. Vol. 99, pp. 157-84.

——————— 1991. 'Naturalising Content'. In Loewer, B. & Rey, G. (eds.), *Meaning in Mind: Fodor and his Critics*. Oxford: Blackwell.

——————— 2005. 'Is Meaning Normative?' In Nimtz, C. & Beckermann A. (eds.) *Philosophy – Science –Scientific Philosophy*. Paderborn: Mentis.

Brandom, R. B. 1994. 'Unsuccessful Semantics'. *Analysis*, Vol. 54, pp. 175-178.

Burge, T. 1979. 'Individualism and the Mental'. *Midwest Studies in Philosophy* Vol. 4, pp. 73-121

Carston, R. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Cheng, K. 2009. 'Semantic Dispositionalism, Idealization, and Ceteris Paribus Clauses'. *Minds & Machines*, Vol. 19, pp. 407–419.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

——————— 1993. *Language and Thought*. Rhode Island: Moyer Bell.

——————— 2000a. *New Horizons in the Study of Language and Mind.* Cambridge: Cambridge University Press.

——————— 2000b. 'Linguistics and Brain Science'. In Marantz, A. Miyashita, Y. and O'Neil, W. (eds.) *Image, Language, Brain*. Cambridge, MA: MIT Press.

———— 2003. 'Reply to Millikan'. In Antony & Hornstein, 2003, pp. 308-315.

Churchland, P. 1981. 'Eliminative Materialism and the Propositional Attitudes'. *Journal of Philosophy*, Vo. 78, pp. 67-90.

Collins, J. 2008. *Chomsky: A Guide for the Perplexed*. London: Continuum.

———— 2009. 'Methodology, not Metaphysics: Against Semantic Externalism'. *Proceedings of the Aristotelian Society*, Supplementary Vol. 83, pp. 53-69.

———— 2010. 'Naturalism in Philosophy of Language'. In Sawyer 2010, pp. 41-59.

Crane, T. 2003. *The Mechanical Mind*. 2nd edition. London: Routledge.

Dancy, J. 2000. *Practical Reality*. Oxford: Oxford University Press.

Davidson, D. 1963. 'Actions, Reasons and Causes'. *Journal of Philosophy*, Vol. 60, pp. 685–700.

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

———— [1986] 2006. 'A Nice Derangement of Epitaphs'. In his *The Essential Davidson*, Oxford: Clarendon Press.

———— 1994. 'Davison, Donald' in Guttenplan, 1994, pp. 231-235.

Davies, D. 1998. 'How Sceptical is Kripke's "Sceptical Solution"?' *Philosophia*, Vol. 26, No.1-2, pp. 119-140.

Dennett, D. 1984. Foreword to Millikan, R. G. *Language, Thought and Other Biological Categories*. Cambridge MA: MIT Press.

Dretske, F. 1986. 'Misrepresentation'. In Bogdan, R. J. (ed.), *Belief: Form Content and Function*. Oxford: Clarendon Press.

Dretske, F. 2000. *Perception, Knowledge and Belief: Selected Essays*. Cambridge: Cambridge University Press

Dummett, M. 1978. *Truth and Other Enigmas*. London: Duckworth.

Dummett, M. 1986. 'A nice derangement of epitaphs: some comments on Davidson and Hacking'. In Lepore, 1986, pp. 459-76.

Fodor, J. 1980. 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology'. *The Behavioural and Brain Sciences*, Vol. 3, pp. 63-109.

———— 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

————1990. *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

————1991. 'Reply to Boghossian'. In Loewer & Rey 1991, pp. 271-277.

Frankfurt, H. 1971. 'Freedom of the Will and the Concept of a Person'. In Watson 1982, pp. 81-95.

Frege, G. [1892] 1993. 'On Sense and Reference'. In Moore, A. W. (ed.), *Meaning and Reference*. Oxford: Oxford University Press.

Gampel, E. H. 1995. 'The Normativity of Meaning'. *Philosophical Studies* Vol. 86, No. 3, pp. 221-242

Glüer, K. 1999. *Sprache und Gegeln: Zur Normativität von Bedeutung*. Berlin: Akademie Verlag.

Glüer, K. and Pagin, P. 1999. 'Rules of Meaning and Practical Reasoning'.
　　　Synthese Vol. 117, pp. 207-27

Glüer, K. and Wikforss, A. 2009a. 'The Normativity of Meaning and Content'. In
　　　The Stanford Encyclopedia of Philosophy (Winter 2010 Edition), Edward
　　　N. Zalta (ed.),
　　　URL　=　<http://plato.stanford.edu/archives/win2010/entries/meaning-
　　　normativity/>.

————— 2009b. 'Against Content Normativity'. *Mind* Vol. 118,
　　　No. 469, pp. 31-70

Goodman, N. 1979. *Fact, Fiction & Forecast*. Cambridge, MA: Harvard
　　　University Press.

Grice, P. 1989. *Studies in the Ways of Words*. Cambridge, MA: Harvard
　　　University Press.

Gross, S. 1998. *Essays on Linguistic Context-sensitivty and its Philosophical
　　　Significance*. Unpublished PhD thesis, Harvard University.

Guttenplan, S. (ed.) 1994. *A Companion to the Philosophy of Mind*. Oxford:
　　　Blackwell.

Hale, B. and Wright, C. (eds.) [1997] 2006. *A Companion to the Philosophy of
　　　Language*. 2nd Edition. Oxford: Blackwell.

Hale, B. and Wright, C. 1997. 'Putnam's Model-Theoretic Argument Against
　　　Metaphysical Realism'. In Hale & Wright [1997] 2006, pp. 427-457.

Hattiangadi, A. 2006. 'Is Meaning Normative?' *Mind and Language* Vol. 21, No.
　　　2, pp. 220-240.

———— 2007. *Oughts and Thoughts: Rule Following and the Normativty of Content*. Oxford: Oxford University Press.

———— 2010. 'Semantic Normativity in Context'. In Sawyer 2010, pp. 87-107.

Hattiangadi, A. & Bykvist, K. 2007. 'Does Thought imply Ought?' *Analysis* Vol. 67, No. 296, pp. 277-285

Hume, D. [1739] 1978. *A Treatise of Human Nature* 2nd Edition. Oxford: Clarendon Press.

Isac, D. & Reiss, C. 2008. *I-Language: An Introduction to Linguistics as Cognitive Science*. Oxford: Oxford University Press.

Kaplan, D. 1989. 'Demonstratives'. In Almog, Perry, and Weinstein, 1989, pp. 481-564.

Kripke, S. 1979. 'A Puzzle About Belief'. In Margalit, A. (ed.), *Meaning and Use*. Dordrecht: Reidel.

———— [1982] 2003. *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.

Kusch, M. 2006. *A Sceptical Guide to Meaning and Rules: Defending Kripke's Wittgenstein*. Chesham: Acumen.

Lenman, J. 2009. 'Reasons for Action: Justification vs. Explanation'. In *The Stanford Encyclopaedia of Philosophy (Spring 2010 Edition)*, Edward N. Zalta (ed.),
URL = <http://plato.stanford.edu/archives/spr2010/entries/reasons-just-vs-expl/>.

Lepore, E. (ed.) 1986. *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell.

Lepore, E. & Smith, B. C. (eds.) 2006. *The Oxford Handbook of Philosophy of Language*. Oxford: Oxford University Press.

Loewer, B. [1997] 2006. 'A Guide to Naturalising Semantics'. In Hale & Wright, 2006, pp. 108-126.

Lycan, W. 2003. 'Chomsky and the Mind-Body Problem'. In Antony and Hornstein, 2003, pp. 11-28.

McGinn, C. 1984. *Wittgenstein on Meaning: An Interpretation and Evaluation*. Oxford: Blackwell.

McDowell, J. 1994. *The Mind and the World*. Cambridge, MA: Harvard University Press.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.

Martin, C. B. 1994. 'Dispositions and Conditionals'. *Philosophical Quarterly*, Vol. 44, pp. 1-8.

Martin, C. B. & Heil, J. 1998. 'Rules and Powers'. *Philosophical Perspectives*, Vol. 12: *Language, Mind and Ontology*, pp. 283-312.

Melnyk, A. 1997. 'How to Keep the "Physical" in Physicalism'. *Journal of Philosophy*, Vol. 94, pp. 622-37.

Midgley, G. C. J. 1959. 'Linguistic Rules'. *Proceedings of the Aristotelian Society* Vol. 59, pp. 271-90.

Millar, A. 2002. 'The Normativity of Meaning'. in O'Hear, A. (ed.) *Logic Thought and Language* Cambridge: Cambridge University Press.

————2004. *Understanding People: Normativity and Rationalising Explanation*. Oxford: Oxford University Press.

Miller, A. 2003a. *An Introduction to Contemporary Meta-Ethics*. Oxford: Blackwell.

———— 2003b. 'Does Belief Holism Show that Reductive Dispositionalism about Content Could Not be True?' *Proceedings of the Aristotelian Society* Supplementary Volume 77, pp. 73-90.

———— 2006. 'Meaning Scepticism'. In Devitt, M. & Hanley, R. (eds.) *The Blackwell Guide to Philosophy of Language*. Oxford, Blackwell.

———— 2007. *Philosophy of Language* 2nd Edition. London: Routledge.

———— 2010. 'The Argument from Queerness and the Normativity of Meaning'. In Granjer, M. & Rami, A. (eds.) *Truth, Existence and Realism* (Paderborn: Mentis).

———— forthcoming. 'Semantic Realism and the Argument from Motivational Internalism'. Forthcoming in Schantz, R. (ed.) *What is Meaning* (De Gruyter).

Miller, A. & Wright, C. (eds.) 2002. *Rule-Following & Meaning*. Chesham: Acumen.

Millikan, R. G. 1986. 'Thoughts Without Laws: Cognitive Science without Content'. *Philosophical Review*, Vol. 95, pp. 47-80.

———— 1990. 'Truth, Rules, Hoverflies and the Kripke-Wittgenstein Paradox'. *The Philosophical Review*, Vol. 90, pp. 323-53.

———— 1991. 'Speaking up for Darwin'. In Loewer & Rey 1991.

———— 1995. *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press.

Pagin, P. 2002. 'Rule-following, Compositionality and the Normativity of Meaning'. In Prawitz, D. (ed.), *Meaning and Interpretation*, Konferenser 55, Kungl. Vitterhets Historie och Antikvitetsakademien, Stockholm 2002. Accessed online at: http://people.su.se/~ppagin/papers/comprules.pdf.

Papineau, D. 2006. 'Naturalist Theories of Meaning'. In Lepore & Smith 2006, pp. 175-188.

———— 2007. 'Naturalism'. In *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL<http://plato.stanford.edu/archives/spr2009/entries/naturalism/>.

Piertroski, P. 2003. 'The Character of Natural Language Semantics'. In Barber, 2003, pp. 217-256.

Piertroski, P. 2005. 'Meaning Before Truth'. In Preyer & Peters, 2005, pp. 255-302.

Pinker, S. 1994. *The Language Instinct*. London: Penguin.

Preyer, G. and G. Peters, G. 2005. *Contextualism in Philosophy*. Oxford: Oxford University Press

Putnam, H. 1975. 'The meaning of "meaning"'. In his *Mind Language and Reality*, 215-271. Cambridge: Cambridge University Press.

Quine, W. V. O. 1960. *Word and Object*. Cambridge, MA: MIT Press.

———— 1969. 'Ontological Relativity'. In *Ontological Relativity and Other Essays*. New York: Columbia University Press.

———— 1970. 'On the Reasons for Indeterminacy of Translation'. *Journal of Philosophy*, Vol. 67, No. 6, pp. 178-83.

Rosenberg, A. 2005. *Philosophy of Science*. 2nd Edition, Abingdon: Routledge.

Sawyer, S. (ed.) 2010. *New Waves in Philosophy of Language*. Houndmills: Palgrave MacMillan.

Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Smart, J. 1978. 'The Content of Physicalism'. *Philosophical Quarterly*, Vo. 28, pp. 339-41.

Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell.

Smith, N. 2004. *Chomsky: Ideas and Ideals* 2nd Edition. Cambridge: Cambridge University Press

Stainton, R. 2006. 'Meaning and Reference: Some Chomskyan Themes'. In Lepore & Smith 2006, pp. 913-940.

Stanley, J. and Szabó Z. 2000. 'On Quantifier Domain Restriction'. *Mind & Language* Vol. 15, pp. 219-61.

Watson, G. 1975. 'Free Agency'. In Watson 1982, pp. 96-110.

Watson, G. (ed.). 1982. *Free Will*. Oxford: Oxford University Press.

Weiss, B. 2010. *How to Understand Language: A Philosophical Investigation*. Durham: Acumen.

Whiting, D. 2007. 'The Normativity of Meaning Defended'. *Analysis* Vol. 67, No. 2, pp. 507-549.

Whyte, J. T. 1990. 'Success Semantics'. *Analysis*, Vol. 50, pp. 149-57.

———— 1991. 'The Normal Rewards of Success'. *Analysis*, Vol. 51, pp. 65-

73.

Wikforss, A. 2001. 'Semantic Normativity'. *Philosophical Studies* Vol. 102, No. 2, pp. 203-226.

Wilson, G. [1994] 2002. 'Kripke on Wittgenstein on Normativity'. In Miller & Wright 2002, pp. 234-259.

Wittgenstein, L. [1953] 2001. *Philosophical Investigations*. 3rd edition. Trans. G. E. M. Anscombe. Oxford: Blackwell.

Woodward, J. 2011. 'Data and phenomena: a restatement and defence'. *Synthese* Vol. 182, pp. 165–179.

Wright, C. 1984. 'Kripke's Account of the Argument Against Private Language'. *The Journal of Philosophy* Vol. 81, pp. 759-77.

———— 1989. 'Wittgenstein's Rule Following Considerations and the Central Project of Theoretical Linguistics'. In George, A. (ed.), *Reflections on Chomsky*. Oxford: Blackwell.

Zalabardo, J. 1989. 'Rules, Communities and Judgements'. *Critica*, Vol. 21, No. 63, pp. 33-58.

————1997. 'Kripke's Normativity Argument'. In Miller & Wright, 2002, pp. 274-293.