

University of St Andrews



Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

SEMANTIC THEORY AND SENTENTIAL UNDERSTANDING

A discussion of tacit knowledge and compositionality

Thesis submitted for the degree of M.Litt

by

ALEXANDER MILLER

St. Andrews, April 1989



Th 1008

DECLARATIONS

(i) I, ALEXANDER MILLER, hereby certify that this thesis, which is approximately 30,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date. 27/4/89.. signature of candidate.

(ii) I was admitted as a candidate for the degree of M.Litt in October 1987. The higher study for which this is a record was carried out in the University of St. Andrews between May 1988 and April 1989.

date. 27/4/89....signature of candidate..

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of M.Litt in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date. 28/4/89. signature of supervisor.

(iv) In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

date. 27/4/89.....signature of candidate

ABSTRACT

This essay is an attempt to look critically at some issues arising from the idea that competent speakers of a natural language can *tacitly know* an extensionally correct theory of meaning for their language.

In the first chapter I outline the idea that sentential understanding is to be construed as an ability, and argue that a speaker's tacitly knowing a theorem of a semantic theory should be construed as his possessing an *intention* to uphold the regularity prescribed by that theorem. I then go on to deal with objections that have been raised against the idea that speakers can have such tacit knowledge by Baker and Hacker, Appiah, and McDowell.

In the second chapter I ask whether a similar story can be given concerning tacit knowledge of the axioms of a semantic theory, and come to the conclusion that it cannot : using modifications of arguments by Evans, Wright, and Stich, I show that speakers cannot be ascribed intentional states corresponding to semantic axioms. I then go back to argue that the ascription of such intentional states corresponding to theorems does not violate a plausible constraint on the ascription of such states.

In the third chapter I look at the prospects for the use of the notion of tacit knowledge in the justification of the compositionality constraint. Beginning with Evans, I eventually come to the conclusion that tacit knowledge of a semantic theory is a sort of causal explanatory structure, to be found in the abilities which constitute language mastery : this finds expression in Davies' version of the mirror constraint. *Inter alia* , I sketch responses to difficulties that have been pointed out concerning this account by Wright, Davies, and Sainsbury, but conclude that there are still at least some difficulties looming which stem from the later Wittgenstein's attack on the objectivity of meaning.

CONTENTS

INTRODUCTION (1 -2)

CHAPTER 1 : KNOWLEDGE OF SENTENCE MEANING

1.1 *Understanding sentences is a recognitional ability* (3 - 4)

Introduction of, and summary of arguments for, the idea that sentential understanding is to be construed as being at least akin to a recognitional ability. Importance of this construal for the debate between realists and anti-realists taken note of.

1.2 *Two sorts of knowledge ?* (4 - 9)

Discussion of the distinction between *knowledge that* and *knowledge how*. Wittgenstein's arguments concerning "a rule for interpreting a rule" used to discredit Dummett's construal of sentential understanding as implicit propositional knowledge. Consequences for e.g. Foster / Davidsonian lines on the role of knowledge in semantic theorising taken note of.

1.3 *Ability and Intention* (9 - 11)

Our account thus far leaves opaque the precise location of the mental aspect of language mastery. Introduction of intention as the fundamental psychological notion underlying the ascription of tacit knowledge in an attempt to remedy this.

1.4 *Some objections dealt with* (11 - 17)

Objections raised by Anthony Appiah, Baker and Hacker, and John McDowell, against the implicit knowledge / recognitional ability view of language mastery, are disposed of.

CHAPTER 2 : THE LIMITS OF INTENTION

2.1 *Tacit knowledge of axioms : inferential integration and insulation* (18 - 26)

Arguments from Steven Stich, Crispin Wright, and Gareth Evans, are corrected in such a way that they can plausibly be viewed as motivating a constraint which rules out the possibility of there being intentional states of speakers corresponding to semantic axioms : such states are merely causal states which play a part in the proximate causal history of the (arguably) intentional states of speakers corresponding to semantic axioms.

2.2 *The generality constraint* (26 - 27)

A modified version of this constraint (which was first formulated by Evans) is presented, the motivation for whose imposition is made clear, using the results of the previous section.

2.3 *Intentions and concepts* (27 - 37)

Discussion of the idea that speakers can possess intentions in the absence of the concepts necessary for a proper statement of their content, and its consequences for the claim that speakers can have intentional states corresponding to meaning delivering theorems. Wright's agnosticism on the issue not clearly consistent with remarks of his elsewhere (concerning his "extension-determining" thesis). Argument to the effect that there is no real tension between the idea and the account which we wish to offer, with reference again to the rule-following considerations in Wittgenstein.

CHAPTER 3 : COMPOSITIONALITY

3.1 *Wright's challenge and Evans' response* (38 – 42)

Outline of Wright's challenge to the semantic theorist to provide a justification of the compositionality constraint. Brief sketch of Evans' response to this challenge.

3.2 *Wright's criticisms of Evans and Davies defence thereof* (43 – 53)

Outline of the objections raised by Wright against Evans concerning the interdefinability of the dispositions which are thought to constitute tacit knowledge of semantic axioms, and the reflection of causal structure in semantic theories. Davies' response to these objections is, as it stands, ineffective ; alternative attempt at meeting Wright's objections is provided.

3.3 *The 5% difference* (53 - 59)

Outline of Wright's objection concerning the 5% difference, and a sketch of a response to that objection from Davies.

3.4 *The mirror constraint and a perspective on semantic facts* (59 – 63)

Outline of Fricker's Metaphysical Perspective on the nature of semantic facts, and suggested that the mirror constraint is a natural way of showing how semantic structure can meet the demands of the Metaphysical Perspective.

3.5 *The doctrine of essential linguistic structure* (63 – 71)

Argument to the effect that a semantic theory need not discern semantic structure met via an argument for the weak doctrine of essential linguistic structure : if a theory of meaning does not discern semantic structure then it will be indeterminate with respect to the assignment of meanings to sentences. Argument to the effect that the equation of semantic structure with causal explanatory structure leaves open the possibility that a language might not have a unitary semantic structure met via an argument for the strong doctrine of essential linguistic structure : that the causal structure of the ability which constitutes language mastery is one of its essential features.

3.6 *Causal structure, novel utterances, and learnability* (71 – 78)

The imposition of the mirror constraint can contribute towards an understanding of how speakers can understand novel utterances, and how natural languages can be learnable. Proof that a theory of meaning which satisfies the mirror constraint for a natural language can have only finitely many proper axioms.

3.7 *The structural constraint and some thoughts on computer programming* (78 – 82)

Comparison of the mirror constraint with the structural constraint and Wright's "computer program" suggestion.

CONCLUSION (83 – 85)

NOTES (86 – 94)

BIBLIOGRAPHY (95 – 96)

INTRODUCTION

Competent speakers of a natural language *know* what the sentences of that language *mean* ; they *understand* those sentences. A theory of meaning for a natural language, if correct, specifies what each well-formed declarative sentence of that language means. Thus, the following question naturally suggests itself : what sort of relationship, if any, obtains between the speakers of, and a correct theory of meaning for, a given natural language ? This essay is an attempt to look critically at some possible answers which might be given in response to this question.¹

Is there a difference in kind between the relationship that obtains between the *axioms* of the theory and the speakers, and the relationship between the *theorems* of the theory and those speakers ? Intuitively we expect the answer to this question to be negative, but according to some philosophers our inclination in this respect turns out to be misguided – the construal of the relationship between the theorems and the speakers is held to be the less problematic of the two.² So a good starting point, or strategy, for our enquiry would appear to be this : first, outline the proper relationship that can be said to obtain between the theorems of the semantic theory for a language and that language's speakers (i.e. between the rules governing the use of the language's *sentences* and the speakers) ; then attempt to carry the story obtained over into the case of the axiomatic base of the semantic theory. If this attempt succeeds, then our task is happily done. If, on the other hand, it ultimately ends in failure, we might at least be better placed to see why this failure has occurred, and as a consequence better placed to provide an account of the relationship which circumvents the problems to which those reasons give rise.

This way of proceeding seems desirable from another point of view. Too often in discussions of language, accounts of meaning have been offered to which accounts of understanding have subsequently had to be tailored , in my view this order of priorities is unacceptable. If either of the notions is primary it has to be the notion of understanding – only once we have achieved a correct model of the nature of our *understanding* of a language can we proceed to draw conclusions concerning the nature of meaning. This might seem like a reversal of the natural order of things, for knowledge of meaning, and understanding, are generally equated. But we would not ask for an account of what our knowledge of tables consists in, or how such knowledge is possible, independently of, or prior to, settling exactly what tables *are* . Accordingly, it seems to have been

thought that an account of meaning must be provided prior to an account of knowledge of meaning, and thus prior to an account of understanding. This claim concerning our order of priorities is, I think, of fundamental importance – and I also believe that the failure to perceive it has been at the root of much philosophical confusion concerning language. We should not proceed as follows – "*This* is what the meaning of a sentence *must* be – now, *how* is it possible for us to become acquainted with it ?". Rather, we should say "*This* is what understanding a sentence *is* – now, what can we say concerning its meaning ?". As Wittgenstein would have put it , "A simile that has been absorbed into the forms of our language produces a false appearance, and this disquiets us.'But *this* isn't how it is !' - we say.'Yet *this* is how it has to *be* !'."3

Our discussion thus begins with the topic of sentential understanding, an enquiry into what model of a speakers understanding of sentences is the correct one.

1.1 Understanding sentences is a recognitional ability

Since the work of the later Wittgenstein has become widely known, it has become common for philosophers of language to claim that sentential understanding cannot be construed as a certain occurrent mental state, or as the occurrence of certain mental processes, construed either as processes that can be brought to consciousness, or as occurring outwith consciousness in some "hypothetical mental mechanism". Instead, it is urged, we should view understanding as being at least akin to the possession of a distinctive *capacity* or *ability* – someone who understands a sentential expression has a capacity to use it in ways not shared by someone who lacks that understanding. That is to say, in the familiar Rylean terminology, understanding a sentence is more like *knowledge how* than *knowledge that*¹: someone who understands a sentence knows how to use it in accordance with its meaning and in a rule-governed manner, and *knowing how* cannot be construed as a particular conscious or unconscious mental state, for the reasons made familiar most forcibly by Wittgenstein in the *Philosophical Investigations* – namely, that the existence of such states is neither necessary nor sufficient for the correct ascription of understanding, and therefore cannot be constitutive of that understanding.²

So I shall take as my starting point this idea that understanding a sentence must be construed, at least in part, as the possession of a particular practical ability, skill, or capacity, in other words as the possession of *knowledge how*. The question which now arises is this: though we have described understanding a sentence as possession of a practical ability, is it possible to give any more detailed characterisation of what that understanding consists in? We have said that understanding is an ability, but can we give an informative account of what it is an ability to *do*? What sort of capacity or ability *is* sentential understanding?

The answer generally given to this question is that understanding is a *recognitional capacity* or *ability*; the ability, that is, to recognise that certain conditions associated with the sentences do or do not obtain, and to be suitably aware of the consequences thereof. As Crispin Wright has put it, "The understanding of any sentence consists in the possession of a distinctive concept, and...the possession of any concept consists in a capacity or network of capacities of discrimination"³, and "Knowledge of use is not essentially verbalisable knowledge. So if knowledge of declarative

sentence meaning is to be knowledge of truth conditions, we should not construe the latter as essentially articulate. Rather it must be a recognitional capacity : the ability to recognise, if appropriately placed, circumstances which do, or do not, fulfil the truth conditions of a sentence and to be prepared accordingly to assent to, or to withhold assent to, its assertion."⁴

The importance of thus construing sentential understanding shows up clearly in the dispute between realist and anti-realist conceptions of meaning. Roughly, the realists contend that understanding must be construed as an ability to recognise the *truth* conditions of sentences, while those of anti-realist persuasion would contend that the conditions, the ability to recognise which constitutes understanding, must be conditions under which the sentence in question is *warrantedly assertible* , and as such are subject to more stringent epistemic constraints than the realists' truth conditions. In this essay I shall not be concerned with the realist / anti-realist dispute in the theory of meaning, and accordingly the question as to the precise *nature* of the conditions ability to recognise which constitutes language mastery will be left on one side. Rather I shall be concerned, in the first instance, with criticisms of the construal of understanding as a recognitional ability which do not turn on any bones of contention between the realists and the anti-realists. Thus, for the sake of convenience, we can discuss the issues here in terms of truth conditions, whilst bearing in mind that any points made would have equal force if made in terms of conditions which warrant assertion, or any other candidate conditions.

Before looking at these criticisms however, I shall look briefly at some of the consequences the Wittgensteinian insights just mentioned have for Michael Dummett's use of the term "knowledge".

1.2 Two sorts of knowledge ?

When I speak of *knowledge that* I am speaking of what might be called propositional knowledge viz. knowledge that a given proposition is true. Consequently, the fact that a speaker X possesses propositional knowledge of a given proposition P can be described by a clause such as "X knows that P". In the first instance I shall take this notion of knowledge to be equivalent to what Michael Dummett (see below) terms "verbalisable knowledge", namely, knowledge which its possessor can formulate verbally or express linguistically. Thus, propositional knowledge that grass is green

will only be ascribable to Jones, when Jones can make it clear to us, by linguistic means, that he believes that grass actually is green, paradigmatically by uttering the sentence "Grass is green". In order to avoid the familiar problems from the theory of knowledge, I claim only that this is a necessary condition for the ascription of knowledge that P, and leave the demarcation of conditions jointly sufficient for its ascription to the epistemologists.

How does all of this carry over to the case of knowledge of meaning? What is involved in saying that "X knows that the meaning of S is p"? Following Dummett we shall construe verbalisable (or propositional) knowledge of meaning as "...knowledge which consists in the ability to state the rules in accordance with which the expression or symbol is used or the way in which it may be replaced by an equivalent expression or sequence of symbols"¹. Thus, we shall ascribe knowledge in this sense of the meaning of an expression only when the speaker in question is capable of giving an explicit account of the rules governing that expression's use, or is able to provide us with a verbal reformulation or linguistic paraphrase of that expression.² And as such, we stipulate that such knowledge can be termed *explicit* knowledge of a sentence's meaning.

What then of *knowledge how*? In the first instance I shall take the ascription of knowledge of how a sentence should be used to be no more and no less than the ascription of the practical ability possession of which is involved in mastery of the use of that sentence. It is clear that *knowledge how* in this sense need not imply the possession of *knowledge that* – if it is the case that I know how to use a given linguistic expression it need not be the case that I am able explicitly to state the rules for its correct use or able to provide a linguistic paraphrase of it. Thus, to get the terminology straight, if I understand a sentence P I can be said to know how to use that sentence, but it need not be the case (although it might) that I can be said to know that the meaning of the sentence is something or other in the sense of explicit or propositional knowledge. Thus this sort of knowledge is "...knowledge which does not reside in the capacity to state that which is known"³ – and since Dummett calls this sort of knowledge "implicit knowledge", I shall do so too.

Can we have a parallel notion of implicit knowledge (or belief) in the cases where we are not concerned with knowledge of meaning, or with the possession of a practical skill? – can we say that someone implicitly knows or believes that a certain state of affairs obtains? I think we can, and frequently do, say this sort of thing; for instance, we can say that Jones implicitly believes that his next door neighbour's dog is vicious, in the case where, although Jones has never uttered that

sentence or any of its synonyms, he takes great pains to get out of the dog's way whenever he sees it approaching. Thus I suggest that we can say that a speaker has implicit knowledge or belief concerning a proposition in the case where, although that belief is unarticulated, or perhaps inarticulable, linguistically, by means of that speaker uttering a sentence which expresses that proposition, it is clear from his other linguistic acts or his other non linguistic behaviour that he actually does believe that the proposition is true, or where that behaviour together with his other (explicit) propositional attitudes licences the ascription of that belief, given the usual holistic constraints.

Thus, on the one hand we have *knowledge that* = propositional knowledge = explicit knowledge, while on the other we have *knowledge how* = implicit knowledge (in the case of knowledge of meaning and more generally in the possession of practical skills), and implicit knowledge = unarticulated or inarticulable *knowledge that* (in the non practical cases). So we have two distinct senses in which a speaker can be said to be the possessor of implicit *knowledge that*, or implicit propositional knowledge. On the one hand, where it seems sensible to say that that speaker knows or believes that a given proposition is true, but never actually gives vent to this knowledge or belief in language; on the other, where he is the possessor of a practical skill or ability.

I take it that each of these notions is, *prima facie*, philosophically unobjectionable; as far as the former case goes it is a matter for debate whether we can have unconscious beliefs, but I think that the onus of proof is on those who say that we cannot⁴; in the latter case we have simply stipulated that ascription of implicit *knowledge that* is ascription of a practical ability⁵, and since there is no suggestion that in this case the speaker possesses anything more than that practical ability, our characterisation as it stands can hardly be objected to.⁶

It is important to note that in the way I have set this up the two sorts of knowledge need not be everywhere mutually exclusive, i.e. it is possible to have both explicit and implicit knowledge concerning the meaning of a given sentence P – although it is clearly not possible to have both explicit and implicit propositional knowledge where this is construed in the former of the two senses above. Someone can have implicit knowledge of the meaning of P but not explicit knowledge (in the case where he is able to use it in the appropriate ways but is unable to formulate the rules governing its use or to produce a linguistic synonym), and he can also have implicit

knowledge together with explicit knowledge (in the case where he is both able to use it and able to give that statement of rules or produce the synonym). In both of these cases we can say that the speaker understands the expression, simply because he possesses the requisite sort of *knowledge how*. But is it possible for a speaker to have explicit knowledge that the meaning of P is such and such without possessing the appropriate implicit knowledge? The answer to this is that it is possible, but only in the unusual case where he can state the rules and give paraphrases but is simply unable to put the expression to its proper use, and as such cannot be credited with understanding. Thus, because of our terminology, we have the apparently paradoxical situation that someone might possess explicit knowledge of a sentence's meaning although he does not actually understand that sentence. In fact this is only an *appearance* of paradox, a mere reflection of the fact that *knowledge how* is the more fundamental notion of the two.

Thus, in summary, the situation is as follows. Take any speaker S and any sentence P. Then, either S understands P or he does not. In the former case S might have implicit but not explicit knowledge of the meaning of P, or he might have both. In the latter case, he might — unusually — have explicit but not implicit knowledge of P's meaning, or — the more usual case — he might have neither.

This way of looking at things differs somewhat from Dummett's, because he writes as if someone who possessed explicit knowledge of meaning but not implicit knowledge might still be credited with understanding. For instance, he says that "...understanding plainly cannot in general consist in the ability to find a synonymous expression"⁷, the suggestion of which is that it sometimes can. This is what I deny, and this denial renders Dummett's claim that understanding cannot everywhere consist of explicit knowledge of meaning somewhat vacuous, for the simple fact is that it cannot *anywhere* consist of such knowledge alone.

This points to the connection between Ryle's arguments and the passages in Wittgenstein's *Investigations* leading to the "paradox" of §201. Wittgenstein's negative claim in these sections is that interpretations by themselves do not determine meaning, where interpreting a rule is conceived of as understanding a *representation*, *expression*, or *translation* of that rule. Now the most plausible candidate for *knowledge that* which would suffice for understanding would presumably be propositional knowledge of a sentence which provided a translation or stated a synonym of the original expression — and Wittgenstein's argument is precisely that such knowledge would not suffice for understanding, because the interpretation of the original sentence *itself* stands in need

of interpretation. Wittgenstein's argument can thus be viewed (given the lack of any other proposed species of propositional knowledge which would turn the trick) as an argument to the effect that propositional knowledge is never by itself sufficient for the possession of sentential understanding.

It may seem that I have been unfair to Dummett here inasmuch as I have taken his notion of explicit knowledge to be a *mere* ability to state rules or to provide verbal reformulations – couldn't he rather define explicit knowledge as the possession of that ability *together* with possession of the recognitional capacity which constitutes mastery of the sentence use? This would then leave the claim that understanding cannot everywhere consist in the possession of explicit knowledge with some substantial content.

However I think that that way of characterising explicit knowledge of sentential meaning is unsatisfactory. Construing it thus either has the effect that *knowledge that* simply becomes a species of *knowledge how* – which is plainly at odds with the reasons for introducing the distinction between the two sorts of knowledge in the first place – or that, in an attempt to avoid that undesirable consequence, we must refuse to equate explicit knowledge of the meaning with *knowledge that* the meaning is such and such. But this latter course of action is not much better than the former alternative because we are then left wondering precisely what explicit knowledge is supposed to be – if it cannot be construed as *knowledge that*, what is it?

My way of setting up the implicit / explicit knowledge distinction may be claimed to be objectionable on another score. Some semantic theorists have claimed that what we are after in a theory of meaning is a theory explicit knowledge of which *would* suffice for mastery of the language.⁸ On my present way of construing the notion of explicit knowledge this does not in fact turn out to be a possibility – explicit knowledge of such a theory would not suffice for understanding the language because someone apprised of that knowledge would have to apply it, and apply it correctly, in order to be deemed an understander of the language; and knowing how to do that would seem simply to be to possess the recognitional abilities which constitute mastery of the language, that is, implicitly to know the meanings of its constituent expressions. So it is simply false that a theory of meaning states something explicit knowledge of which *would* suffice for understanding the language, and to that extent I have to part company with Foster, McDowell, and Davidson. But again there is no good objection to my construal here – the price of retaining the rejected conception of what a theory of meaning is would be making *knowledge that* a species of

knowledge how, or changing the characterisation of explicit knowledge, both of which routes we have already found to be unacceptable.

1.3 Ability and Intention

In sections 1.1 and 1.2 we stated an account according to which a speaker's understanding of a given sentence is constituted by his possession of a certain recognitional ability. I want now to argue that as it stands this account is not wholly satisfactory, for the reason that it appears to leave out (or rather that it is not clear how it includes) something essential to any account of the nature of language mastery, namely, that such an account should represent the participants in the practice of speaking the language as *intentional* agents, and thus reserve a place for the presence of the *mental* aspect of language use.

While the account adumbrated in 1.1 and 1.2 may not be directly inimical to such a representation, it does leave somewhat opaque precisely where the mental aspect of language use is located: it seems that the ascription of an ability to something will not by itself suffice to confer the status of an "intentional agent" upon it. A few examples should be sufficient to establish this point: certain small fish can be said to possess an ability to swim, amoeba are said to possess an ability to divide into smaller amoeba, and pocket calculators are said to be able to compute the answers to arithmetical and statistical problems – but in none of these cases do we want to describe the creature or system in question as an intentional agent or system. Notice also that there is (at least to my ears) something odd about saying that the fish, for instance, *know how* to swim; this veering away from use of the intentional term "knows" in these cases backs up the point just made: the possession of an ability is not generally sufficient to guarantee the involvement of mind, or the mental, in a given practice, be it linguistic or otherwise.

How then can we, as advocates of the account of 1.1 and 1.2, avoid the relapse into some crude and philistine form of behaviourism about the use of language? Are we thus committed (or at least leaving ourselves open) to the view that there is, strictly speaking, nothing more to the use of language than the audible and visible events which make up speakers' overt linguistic behaviour?

This point connects up with a demand which, to my knowledge, was first made by W.V.O

Quine.¹ We want to describe language as being, fundamentally, a rule-governed activity. The exercise of the abilities which speakers of a language possess are constrained normatively by these rules ; the rules are viewed as *guiding* the behaviour of those competent with the language. Quine's problem was this : what content can we give to the distinction between a rule, which guides and normatively constrains the behaviour of a creature, and an empirical generalisation which describes, or *fits* , that very same behaviour ? If no substance can be afforded to the former save what is afforded to the latter then again the threat of behaviourism looms : we do not think of the complex biological story which describes the amoeba's splitting in two as being a rule which guides and normatively constrains that amoeba's behaviour. What is the difference between a speaker of a natural language and the amoeba in this respect ?

The solution to this difficulty is, I think, as follows. The behaviour of the speaker and the amoeba are both to be subsumed under generalisations, and they both exhibit definite patterns of regularity. The crucial difference is that the language user possesses certain *intentions* , to the effect that his behaviour will *conform* to those regularities, that it will *serve to uphold* those regularities. In the case of our language user who understands sentence S, the content of the intention possessed is that the regularity prescribed in the T-theorem which states that S has such and such a truth (assertibility) condition will be upheld or conformed to.

Of course, this is not so much a solution to Quine's difficulty, rather the reappearance of that difficulty in a different (but usefully more sharply delineated) guise : how is it possible for anything anywhere to possess intentions ? In virtue of what, if anything, do they possess them ? What entitles us to say that the speaker has intentions but not the amoeba ? How are the truth-values of ascriptions of intention (1st person or otherwise) determined ? What sort of states (if indeed there are any) confer truth and falsity upon them ? We shall return to at least some of these fundamentally important questions in due course. For the moment I want to concentrate on the following, merely taxonomical, question : how can we supplement the account of understanding as possession of recognitional abilities in such a way that the speaker's possession of the requisite intentions is made clear ?

Speaking of the recognitional ability ascribed to a competent understander of a sentence, Crispin Wright says,

"There is no hiatus between this ability and knowledge of the sentence's meaning ; to attribute to someone the ability to recognise as such circumstances which do, or do not, fulfil

the truth-conditions of a sentence – however problematic the criteria for such an attribution – *is* to attribute to him understanding of the sentence."²

Obviously Wright too wants to avoid the behaviouristic threat we took note of earlier, so I take it that he wants to find a place for the needed intentions *within* the characterisation of the recognitional ability. On such an account, possession of some sorts of ability would qualify us as intentional agents, as possessors of minds ; of others, not. My challenge to Wright is : how are the two sorts of ability to be distinguished ?

Or perhaps it is simply misleading to speak of two sorts of *ability* here – we would not describe gold as having an ability to remain solid when immersed in water, but merely as having a *disposition* to do so. So what we are actually after here is an account of what the distinction between abilities and mere dispositions comes to – how do abilities involve intentional states in ways not shared by mere dispositions ? I will not attempt to answer this in detail here, but content myself with the characterisation of abilities as some sort of "intentional dispositions" : someone who has an ability to X has a disposition to X which is guided by certain of his intentions, beliefs , and desires.

So I think we can summarise the position reached as follows. To say that S implicitly knows a given T-theorem of a materially correct semantic theory for his language is to say that S has an intention to uphold the regularity prescribed by that T theorem, which partially constitutes the recognitional ability he possesses which enables him to behave in such a way that the constraints imposed by the T-theorem in question are respected. Ascription of implicit knowledge is in this sense ascription of an ability *via* the ascription of intentions, and we can say that a speaker understands in virtue of that implicit knowledge, because he understands in virtue of his possession of the specified intentions.³

1.4 Some objections dealt with.

As I said in the previous sections, sentential understanding is essentially *knowledge how*, implicit knowledge of sentences' meanings, and it is not sufficient for the possession of understanding that a speaker be apprised of a particular body of propositional knowledge. In this

section I outline some criticisms that have been made against the account which I have adumbrated ; first, some that are made by Anthony Appiah in his book *For Truth in Semantics* ; secondly, one made by Baker and Hacker in their book *Language, Sense, and Nonsense* ; and lastly, one made by John McDowell, in his paper "In defence of Modesty".

Anthony Appiah says "I think we should treat knowledge of meaning as theoretical knowledge, 'knowledge that', rather than 'knowledge how' ", and "My view is not that knowledge of meaning is just 'conveniently' represented as propositional ; I think it is propositional."¹

Appiah's argument for this contention goes as follows : generally, in giving an explanation of why a speaker S uttered a sentence P on a given occasion, we have to advert to the fact that that speaker has certain beliefs, of which he may or may not be conscious. Appiah gives an example to illustrate his point – I utter "There is a car we can use", and the fact that I made this utterance is explicable only if certain beliefs are attributed to me e.g. I believe that if I say this you will come to believe that there is a car we can use, I believe that my saying it will induce in you, provided you think I am honest, the belief that I believe that there is a car we can use, and so on. Now with this much I am not in disagreement – what I do deny is that it is in virtue of *those* beliefs that the speaker understands the utterance in the way he does, for his possession of those beliefs is perfectly compatible with his lacking the abilities constitutive of understanding. Appiah appears to have illegitimately equated two distinct enterprises – on the one hand that of giving a *rationalistic explanation* of why a speaker uttered a sentence P, and on the other, that of giving an account of what his understanding of P consists in. Provision of the former does not imply provision of the latter. Following the example cited above Appiah says that "Nothing in this explanation involves ascribing to me beliefs about what the sentence 'There is a car we can use' means ". Now I think we can say two things about this. First, I'm not at all sure that it is true – isn't an essential part of the rationalistic explanation that I understand "There is a car that we can use', that I realise that the utterance of that sentence is appropriate to the situation at hand, that I have an ability to discriminate those situations in which the utterance of the sentence is appropriate from those in which it is not ? I would say that e.g, the ascription of the belief that "if I say this you will believe that there is a car we can use" is intelligible only if I am credited with an understanding of the sentence "There is a car we can use". Secondly, even if it were true this would not show that understanding could be construed as the possession of propositional knowledge – the explanation (allegedly) provided was not an attempt to provide an account of what my understanding of "There is a car we can use"

consists in. If it were, we *would* have to refer to the ability that I have to use that sentence, and thus, on my account, to the implicit knowledge I have of its meaning.

Appiah describes his position as "simple", and he summarises it thus : "Speakers have certain beliefs in virtue of which they understand languages."² Now if he means implicit beliefs in the sense of recognitional abilities then we are not in disagreement, except that I would say that his description of this knowledge as propositional is at best misleading. If he means implicit beliefs in the sense of unarticulated beliefs (corresponding to the former of the two senses of implicit knowledge that were distinguished in the previous section) then we are – the possession of such beliefs does not constitute understanding, although the possession of them by a speaker may be a necessary condition for his understanding the sentence in the way he does, or for the viability of our interpretation of what it was that he said.

Now on to a criticism made by Baker and Hacker. They claim that Dummett's contention that a theory of meaning provides a theoretical representation of a practical ability rests upon the quite faulty assertion that "...to understand a sentence is to assign (or to be able to assign (?)) to the sentence a set of truth conditions".³ It should now be relatively clear that the attribution of this idea either to Dummett or Wright is quite unwarranted (since they clearly see the possession of recognitional abilities as basic), and also that the view they attribute to Dummett is in fact itself in direct conflict with some of the points that I made in my exposition above.

In our account the ability to assign truth conditions to sentences would amount to possession of *explicit* knowledge of their meanings – this is not what I take to be constitutive of linguistic competence, and I have provided arguments in the previous sections as to why this should be so. Competence is constituted by the possession recognitional abilities, abilities to discriminate between different sorts of condition and to respond accordingly to those conditions by effecting certain linguistic acts – whatever this ability is I think it is at best highly misleading to describe it as an ability to "assign" truth conditions to sentences, and as a result of this gross inaccuracy the remainder of Baker and Hacker's attack on the claim that theory of meaning is a theoretical representation of a practical ability loses much, if not all, of its plausibility.⁴

They suggest that proponents of the Dummettian view rely on an analogy between the mastery of language and the ability to perform arithmetical computations. The axioms and recursion clauses of an arithmetical theory are supposed to represent the latter, because by means of them we can

produce an answer to an arbitrary arithmetical problem – and this because the mastery of arithmetic is manifested in the production of answers to those very problems. The point is then made that "There is no such thing as an act of assigning truth conditions to a sentence which stands to mastery of a language as giving an answer to a multiplication problem stands to mastery of elementary arithmetic. The analogy is altogether lame."⁵ With this much I agree – I disagree only that someone construing understanding as a recognitional ability need have recourse to any such analogy. If Baker and Hacker had really been prepared to give their opponents a run for their money here they would surely have concentrated instead on the example, often used by Dummett and Wright (who as far as I can see never use the arithmetical analogy in this context), of the case of the ability to play chess. Mastery of the game surely can here be plausibly construed as possession of recognitional abilities, viz. the abilities to recognise that certain board configurations obtain and to respond accordingly to these configurations by making, or by refraining from making, certain moves. There is such a thing as an act of recognising that the truth conditions of a sentence obtain and realising the consequences thereof (prescinding from the familiar anti-realist worries) which *does* stand to mastery of a language, as recognising that certain board configurations obtain and realising the consequences thereof stands to mastery of the game of chess – and this analogy is *not* altogether lame.

I now move on to consider an objection that has been levelled against the notion of implicit knowledge, or implicit guidance by rules, by John McDowell.⁶ He sums the objection up himself thus :

"Consider Dummett's suggested explanation of the concept 'square'. Perhaps a speaker of English might be induced to acknowledge it as correct. But the explanation contains uses of the word 'square'. So comprehension of it as an object of acknowledgement would be an exercise of the very capacity that we were trying to see as guided by the implicit knowledge that the acknowledgement supposedly reveals. If we needed guidance in our overt practice we should need it just as much in our understanding of the supposed guide".⁷

We say that X implicitly knows the rule R, or has his practice guided implicitly by R, when X has an intention to uphold the regularity prescribed by R, together with a recognitional ability that enables him to conform his behaviour to the content of that intention – and generally X will be unable to formulate R linguistically. Now McDowell's argument seems to turn on the claim that a speaker cannot be viewed as implicitly following a rule in his use of an expression E where an

account of the content of that intention would involve uses of the expression E itself – the thought being that in order to have the intention I would have to understand the expression E anyway. How can my use of E be guided by the intention if the explanation of the content of that intention involves use of E itself?

I think the key to the rejection of McDowell's argument here lies in Wittgenstein's remarks concerning rule-following and interpretation. I suggest that McDowell's argument can be successful only if he makes the following assumption about the nature of rule-following: that following a rule R, or intending to uphold a regularity R, always consists in *interpreting* R – where interpreting R is conceived of as understanding a representation or translation of R and on the basis of that understanding deciding whether or not a given application is in fact warranted by the rule. If I have to interpret R in order to follow it, or in order to intend to uphold the regularity which it prescribes, then that accomplishment will indeed be rendered mysterious if whatever functions as the rule contains appearances of the concept whose usage it is supposed to be guiding; but of course the fundamental insight which Wittgenstein gives us concerning the nature of the rule-followers accomplishment is precisely the opposite of this, namely, "that there is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call 'obeying the rule' and 'going against it' in actual cases."⁸ Eventually we reach a level where the speaker is still following a rule, but where the correctness or otherwise of his applications of this rule cannot now be assessed by reference to some other rule or set of rules – and because no interpretation is taking place at this level it matters not that what functions as the rule itself uses the very concept whose usage it is supposed to govern.⁹

Crispin Wright has suggested a different way of rebuffing McDowell's objection here.¹⁰ He suggests that we can draw a distinction between, on the one hand, a statement of a rule R, and on the other, that which actually functions as the rule R, and he seems to accept McDowell's claim that an explanation such as the one proffered by Dummett cannot *function* as a rule while claiming that it is nevertheless a *statement* of the constraints imposed by the rule. This would apply to the homophonic axioms and theorems of a semantic theory as well – Wright accepts that they cannot be what function as rules, but that they can nevertheless still be statements of the constraints imposed by those rules. Indeed, in addition to accepting the gist of the foregoing argument of McDowell's he provides another¹¹: the homophonic axioms and theorems cannot be what function

as the rules because they do not carry normative force in such a way that we can settle disputes by reference to them, for any disputes that arise simply become transformed into disputes about the contents of the rules themselves. For example, the axiom "x satisfies 'red' iff x is red" could not be used effectively to settle a dispute over whether a given application of the predicate "red" was correct or not – we would be left with another dispute arising from the interpretation of the RHS of the axiom. For example, suppose that some deviant points to the grass and utters "Red". Then it will be of no avail to attempt to persuade him of the error of his ways by invoking the homophonic axiom above, because the disagreement will simply reappear as a disagreement about the correct application of its RHS.

I am inclined to reject Wright's diagnosis here – he gives us no clue as to what *does* function as the rule in the cases considered, and says only that it cannot be the homophonic axioms and theorems of a semantic theory. So at best we are left with the following unanswered question : what *does* function as the rule ?

Suppose that what *does* function as the rule admits of at least some linguistic expression. Then I suggest that Wright faces the following dilemma. Either he attempts to get non-homophony everywhere in the axioms and theorems – in which case he seems to commit himself to the view that his semantic theory will not contain only finitely many semantic primitives, a view with at least prima facie uncomfortable consequences for the learnability of language¹² – or, in order to avoid this, he accepts that we are going to have homophony at some level, and that consequently is going to come up against the original problem again anyway.

So far as the point about the settling of disputes is concerned I think we can say the following. Another of the upshots of Wittgenstein's rule-following considerations is that we are inevitably going to reach a level at which disputes concerning the applications of a given rule are not settleable by recourse to any further rules – the disputants will simply disagree about which applications of the given rule are correct, depending on their respective irrational and wholly subjective inclinations to respond to the existing perceptual conditions in certain ways : and this is precisely the shape the dispute appears to take in the cases in which what functions as the rule is a homophonic axiom or theorem.¹³

In summary, then, acceptance of Wright's distinction leaves us with the following options. Either we are left completely in the dark as regards what do function as the rules in question, or we accept that what function as the rules admit of linguistic expression and either (a) finish with a

theory globally non-homophonic but with an infinite number of semantic primitives, or (b) accept that homophony must inevitably enter the scene at some stage anyway. My view is that we should perhaps take note of Wright's distinction but opt, unlike Wright, for (b) – a semantic theory is inevitably going to contain homophonic axioms or theorems, but given the Wittgensteinian insight alluded to earlier I suggest, pace McDowell, that this provides us with no genuine cause for concern.

THE LIMITS OF INTENTION

The question of the limits of intention is a difficult one to answer. It is a question that has been asked in many different ways, and the answers have been equally diverse. Some have argued that intention is limited by the physical world, while others have argued that it is limited by the mind. The question is still open, and it is one that deserves further investigation.

2.1 Tacit knowledge of axioms : inferential insulation and integration.

In the first chapter we saw that an essential part in the location of the mental aspect of language mastery was the ascription of intentions to those competent with the language, namely, intentions to uphold the regularities prescribed by the theorems of an extensionally acceptable semantic theory for that language ; and I alluded, in note 3 to section 1.3, to the fundamental difficulty which this account apparently faces. I will address this difficulty later in this chapter ; before doing so, however, I shall discuss a difficulty that has been raised, in particular by Garth Evans and Crispin Wright, concerning the claim that speakers can have genuine intentions to uphold the regularities prescribed by a semantic theory's *axiomatic base* .

To begin by summarising the Evans / Wright objection.¹ They claim that there is a necessary condition which all genuine intentional states must satisfy to qualify as such, and that putative states of "tacit knowledge" of meaning-theoretic axioms do not satisfy this condition. In order to render the imposition of this condition plausible they ask us to contrast, on the one hand, the belief that a man might have to the effect that a certain substance is poisonous, with the disposition that a rat might have to avoid a similarly contaminated substance. Can we describe the rat as having a genuine belief that the substance is poisonous ? They suggest not : for whereas in the case of the man the belief is, to use Evans' phrase, "at the service of many distinct projects", and can interact with others of his beliefs and desires to produce new beliefs and desires, none of this obtains in the case of the rat's disposition. In the case of the man, for instance, the belief could be at the service of projects such as killing an adversary, retaining maximal good health, and getting out of an obligation, to name but a few. None of this is possible in the case of the rat : the putative "belief" is harnessed to the single "project" of avoidance of the substance which has, we might suppose, caused it some distress in the past. All this is supposed to be a reflection of the fact that propositional attitudes and intentional states, such as beliefs and desires, come in *articulated systems* or *holistic networks* . And it is because genuine beliefs come in such networks and can thus interact with other beliefs, that they can indeed be at the service of many distinct projects (e.g. the man's belief that the substance is poisonous can be at the service of the project of getting out of a particular obligation because that belief can, together with the beliefs that a small amount of the substance causes only a small amount of harm and that a slight illness

will assuage the demand that the obligation be fulfilled, lead to the belief that taking a small amount will enable him to avoid fulfilling the obligation).

We can thus suggest the following rather crude version of the constraint suggested by the Evans / Wright discussion :

Constraint 1 : A state P of an agent W is a genuine propositional attitude or intentional state *only if* P can interact with others of W's propositional attitudes and intentional states to produce new propositional attitudes – thus putting P at the service of many distinct projects of the agent.²

Where does this leave the "tacit knowledge" a speaker might have of a meaning theoretic axiom ? Evans and Wright both claim that such a state of a speaker violates the constraint above. Far from being at the service of many distinct projects the "tacit knowledge" is, says Evans,

"...exclusively manifested in speaking and understanding a language ; the information is not even potentially at the service of any other project of the agent, nor can it interact with any other beliefs of the agent (whether genuine beliefs or other tacit 'beliefs') to yield further beliefs."³

While Wright puts it like this :

"...the (implicit) knowledge of a meaning theoretic *axiom* would seem to be harnessed to the single project of forming beliefs about the content of sentences which contain the expression, or exemplify the mode of construction, which it concerns",

and he asks the following (rhetorical) question :

"...what is supposed to be the role of *desire* ? What is the (implicit ?) desire which explains why the subject puts his axiomatic beliefs to just this use, and what are the different uses to which they might be put if his desires were different ?"⁴

Much the same point has been made by Stephen Stich.⁵ He draws a distinction between, on the one hand, genuine belief states, and on the other, what he calls *subdoxastic states* : that is, cognitive states which represent information – which have informational content – and which play a part in the proximate causal history of beliefs, but which are not in fact genuine beliefs themselves.

Stich suggests a necessary condition on a cognitive state's counting as a genuine belief which is almost identical to the constraint extracted from the Evans–Wright discussion above. This uses the notions of *inferential integration* and *insulation* .Whereas genuine beliefs form inferentially integrated subsystems of an agent's cognitive states (in the sense that there are *many* inferential

routes, both deductive and inductive, from a given belief to almost any other belief), subdoxastic states, in contrast, are inferentially insulated from the vast majority of the subject's genuine beliefs (in the sense that the subdoxastic state will be linked inferentially to only a very limited range of the agent's beliefs).

Whereas Stich sets his necessary condition up as a necessary condition on beliefs rather than as a condition on intentional states generally, the extension of that condition to the more general case is unproblematic, so long as we are willing to allow that there can be inferential relations between propositional attitudes other than belief. Without expending too much time on arguing for this I think we can say that it is an intuitively acceptable enough move. Take desire for instance. Initially it seems odd to speak as if a desire could function as a premise, or be the conclusion of, a process of inference ; but this oddness should vanish on reflection upon some simple examples e.g. if I have a belief that that book is interesting, and a desire to read something interesting which together produce the desire to read that particular book, then this can I think harmlessly be described as an inferential process. (In fact, this is just the old idea of practical inference).

While Evans appears to view the objection as applying to tacit knowledge *tout court* , that is, not only to states of "tacit knowledge" of axioms, but also to states of "tacit knowledge" of the rules governing sentences, Wright quite clearly sees his objection as applying only to "tacit knowledge" of the axioms – as he says, "someone who is credited with implicit knowledge of a meaning-delivering theorem may express his knowledge in an indefinite variety of ways, including, in appropriate contexts, lying, assent, and silence", so that no reason emerges "to doubt the propriety of crediting them [speakers] with implicit knowledge of the content of the meaning-delivering *theorems* . "6 Thus, Wright's defence of genuinely intentional tacit knowledge of meaning-specifying theorems is in effect a claim that tacit knowledge of such theorems *can* be inferentially integrated with the rest of the agent's propositional attitudes in the manner required by Stich. How can my semantic belief concerning the meaning of a given sentence interact with my propositional attitudes to give rise to new propositional attitudes ? A few examples should suffice to convince us that this is indeed possible. I have a certain intentional state the possession of which is constitutive of my understanding of the sentence "The pillar-box is red", and this intentional state can interact with e.g. my belief that there are pillar-boxes in the High-Street to produce the belief that there are some red things in the High Street. Similarly, the intentional state constitutive of my understanding of " $2 + 2 = 4$ " can combine (in the sense that I can deploy the information contained

therein) with the belief that there are two keys and two pennies in my pocket to form the belief that there are at least four things in my pocket. Examples can be quite easily multiplied – and this shows how, apparently unlike the putative implicit knowledge of axioms, the implicit knowledge of semantic theorems is indeed at the service of many distinct projects.

Let's suppose, as seems plausible, that Wright is correct in his insistence that the particular objection just considered does not apply to states of "tacit knowledge" of the meaning delivering theorems, for the reason stated. Then my question is simply this : why cannot we view the "tacit knowledge" of an axiom as a genuine intentional state, in virtue of the fact that although it is *directly* harnessed to the single project of forming beliefs about the content of sentences in which it figures, it can make an *indirect* contribution to the other projects of the agent *via* the states of "tacit knowledge" of the theorems corresponding to the sentences in which it figures ? The state of "tacit knowledge" of an axiom is at the service of many of the agent's projects because the project to which it is *directly* harnessed is itself at their service. Or equivalently, a state of tacit knowledge of an axiom can lead inferentially to a vast number of other propositional attitudes because it *can* lead inferentially to genuine intentional states (i.e. those that consist in implicit knowledge of the appropriate theorems), which *in turn* can lead to almost any other intentional state, *modulo* the other intentional states which we suppose the agent to possess.

So, it is not the *number* of intentional states to which a given cognitive state can give rise, or the *number* of the agent's projects which the information represented by such a state is at the service of, which are important in drawing the distinction between genuinely intentional and merely subdoxastic states ; rather, it is the nature of the *route* from the given state to the rest of the propositional attitudes, and the nature of the *route* via which the information in question is placed at the service of a multiplicity of the agent's projects, that are crucial.

I suggest, then, that in order to draw the required distinction in such a way that states of tacit knowledge of axioms are excluded, we need something along the lines of the following amended version of constraint 1 :

Constraint 2 : A state P of an agent W is a genuine propositional attitude or intentional state *only if* P can interact *directly* with others of W's propositional attitudes and intentional states to produce new propositional attitudes – thus putting P *immediately* at the service of many distinct projects of the agent.⁷

Of course, the questions facing us now are (i) What exactly do we mean by "directly" as it appears in our statement of the revised constraint ?, and (ii) Does the constraint provide us with a plausible means of drawing the distinction between intentional and non-intentional states ?

Let's attempt to answer the former question first. We can say that *a cognitive state P interacts directly with a given propositional attitude only if the interaction can generally take place without the mediation of some other propositional attitude R in the causal generation of which P plays a part*. Seeing why states of tacit knowledge of axioms fail to satisfy the constraint will help us obtain a more secure grip on this notion of directness. Such a state can interact with other states only via the states of tacit knowledge concerning the sentences in which the expression corresponding to the axiom figures ; this is because of Frege's context principle – it is always by means of complete sentences that we perform linguistic acts, or, more figuratively, make moves in a given language game : a speaker's understanding of a subsentential expression can be manifested only through the use that he makes of *whole* sentences in which that expression figures.

Thus, suppose that there is a cognitive state of mine which represents the information that a given predicate e.g. "horse" has such and such satisfaction conditions. Suppose also that I hear someone saying that he has a horse with five legs. Then I might form the belief either that the person in question simply has an understanding of the predicate which differs from mine, or that he has a very rare and unusual sort of horse : but these beliefs can be formed *only* via my implicit knowledge of the meaning of the whole sentence "I have a horse with five legs", because it is only in the context of a whole sentence that a linguistic act involving the predicate can be effected.

So, tacit knowledge of an axiom can never (with one exception) interact directly with other putative intentional states, because it always has to interact with them via the states of tacit knowledge of the appropriate theorems, in whose causal generation it plays a part. The one exception is of course it's interaction with the states of tacit knowledge of the theorems themselves – but this is not sufficient on its own for the satisfaction of the constraint.

We should perhaps note that Stich appears to see his objection concerning inferential integration and insulation as applying to tacit knowledge of *theorems* as well as to tacit knowledge of the axiomatic base. He makes the following remarks concerning the states underlying our ability to discriminate between grammatical and ungrammatical sentences, but they might equally have been directed against semantic as well as syntactic competence :

"...suppose that, for some putative rule *r*, you have come to believe that if *r* then Chomsky is seriously mistaken. Suppose further that, as it happens, *r* is in fact among the rules stored by your language processing mechanism. That belief along with the subdoxastic state will not lead to the belief that Chomsky is seriously mistaken. By contrast, if you believe (perhaps even mistakenly) that *r*, then the belief that Chomsky is seriously mistaken is likely to be inferred. It would be easy enough to marshal many more illustrations of the fact that the subdoxastic states which store grammatical information are largely inferentially isolated from beliefs."⁸

I think we can make a number of points about this example. Firstly, I am rather suspicious of the idea that we can credit a speaker with a genuine belief that if *r* then Chomsky is wrong without crediting him with at least the conceptual resources needed to form a belief to the effect that *r*. How can I possibly have a genuine belief in a conditional, when I do not possess conceptual resources sufficient for the formation of a belief in the antecedent? So – on Stich's assumption that we need the belief that *r* in order to get the inference underway – it would be *possible* for someone in the sort of situation which Stich envisages to move to the belief that Chomsky is mistaken, although he might not in fact do so. In this respect the situation does not differ from that in which we have two genuine beliefs – it is far from clear that Stich's claim concerning the case in which each state involved in the example is a genuine belief is itself plausible (that where I have the *belief* that *r* and the *belief* that if *r* then Chomsky is wrong I will typically draw the conclusion that Chomsky is wrong): it would require for its tenability some assumption to the effect that a speaker is likely to draw out the believed logical consequences of his beliefs, an assumption which Stich himself explicitly repudiates a few pages earlier.⁹

Secondly, suppose that it were impossible for the speaker in the first example to move to the belief that Chomsky is mistaken. What would the explanation of this consist in? Is it supposed to turn on the fact that *r* is (currently?) inaccessible to consciousness? Stich has already said that he is willing to allow unconscious inference (he quite explicitly calls the relationship between a subdoxastic state and the belief to which it gives rise a case of inference), and in any case it seems clear from his paper that he (in my view, quite rightly) sees inferential integration to be a more fundamental factor in drawing the distinction than accessibility to consciousness.

Thirdly, what is important is not whether the speaker explicitly believes that *r*; rather, what is crucial is whether the speaker can have a genuine intention to conform to *r*, and thus whether he can employ the information encoded within in the production of new intentional states by the

exercise of the ability partially constituted by his possession of that intention (in the direct manner required by our constraint). Stich has provided us with no argument to the effect that this is impossible.

So : that is our proposed constraint, which implicit knowledge of axioms fails to satisfy because of the Fregean context principle, and which does not appear to rule out implicit knowledge of the truth-theoretic output of a semantic theory. The pressing question now of course concerns its plausibility : is the fact that a given state fails to satisfy it good grounds for refusing to describe that state as genuinely intentional ?

Evans and Wright laid down their original constraint as an *a priori* condition on a given state's being an intentional one ; Stich took a more humble approach, noting that intuitively we are not willing to call all content-bearing cognitive states genuine beliefs, and he sought only to uncover the rules of classification which we might be tacitly applying in drawing the distinction, of which constraint 1 formed a part.¹⁰ So two questions we might ask in attempting to decide on the plausibility of constraint 2 are : (a) Is there any good *a priori* motivation for the constraint ?, and (b) Does the constraint rule out the states which our intuitions suggest ought to be ruled out ?

I won't spend a great deal of time on (b) : I will limit myself to noting that no genuine belief state can be ruled out by the constraint as it now stands since I can move from a belief that P to almost any other belief Q quite simply, by coming to possess the belief that $P \rightarrow Q$ and drawing out the appropriate inference – where the interaction between the beliefs that P and that $P \rightarrow Q$ needn't take place via any further intentional state casually generated by P ; and to noting that it does seem to rule out at least some of Stich's intuitive examples of subdoxastic states – e.g. in the example of Hess's experiment with the retouched photographs it seems that it is only via the conscious belief that the pupil sizes have been enlarged in one picture that the states representing information about pupil size can indeed interact with other propositional attitudes.¹¹

This latter remark might seem to be in conflict with my earlier remarks concerning Stich's use of the Chomsky example, but it is clear after a little reflection that any conflict is merely apparent. I can deploy the information in r (even though I do not have an explicit belief to the effect that r) in such a way that it can interact directly with other propositional attitudes, because I have an ability partially constituted by my intention to confer a particular status upon r – I can then exercise this ability in such a way as to obtain the necessary interaction despite being unable to give an explicit

statement of the rule. There is no analogue of this in the Hess example – no way in which the notion of intention (and thus of ability) can get a grip – hence, if I don't have the explicit *belief* that the pupils have been enlarged the cognitive states underlying this really will be inferentially insulated from the rest of my propositional attitudes.

Thus I would tentatively suggest that we can give an affirmative answer to (b). But what about (a) ? Why should a state which fails to satisfy constraint 2 be discounted from being a genuine intentional state ?

I think this latter question can only be answered after some reflection on the role played by the postulation of intentional states in the rationalistic explanation of human behaviour. Some sorts of behaviour exhibited by a human agent call for explanation in terms of the beliefs, desires, and other propositional attitudes possessed by that agent ; the use of language is clearly one such kind of behaviour, and I argued in Ch. 1 that the notion needed for explaining this behaviour is that of intention. So we attempt to rationalistically explain a person's use of his language by crediting him with a range of intentions ; but the crucial point is that once we have credited him with the appropriate intentions concerning the use of whole sentences, it seems that our explanatory task is done. Crediting the speaker with *intentions* concerning the *axioms* seems to add nothing whatsoever to the rationalistic explanation of the speaker's behaviour provided by ascribing to him the intentions corresponding to the theorems : the explanatory redundancy of the ascription of tacit knowledge of the axioms is guaranteed by the fact that they only ever play a part in the explanation of behaviour via states of tacit knowledge of theorems – if it *were* possible for states of tacit knowledge of axioms to interact directly with other intentional states then this crucial point about explanatory redundancy could not be made.

Of course, there is the residual task of explaining why he has *those* intentions in the first place, but it is not clear how much scope there is for rationalistic explanation of this as opposed to merely *causal* explanation – we might attempt to give an explanation of the former sort by claiming that he has a *desire* to communicate successfully with others and a *belief* that this aim will best be achieved if he has those particular intentions ; but even if rationalistic explanation is possible here (and I'm not fully convinced that it is) ascribing to the speaker intentions to the effect that the regularities prescribed by the axiomatic base ought to be upheld will again be explanatorily redundant – the question has already been given as full an answer as is possible once the belief and the desire mentioned above have been alluded to.

So it seems that room can be found for the states corresponding to the axioms only within a *causal* explanation of speakers' behaviour, and accordingly we can legitimately view such states only as *causal* states which play a part in the proximate causal history of the (intentional) states corresponding to the theorems in the semantic theory. Ascribing speakers with intentional states corresponding to the various parts of the axiomatic base would simply be to load our theory with more baggage than is warranted by its explanatory brief : if P only ever interacts with other propositional attitudes via a state P* of which it is a causal antecedent, and if describing P* as intentional exhausts the explanatory demands on the theory – in the sense that describing P as an intentional state makes *no* contribution whatsoever to that explanation – then it seems that the most that we can claim concerning P is that it is a *causal* state which has a part to play in the proximate causal history of P*.¹²

2.2 The Generality Constraint

In his paper "Tacit knowledge and the structure of thought and language"¹ Martin Davies suggests two other ways in which the distinction between intentional and non-intentional states can be drawn. The first of these rests on the claim that genuine belief states are generally accessible to the consciousness of their possessor ; but I agree with Davies comment that "it would be a particularly bold theorist who would rest a distinction upon accessibility *versus* inaccessibility to consciousness", and in any case I don't want my argument here to rest on anything as theoretically flimsy as that particular distinction.

The second uses the "generality constraint" discussed by Gareth Evans in *The Varieties of Reference*.² This constraint can be stated in its essentials as follows :

If a cognitive state P with the content *a is F* of an agent W is to be properly called a propositional attitude then the ascription to W of conceptual resources sufficient for the formation of states with contents *b is F*, *c is F*, and so on, for each object *b, c ...* of which W has a conception, and of which *F* can sensibly be predicated, and for the formation of states with contents *a is G*, *a is H*, and so on, for each predicate *G, H ...* of which W has a conception, and under which *a* can sensibly be thought of as falling, must have

some possible explanatory function to fulfil in the rationalistic explanation of W's behaviour.

The rationale for the imposition of this constraint should be clear in the light of the insights provided by section 2.1. In the case of propositional attitudes and intentional states, the direct interaction required by constraint 2 can be achieved only if W has indeed the conceptual resources needed to form the other cognitive states mentioned in the generality constraint – without structure, inferential interaction is impossible. For example, in order to move directly from the belief that "Snow is white" to the belief that "Snow is the same colour as chalk" we have to credit the agent concerned with, amongst other things, conceptual resources sufficient for the formulation of the belief that "Chalk is white".³ In the case of subdoxastic states – e.g. tacit knowledge of a semantic axiom – no such requirement concerning conceptual resources is needed, because the interaction between that state and other cognitive states is generally mediated by another state of which the original was a causal antecedent.

My formulation of the generality constraint differs slightly from the one actually found in Evans and used by Davies. Firstly, I have weakened the constraint somewhat by adding in the condition that W should have conceptual resources sufficient for the formulation of *x is F*, only for those objects *x* of which W has a conception and of which *F* can sensibly be predicated. This is trivial – Evans' constraint as it originally stood would have had us attributing to someone with the belief that "I am in pain" conceptual resources sufficient for the formation of a belief that e.g. "The eiffel tower is in pain". This is clearly too strong, and most probably due to a simple oversight on the part of Evans and Davies. Secondly, on Davies' view subdoxastic states would fail to satisfy the constraint on the grounds that there was some *conceptual bar* to the formation of the other appropriate states. My view is not so much that there is a conceptual bar, but simply that there is no *explanatory warrant* for the ascription of resources sufficient for such a formulation. This seems more reasonable in the light of the conclusion of 2.1.

2.3 Intentions and Concepts

I now come to discuss the "fundamental difficulty" which apparently loomed over our account of implicit knowledge of the rules governing a competent speaker's mastery of his language. The

problem was that in attributing such implicit knowledge to a speaker we seemed thereby to be attributing to him intentional states for the explication of whose contents he manifestly did not possess the requisite sorts of concepts viz. satisfaction, infinite sequence, the star functor, and all the rest of the Tarskian machinery. We thus appeared to be violating what seems to be a very plausible and intuitive constraint on the ascription of propositional attitudes :

Constraint 3 : A propositional attitude P should not be ascribed to an agent if that agent clearly does not possess the concepts needed for a proper explication of the content of P.

Crispin Wright has criticised the view that this constraint *automatically* rules out intentionalistic implicit knowledge of meaning theoretic axioms and recursion clauses on the grounds that it enjoins a revision in some of our everyday practices of rationalistic theorising, and as such requires serious argument in the philosophy of mind.¹ For instance, we normally and unthinkingly offer rationalistic explanations of the behaviour of dumb animals with no pretence that they actually possess the concepts which figure in the intentional states ascribed ; and if we were to come across a group of children who play chess inventively and well, but who are unable to give a verbal statement of the rules of chess, we would still, claims Wright, ascribe to them the intentions , knowledge, and abilities, constitutive of mastery of the game this is not intended by Wright to be an attack on the credibility of the constraint in itself, but merely as a pointer to the fact that *argument* , and not just allusion to common sense and everyday practice will be required in order to discredit the view that speakers can have intentional states corresponding to the axioms and theorems of a semantic theory for their language.

In the previous two sections I provided arguments which I believe effectively rule out any such intentionalistic construal of the speaker-axioms relationship : however, no reason emerged from those sections to cast doubt on the propriety of such a construal of the speaker-theorems relationship – but my worry now is whether the problems posed by constraint 3 also vitiate the idea that a speaker can have intentions concerning the use of whole *sentences* of his language. Granted, we do not now have to credit speakers with sophisticated notions like satisfaction and so on these belong to the axiomatic base and if we are dealing with a homophonic truth theory, the only concepts that will appear in the theorems will be the concepts which figure in the object language sentences (and we can suppose the speakers to be in possession of them), together with the disquotational truth predicate. Can speakers be credited with a grasp of this latter concept ? It

seems to me that we can suppose so, quite unproblematically ; to this extent, a realist who accepts a homophonic semantic theory will be left untouched by constraint 3. But how does an anti-realist fare in this respect ? It seems to me that as far as the anti-realist is concerned, the theorems of his semantic theory are in much the same position, as regards their violation of constraint 3, as the theory's axiomatic base : such theorems will feature notions such as warranted assertibility and superassertibility, of which the average competent speaker will not, on the face of it, be deemed to be in possession. So I think it's clear that constraint 3 not only threatens the intentionalistic construal of the speaker-axioms relationship, but also, at least for an anti-realist who attempts to construct theories in the manner of Wright, the relationship thought to obtain between the speaker and the theorems.

In fact, despite what I said concerning the constraint's apparent lack of effect on homophonic truth-theories, the realist is in just as much trouble as the anti-realist here, for the following reasons. Firstly, if the anti-realist attacks on the intelligibility of the realist's concept of truth are well founded, the realist shouldn't be involving that concept in theories of any sort anyway : the anti-realist would say that speakers do not in fact possess the realist concept of truth because there is no intelligible such concept.² But given our professed agnosticism on the anti-realist / realist debate for the duration of the present essay, the following point is perhaps more immediately damaging for the realist. Although Anscombe, in her famous discussion of intention, concentrates on the notion of intentional action – i.e. on the intention with which an action was done – in contrast to the notion of an intentional state, it's not difficult to see how our problem can be stated in terms of the former. As Anscombe says "...the intentional character of the action cannot be asserted without giving the description under which it is intentional since the same action can be intentional under one description and unintentional under another."³ Thus, a man's action could be described both as "pulling the handle of the pump" and as "poisoning the inhabitants of the house" – but it could only be described as intentional under the former description (assuming for the sake of argument that the man is ignorant of the fact that the operation of the pump handle enables poison to enter the water supply of the house, and is aware that he is in fact pulling the handle of the pump). The construal of the action as intentional under the latter description is ruled out because of Anscombe's insistence that (a) an intentional action is one to which a certain sense of the question "Why ?" is given application⁴, and (b) that this question is refused application by the

answer "I was not aware I was doing that" . This is precisely the sort of answer likely to be given by our pump attendant above, when asked "Why are you poisoning the inhabitants of the house ?", and this shows, in Anscombe's view, that his action construed in the latter of the two senses above cannot properly be described as intentional.

All of this seems plausible enough, but how does it apply to the language case ? We want to view someone competent with a given sentence of his language as possessing an intention to uphold the regularity prescribed by a meaning-delivering theorem of a correct semantic theory of his language. Suppose then, that such a speaker utters the sentence "The sun will rise tomorrow". Then, this action can be described as "an utterance of 'The sun will rise tomorrow' ", or as "an act of conforming to (#)", where (#) states the truth-condition of the sentence "The sun will rise tomorrow". Now, if we ask him "Why did you conform to the regularity prescribed by (#) ?" we are almost certainly going to be met with the answer "I wasn't aware that I was doing *that* ", which rules out calling the action intentional under the more complex of the two descriptions. So even if we grant, contra the anti-realist, that a speaker might have a grasp of a verification-transcendent concept of truth, Anscombe's account seems to show that the realist is affected just as much : it's not just possession of concepts that counts – although lacking the concepts is a sufficient condition for failing Anscombe's test, possessing them is not by itself sufficient for passing it. In any case the answer "I was not aware I was doing that" is guaranteed, if the description under which the action is given contains concepts outwith those possessed by the agent in question. Thus, given that Anscombe's account rules out in a general way the description as intentional of the agents utterance as "Upholding the regularity prescribed by ...", it seems to cast doubt on the ascription of intentional states to agents whose content is precisely that. In fact, we can frame a constraint similar in spirit to constraint 3 using the notion of intentional action instead of intentional state :

Constraint 4 : an action A of an agent W can be described as intentional under a given description X only if the concepts used in the description X do not actually outstrip those possessed by W,

which, when taken together with the following

Constraint 5 : an agent can be ascribed the intentional state "intending to X" – i.e. the intention to X – only if (at least some of) his actions count as intentional when taken under the description X,

has the same effect as constraint 3 : we cannot properly describe speakers as possessing intentional

states corresponding to the theoretic output of a semantic theory, if the concepts used therein are not actually possessed by those speakers.

Whither now ? Must we give up the whole enterprise of semantic theorising as we have conceived it ? Wright takes his allusion to everyday rationalistic theorising as providing a basis for the rejection of the third constraint, and remains agnostic, in the absence of any compelling argument to the contrary, on the ascription of intentional states which violate it. I will now argue that the constraint is in fact plausible, but that its plausibility need do nothing to vitiate our conception of the theorist of meaning's enterprise – in effect, where Wright denies the speakers possess the relevant concepts and saves his skin by refusing to embrace constraint 3, I will accept the constraint, but accept also that the speakers *do* possess the requisite concepts, even though they are unable to give their possession of them verbal expression. First, however, I want to make some comments regarding Wright's agnosticism on the legitimacy of constraint 3 : I want to argue that whatever independent merits or demerits attach to this latter course of action, it is not unproblematically available to Wright because it *prima facie* conflicts with his own views on how the truth values of ascriptions of intentional states come to be determined.

Wright suggests that "intentional and sensational states...are, in effect, 'secondary' : that subjects' best judgements fix the extension of the truth predicate among ascriptions of belief, desire, and feeling to them."⁵ The concept of intention can be said to be secondary in Wright's sense if and only if a subject's *best opinions* serve to *determine*, rather than *reflect*, the extension of the truth-predicate with respect to the sentences through which his avowals of intentions are expressed – where best opinions are those formed under what are, with respect to the particular subject matter under consideration, *cognitively ideal* conditions of both the subject and the circumstances in which he acquires the appropriate opinion. Best opinions are thus viewed by Wright as determining constitutively the truth-values of sentences expressing avowals of intention, as opposed to *tracking* independently constituted states of affairs which confer truth or falsity upon them.

Wright argues for this claim by reflecting on what he calls the *provisional equation* :

S judges under conditions C --> (P <--> S believes that P)

where the C-conditions are in fact the cognitively ideal circumstances we spoke of above. According to Wright it is sufficient for our best judgements concerning sentences of the same class as P to be *extension determining* that each of the following four conditions are satisfied:

(1) The equation should be *a priori* true – we should have a priori covariance of best opinions and truth.

(2) The C-conditions should admit of non-trivial specification – they should not be conditions that have 'everything it takes' to ensure the correctness of the subject's opinion.

(3) The question as to whether the C-conditions are satisfied in a particular case must be answerable independently of any truths concerning the details of the extensions of the class of concepts under scrutiny.

(4) There must be no better way of accounting for (1), (2), and (3) above.

Wright attempts to substantiate his claim by showing that (1) - (4) hold in the case of self ascription of intention, and his account of how the special authority possessed by avowals is possible subsequently turns on a story concerning the relative ease with which best opinions can in this case be accomplished. His main worry about his account concerns the possibility of the satisfaction of the non-triviality condition (2), and he provides what seems to me to be a reasonable way of assuaging this worry⁶ – but the worry I mentioned above arises from the fact that Wright includes the following among his C-conditions: *the agent must have an appropriate grasp of the concepts requisite for the expression of the avowal in the first place*. I think that this points to some tension between the agnosticism which Wright espoused earlier concerning e.g. the description of the dog as having genuine intentional states, and the extension-determining thesis just outlined. The inclusion of the above among the C-conditions means, arguably, that in the case of (really) dumb animals the c-conditions which must obtain before there can be any determinate facts concerning their intentions, can never actually obtain, and that therefore we cannot legitimately speak of them as possessing intentions (at least, not of any significant degree of complexity). And more importantly, if this really is one of the C-conditions, the idea that competent speakers have tacit knowledge of meaning delivering theorems is threatened just as much as their putative tacit knowledge of the axiomatic base. So I suggest that Wright's way of setting up the C-conditions in the intention case effectively prejudices the issue concerning implicit

knowledge *qua* intentional state, upon which he earlier expressed a desire to remain neutral.⁷

It might seem that I have been unfair to Wright here insofar as his thesis is concerned with *avowals* – 1st person ascriptions of intentions – and how they come to have their truth values determined. Couldn't it be the case, it might be objected, that we ought to concentrate our attention on 3rd person ascriptions : on ascriptions of the form "X intends to Ø" as uttered by Y, as opposed to "I intend to Ø" as uttered by X himself ? If we could give an alternative account of how the truth values of the former sort of ascription came to be determined, not involving the ascription to X of the concepts requisite for the expression of "I intend to Ø" then we might indeed be able to bypass the problem raised in the preceding paragraph.

But this suggestion strikes me as unattractive : shouldn't the state of affairs which confers truth or falsity on "I intend to Ø" as uttered by X be the *same* as that which confers truth or falsity on "X intends to Ø" as uttered by Y – in much the same way that the state of affairs which confers truth or falsity upon "I am wearing a blue jacket" as uttered by X is the same as that which confers truth or falsity on "X is wearing a blue jacket" as uttered by Y ? This problem is accentuated if we imagine the 1st and 3rd person ascriptions taking place at the same time and under identical conditions. Suppose John says "I intend to go to the cinema" simultaneously with my utterance of "John intends to go to the cinema tonight". Suppose for the sake of argument, that the relevant C-conditions are satisfied, and that John's avowal is true. What makes this avowal true, on Wright's account is not that it tracks or fits some independently constituted state of affairs, conceived by the Cartesian, for instance, as an event in, or state of, John's inner mental theatre ; rather, what makes it true is simply that John expressed the avowal in circumstances in which all of the C-conditions were in fact satisfied. But is it *this* which makes the 3rd person ascription true ? If so, then the reply we made on behalf of Wright appears to be blocked : 1st and 3rd person descriptions are determined in the same way, so lack of concepts on the part of the agent to whom the intentional states are ascribed vitiates the latter sort of ascription as well as the former. If not, then we are left with the following problem : if the truth values of the two sorts of ascription are *not* determined in the same way, what guarantee do we have that they will not come into conflict i.e. on what can we base our assurance that my utterance and John's avowal must have the same truth value ?

This latter remark suggests the proper response we require from Wright here : what we need is an account which tie together the truth conditions of avowals and 3rd person ascriptions of

intentional states when the C conditions are satisfied. This ensures that there will be no conflict of the sort mentioned above – but in order to retain the possibility of tacit knowledge, we need in addition, an account of how the truth conditions of avowals and 3rd person ascriptions come to be determined *when the C-conditions are not in fact satisfied*. What was wrong with our suggestion above was that it attempted to give an account of how truths about intentional states come to be determined which *cut across* cases in which the C-conditions obtained and cases in which they did not. But the onus is now on Wright to produce such an account – until this is forthcoming his agnosticism on the issue seems at best to be only dubiously consistent with his story concerning the determination of truths about intentional states. If he wishes to retain this agnosticism then he will either have to (a) alter his account of how 1st person ascriptions of intentional states come to be determined – in particular, if he wishes to retain the extension-determining thesis in the present instance, he is going to have to modify his characterisation of the C-conditions in such a way that possession of concepts is not an essential prerequisite for the possession of the intentional states in the explication of whose contents those concepts figure, or (b) provide an account of how the truths in question come to be determined when the C-conditions do not obtain. Either way, more work requires to be done.⁸

In any case, let us leave these ad hominem points against Wright on one side, and examine the plausibility of the third constraint, and its consequences for the notion of implicit knowledge.

Wright presents two examples in an effort to show that the imposition of constraint 3 cannot be justified by means of simple recourse to common sense and to our everyday thinking and modes of explanation. This is because of the following fact : if the constraint were justified then certain explanations which we would normally and unthinkingly give in our pre-philosophical jackets would be ruled out of court – so if we are looking for a justification for constraint 3, we are going to have to look elsewhere than to our everyday explanatory practices.

The first example concerns the rationalistic explanation (explanation via the ascription of propositional attitudes and intentional states) of the behaviour of dumb animals. Wright suggests that if a dog sets off each day around the same time and subsequently meets his master on the road, we would ordinarily explain his acts of setting off at that time in terms of an intention to intercept his master – in spite of the fact that the dog does not possess the concept "master".⁹

I'm not sure what to say about this example – as Wright says, much work remains to be done in

the philosophy of mind before as we can be sure as to the propriety or otherwise of the sort of explanation offered for the dog's behaviour. All I want to say here is that I think Wright has misplayed his hand somewhat, and that we can in fact do better than achieve the agnostic position he finally settles for in the text. This will involve a discussion of the example already alluded to at the beginning of this section, that of the dumb chess players : these children, you will recall, display all the trappings of intelligent chess playing behaviour – behaviour, that is, distinctive of those who have a mastery of the rules of the game – with the difference that that they are unable to state any of the rules of chess verbally, and don't understand any such statements when they are presented to them. Now the rules of chess will contain mention of many concepts – "bishop", "checkmate", and "en passant", to name but a few – none of which can, by hypothesis, be explained or given verbal formulation by the children in our example. But, as Wright quite correctly says, it seems entirely natural to credit such children with "the knowledge and intentions constitutive of an understanding of, and the practice of playing by, the rules of chess".¹⁰ Wright takes this as showing that we ascribe intentional states – in this case implicit knowledge of the rules of chess – to people, when they don't possess the concepts necessary for a proper statement of the contents of those states – in this case "King", "checkmate" and so on. And this in turn is supposed to show that we naturally ascribe intentional states in such a way that constraint 3 is apt to be violated.

This is, I think, a definite error on Wright's part : for what is it to possess the concepts in question other than to possess the abilities constitutive of mastery of the game ? I want to say that the children in the example *do* possess the concepts of "checkmate" and so on, even though they are unable to give vent to this possession in language. And this should hardly be controversial – all it amounts to is the claim that *knowledge how* is more important than *knowledge that* as regards the concepts a person can legitimately be viewed to be in command of : to possess a concept *is* to possess an ability, and it seems to me at best gratuitous to demand that explanations or verbal formulations should be included among the exercises of this ability.¹¹ To this extent I am in agreement with Geach, when he says that

"If a man struck with aphasia can still play bridge or chess, I certainly wish to say he still has the concepts involved in the game, although he can no longer exercise them verbally. But it would be hard to devise non-verbal criteria for the patients having retained a concept of the day after tomorrow."¹²

This latter sentence might seem to cast doubt on the idea that there might be non-verbal criteria for a

speakers having retained a concept of e.g. assertibility, but I think that Geach is, strictly speaking, incorrect, as far as the latter claim goes. To see this all we need to do is to imagine the rules of our current game of chess enriched to include the following (we might call the resulting game "temporal" chess) : if a player at any time waves a yellow handkerchief at his opponent during the time allotted for one of the opponents moves, then the opponent shall defer making his move until the day after tomorrow. My suggestion then is that if the players play in accordance with this rule and use it intelligently etc (perhaps as a means of securing more time for the analysis of a tricky position), then we can indeed credit them, contrary to Geach's suggestion, with the concept of "the day after tomorrow".

Our conclusion in the case of language mastery should now be clear: if speakers linguistic usage is governed by rules – if they can speak the language in accordance with those rules – then this in itself is enough to establish that the speakers *do* possess the concepts involved in the statements of the rules, even where those rules are immensely complex as regards their conceptual content, and even where those speakers are unable to give explicit statements or linguistic paraphrases of the rules. To recapitulate, this should come as no surprise : possessing a concept is possessing an ability, is possessing *knowledge how*, which knowledge is prior to any species of *knowledge that*..

We are now left with the following residual problem, stemming from our brief discussion of Anscombe earlier in this section : even if we grant that speakers possess the concepts which feature in the statements of the truth-conditions of the sentences of their language, the fact is that they will answer "I was not aware I was doing that" when asked why they conform their linguistic practice to the rules which give those truth-conditions. This means, taking Anscombe's stance, that the action which consists in uttering a sentence cannot be described as intentional under the description "conforming to...." and thus, by constraint 5, they cannot be ascribed intentional states with contents along the lines of "intends to conform to...". Note again that this point is not specifically about the concepts they possess : the speakers arguably do possess the crucial concepts, but possession of concepts does not guarantee an affirmative answer to the awareness question.

The solution to this difficulty is, I think, that Anscombe has overstated her case : let's consider our chess playing children again. I have said that these children would plausibly be viewed as knowing the rules of chess – intending to uphold those rules – despite their hypothetical verbal incapacity. Suppose that one of the children moves a pawn to capture *en passant* one of the

opposing pawns, in such a way that the rules governing *en passant* are respected. Suppose also that we ask the child "Why did you move your pawn in such a way that the rules governing *en passant* were upheld?". Ex hypothesi, the answer to this question will be an avowal of lack of awareness on the part of the child – if Anscombe were correct, this would mean that the child's action could not be viewed as intentional when taken under the description "conforming to R", where R gives the conditions under which an *en passant* move can legitimately be made, and hence by constraint 5, we could not ascribe to the child an intentional state with the content that the rule in question should be conformed to. This in itself should be enough to cast some doubt on Anscombe's criterion for intentional action – why should awareness as such be the determining factor? And if we read "consciousness" for "awareness" things begin to look suspiciously like the pictures of understanding which Wittgenstein takes such pains to discredit. I would say that the realm of an agent's intentional actions far outstrips the realm of actions that he is, as it were, consciously aware of executing. Anscombe's example of the pump handler lends her account some plausibility – the agent's action would indeed only be improperly viewed as intentional when taken under the description "poisoning the inhabitants of the house", and we would thus be in error if we ascribed to the handler an intention to poison the inhabitants. But the reason for this is not that the handler lacks an explicit awareness of the poison entering the water supply of the house, or is ignorant of the causal relations between the pump and the flow of the poison, but rather that there is no state which stands in the appropriate sort of relation either to his other mental states, or to his future patterns of behaviour. This is where the pump case differs from the chess case, although each of them exhibit a negative response to Anscombe's test: in one case, there is a state which is suitably related to – has the appropriate causal / functional role in the production of – the agent's future behaviour, whereas in the other there is no such state.

My conclusion then, is that Anscombe's insistence on explicit awareness in action lacks any sound basis once we have taken a proper account of the Wittgensteinian insights concerning the nature of the relationship between understanding and use, between intentional states and behaviour generally, and the related fact that explicit awareness (consciousness) can have no distinctive role to play in establishing that relationship. There is no deep problem with the idea that speakers can have intentions to uphold rules, where the statements of those rules contain concepts which the speakers are unable to formulate, paraphrase, or even recognise on presentation – or, in other words, there is no deep problem with the idea that such rules can be implicitly known.¹³

3. COMPOSITIONALITY

Calculating prodigies who get the right answer but cannot say how. Are we to say that they do not calculate ? (A family of cases.)

(Wittgenstein, *Philosophical Investigations*, §236.)

sentences.

Evans' discussion proceeds with reference to the relatively simple, and finite, language (call it L) consisting of ten names 'a', 'b', 'c',..., which stand for Harry, John, Bill,...., ten predicate expressions 'F', 'G', 'H',..., which stand for happiness, baldness, heaviness,... . This language then has 100 syntactically admissible sentences, each consisting of the concatenation of a name with a predicate.²

Now suppose that a semantic theorist sets out to find the correct theory of meaning for this language ; one constraint on this enterprise is of course that the theory eventually settled for should have the right *output* ; namely, that the meaning specifications which it issues in should be *correct*.. For the sake of clarity let us suppose that what the semantic theorist is after is a correct Davidsonian truth conditions theory for the language L (nothing of any importance hinges on this decision for our present purposes). Then, we will regard such a theory as acceptable for L if it delivers the following set of truth condition specifications :

'Fa' is true in L iff Harry is bald
'Fb' is true in L iff John is bald.....
.....'Ga' is true in L iff Harry is happy.....
.....'Oj' is true iff Engelbert is less than optimally hirsute.

But then the following problem arises for our semantic theorist. Call two theories which issue in the same set of truth conditions specifications "extensionally equivalent". Then the following two theories for L will be extensionally equivalent :

T₁ : the *listiform* theory, which has 100 axioms, one for each individual sentence of L (simply the list of truth conditions specifications given immediately above).

T₂ : the *articulated* theory consisting of 21 axioms, one for each of the proper names (e.g. " 'a' denotes Harry"), one for each of the predicates (e.g. "an object satisfies 'F' iff it is bald"), and an axiom for the subject predicate mode of combination ("a sentence coupling a name with a predicate is true iff the object denoted by the name satisfies the predicate).

Obviously the theory T₂ is closer in spirit than T₁ to the theories which semanticists have in fact been attempting to construct for natural languages, but Wright's challenge can now be stated as

follows : given that the constraint that a theory issue in the correct truth condition specifications for the language is not by itself sufficiently strong to discriminate in favour of T_2 , can any further constraints be imposed which will provide a substantial motive for the preference of T_2 to T_1 ? In other words, can there ever be empirically respectable evidence which will discriminate between two extensionally equivalent semantic theories ?³

Evans suggests that such a constraint can be given, and that there can be perfectly respectable empirical evidence for or against the assertion that a given theory meets that constraint. The constraint, very roughly, is this : the theory should aspire, not only to provide a set of correct truth condition specifications for the language under scrutiny, but also to describe the dispositions corresponding to each of the expressions for which that theory has a proper axiom. Thus if we find that speakers have 100 dispositions of the relevant type we will be justified in accepting T_1 (in which case we will say, somewhat metaphorically, that speakers of L have tacit knowledge of T_1), whereas if we find that they have 20 such dispositions, the acceptance of T_2 will be warranted (and with it the claim that the speakers of L have tacit knowledge of T_2).

But precisely what are the dispositions of "the relevant type" alluded to above ? In the case of T_1 the dispositions corresponding to its primitive expressions (the expressions to which it devotes an individual axiom) are relatively easy to specify : each disposition is simply "a disposition to judge utterances of the relevant sentence type as having such and such truth conditions."⁴ T_2 is somewhat more problematic in this respect, because its primitive expressions are not whole sentences, but rather proper names and predicate expressions, and of course it is only *sentences* which can be said to have truth conditions. As a consequence, the dispositions corresponding to the primitive expressions of T_2 have to be inter-defined. Evans suggests characterising the dispositions, possession of which would constitute tacit knowledge of T_2 , in the following manner :

"... we might say that a speaker U tacitly knows that the denotation of 'a' is Harry iff he has a disposition such that :

$(\pi\phi)(\pi\psi)$ if :

(i) U tacitly knows that an object satisfies ϕ iff it is ψ

(ii) U hears an utterance having the form ϕ -a,
then U will judge the utterance is true iff Harry is ψ .

Connectedly, we say that a speaker tacitly knows that an object satisfies 'F' iff it is bald iff he has a disposition such that :

$(\pi x)(\pi \alpha)$ if

(i) U tacitly knows that the denotation of α is x ,

(ii) U hears an utterance having the form F - α ,
then U will judge that the utterance is true iff x is bald.

In these formulations, ' π ' is a universal substitutional quantifier, with variables having the following substitution classes: ϕ , names of predicate expressions of the (object) language; α , names of names of the (object) language; ψ , predicate expressions of our language (the metalanguage); and ' x ', proper names of our language."⁵

How can we tell, empirically, whether or not a speaker has the dispositions possession of which constitutes, on the present account, tacit knowledge of T_1 or of T_2 ? Evans suggests three possible sources of relevant evidence.

The first source is connected with Evans' insistence that the notion of disposition involved in his account has to be taken in a full-blooded way – the ascription of a disposition is not to be regarded merely as a statement that some regularity obtains :

"These statements of tacit knowledge must not be regarded as simple statements of regularity, for if they were, anyone who correctly judged the meanings of complete sentences would have a tacit knowledge of T_2 . When we ascribe to something the disposition to V in circumstances C, we are claiming that there is a state S which, when taken together with C, provides a causal explanation of all the episodes of the subject's V-ing (in C). So we make the claim that there is a common explanation to all those episodes of V-ing. Understood in this way, the ascription of tacit knowledge of T_2 involves the claim that there is a single state of the subject which figures in a causal explanation of why he reacts in this regular way to all the sentences containing the expression."⁶

then

"The decisive way to decide which [ascription of tacit knowledge] is correct is by providing a causal, presumably neurophysiologically based, explanation of comprehension. With such an explanation in hand, we can simply see

whether or not there is an appeal to a common state or structure in the explanation of the subject's comprehension of each of the sentences containing the proper name a."7

In addition, we can also examine the patterns of acquisition of knowledge of the meanings of sentences manifested in the linguistic behaviour of L-speakers. For example, evidence suggestive of tacit knowledge of T_1 would be that even when a speaker has acquired dispositions to judge correctly of the truth conditions of Ga and Fc, he is not thereby (in the absence of further training and exposure) disposed to judge correctly of the truth conditions of Gc. Evidence suggestive of tacit knowledge of T_2 would be that he *is*, under the same conditions, so disposed.

Further evidence is provided by the patterns of *loss* of knowledge of meanings exhibited in speakers of L. If such a speaker is initially competent with each of the 100 sentences of L, and if, by knocking out his competence with, say, Hd, we thereby disturb his competences with all other sentences containing the expressions 'H' and 'd', then tacit knowledge of T_2 will be ascribable. If, on the other hand, the other competences remain undisturbed by the speaker's loss of competence with Hd, then ascription of tacit knowledge of T_1 will be in order.

It seems then, on the face of it at least, that we have found an empirically respectable way of deciding which of the two theories T_1 and T_2 should be accepted for the language L ; if the speakers of L have tacit knowledge (in the sense outlined above) of T_1 , then we ought to accept T_1 as providing the best theory of truth conditions for that language ; if, on the other hand, speakers have tacit knowledge of T_2 , we should accept T_2 as the more suitable of the two theories.

It is worth noting that at this stage the motivation underlying Evans' account seems to be the following : the derivational structure of a theory of meaning (the canonical routes from its axioms to its theorems) should in some sense reflect the causal structure found among the competences of the speakers of the language under scrutiny (the causal routes leading from the dispositions associated with the language's names and predicates to the intentional states associated with the whole sentences of the language). In the rest of this chapter we will have occasion to elaborate upon and question the constraint suggested by this underlying motivation.

3.2 Wright's criticisms of Evans and Davies' defence thereof

In this section I outline two of the criticisms that Crispin Wright has raised against the dispositionalist account of tacit knowledge of semantic axioms which I adumbrated in the previous section, and the responses that have been offered by Martin Davies on Evans' behalf. I will argue that the responses offered by Davies to the two criticisms in question are unsuccessful as they stand, and will sketch my own alternative defence of Evans.¹

The first of Wright's objections which we shall consider is concerned with Evans' characterisation of the dispositions possession of which constitutes, on his account, tacit knowledge of the theory T_2 .

We can see from the quotations on pages 40-41 above that the dispositions corresponding to the predicate expressions and names of L are interdefined; that is to say, the dispositions supposed to constitute tacit knowledge of the denotation conditions of the proper names of the language are defined in terms of tacit knowledge of the satisfaction conditions of L 's predicate expressions; similarly, the dispositions supposed to constitute tacit knowledge of the satisfaction conditions of the predicates are defined in terms of tacit knowledge of the denotation conditions of L 's proper names.

Why is this a problem? As Wright puts it "...to characterise a disposition ought to be to characterise both what it is a disposition to do and the circumstances under which it will be manifest."² Now suppose, by way of illustration, that we are attempting to give a dispositional account of the metallurgical phenomenon of ductility, and that we come up with the following result: X is ductile iff the observable phenomena c_1, \dots, c_n occur under background circumstances C . Suppose further that these background conditions include the possession by X of the additional dispositions d_1, \dots, d_k . Now Wright's point is that if it turns out that in characterising the manifestations distinctive of some one of the further dispositions d_i , say, we *have* to refer to background circumstances *which include the assumption that X is ductile*, we will thereby have said *nothing* whatsoever as to what ductility consists in – we will simply have failed to say what ductility is.

This point seems to me to be fundamentally correct, and it isn't difficult to see how the objection

applies to Evans' characterisation of the dispositions which make up tacit knowledge of T_2 . Take the example of the disposition which is said to constitute tacit knowledge of the denotation of the name 'a'. What we are after in characterising this disposition is something of the form : X tacitly knows that the denotation of 'a' is Harry iff observable phenomena c_1, \dots, c_n occur under background conditions C. In this case the background conditions include the possession by X of the further dispositions d_i, \dots = tacit knowledge of the satisfaction conditions of certain of the predicate expressions of L. *But*, a characterisation of the distinctive manifestations of the d_i 's is possible only if we are prepared to make reference to background conditions in which X is assumed to have tacit knowledge of the denotation conditions of the names of L – and, it is precisely *this* species of tacit knowledge which we are at present concerned to explicate – so, our account turns out to be viciously circular, and we apparently fail altogether in our attempt to say what tacit knowledge that the denotation of 'a' is Harry consists in.

Wright himself is not pessimistic about the possibility of a solution to this problem :

"...I offer the point more as something which someone who wished to advance Evans's account should say something about than as an objection. Perhaps a more sophisticated account of the notion of a disposition would remove the worry ; my own suggestion would be that Evans's proposal should have proceeded by reference to states of a different sort – his real interest after all, is in the underlying 'categorical' bases."³

and Davies subsequently offers what seems to be a respectable way around the trouble threatened by Evans's characterisation. Davies suggests that we cast our account of tacit knowledge in terms of "underlying explanatory states", rather than in terms of dispositions – instead of defining tacit knowledge in terms of the dispositions a speaker has concerning truth, satisfaction, and denotation conditions, we define it in terms of the states which make up the "categorical bases" underlying those dispositions. We ought to note, in fairness to Evans, that this seems to be precisely what he was after in the first place – witness his claim that dispositions have to be given a full blooded characterisation (see above), and his statement that

"Tacit knowledge of T_2 requires that there should be twenty such states of the subject – one corresponding to each expression of the language which the theory treats separately – such that the causal explanation of why the subject reacts in the way that he does to any sentence of the language

involves two of those states, and any one of these states is involved in the explanation of the way he reacts to ten sentences containing a common element."⁴

In any case, the speaker with tacit knowledge of T_2 will not now be characterised as having 20 dispositions defined in the ways suggested by Evans, but as being the bearer of twenty causal explanatory states each of which is the basis of one of those dispositions. This allows us to sharpen up the constraint that was breaking through the clouds at the end of the preceding section - that constraint could perhaps now be stated as follows :

If, and only if, a speaker who has dispositions to judge correctly of the truth conditions of S_1, \dots, S_n is thereby (and without any further training or exposure) disposed to judge correctly of the truth conditions of S , should the semantic resources sufficient for the canonical derivation of truth conditions specifications for S_1, \dots, S_n be sufficient for the canonical derivation of a truth condition specification for S .

Under our first revision of Evans's account in terms of underlying states this becomes what Davies has termed the "mirror constraint" :

If, and only if, the operative states implicated in the causal explanation of a speaker's beliefs about the meanings of S_1, \dots, S_n are jointly sufficient for a causal explanation of his belief about the meaning of S , should the semantic resources sufficient for the canonical derivation of truth conditions specifications for S_1, \dots, S_n be sufficient for the canonical derivation of a truth condition specification for S .⁵

However, I have the following residual worry about whether this really does satisfactorily avoid the problem concerning interdefinability and vicious circularity which Wright has raised here : if our only means of *individuating* the categorical bases underlying the dispositions is *via* the notion of the dispositions that they underlie, then doesn't the problem simply carry over into the revised account in terms of causally operative states ? Is it possible to characterise the causally operative state which underlies my disposition connected with the denotation conditions of the name 'a' *without* referring to the causal states underlying the disposition I have connected with the satisfaction conditions of L 's predicates ? If not, and if this holds vice versa, then I suggest that we have again failed to say what tacit knowledge that the denotation of 'a' is Harry consists in, because we will have failed to individuate the causal state possession of which allegedly constitutes that tacit knowledge. So we are faced with the following dilemma : either we must provide an

account of how the operative states can be individuated without reference to the dispositions which it is claimed that they underlie – which account is at present lacking – or, on the other hand, the same problem that arose for Evans' original account simply arises again for Davies proposed revision.

Is Wright's objection then fatal to Evans' account? It seems to me that we would be overhasty in concluding that it is: Davies' switch to talk of underlying causal states and categorical bases indeed does nothing to remove the circularity which Wright focuses on, but how "vicious" is that circularity? I would say that if there is in fact a vicious circularity here then it is bound to infect *any* constitutive account of the mastery of individual subsentential expressions, no matter whether that account is couched in terms of dispositions, underlying causal states, or in terms of anything else we might care to choose. This should raise our suspicion, and is quite easily seen.

Any account of what competence with a name consists in will have to contain resources sufficient for an account of what competence with the sentences containing that name consists in (this is because competence with the name simply *is* competence with those sentences). However, this account is incomplete in the absence of an account of what the understanding of predicates consists in; and when we try to give this latter account – an account of what understanding the sentences containing the predicates consists in, we find ourselves back at the point at which we set out i.e. requiring an account of what competence with names comes to. So the circularity really comes to this: *any* account of what competence with a name consists in requires an account of what competence with predicate expressions consists in, and vice versa.

This points to the proper line of response to Wright's objection. Isn't it an accepted fact that a similar interdefinability property is possessed by beliefs and desires? In the simple case, intentional action requires both beliefs and desires to be present, and more generally (and as we noted earlier) beliefs, desires, and propositional attitudes are ascribable to an agent only in systems, and not individually. But when we attempt to give e.g. a constitutive account of the belief which partially rationalises a certain action, we stand in need of a similar account of the appropriate desire, and vice-versa. However, we do not normally take this as signalling the impossibility of providing a constitutive account of either belief or desire: rather, the conclusion drawn is that the relationship between the beliefs, desires, and the physical facts which ground their ascription is irreducibly *holistic*. How then do we say what beliefs and desires are? I think, roughly speaking, that there are two components in an answer to this.⁶ (a) By showing how the propositional attitude

states relate *to each other* by giving the a priori principles which constrain the relations *between* the various sorts of propositional attitude (an example of such a principle might be what David Lewis has called the "rationalisation principle" – the beliefs and desires ascribed to the natives should be, broadly speaking, consistent with the redescription of their physical behaviour as intentional, and with the contents of the speech acts they perform : in a nutshell, what they *think* , should be consistent with what they *say* and with what they *do*).⁷ And (b) by giving the interpretative principles which link the propositional attitudes holistically to the physical facts which ground their ascription (an example of a principle of this type might be the "principle of charity" – the natives should be interpreted as believing and desiring what, broadly speaking, *we* feel they ought to be believing and desiring *vis a vis* the physical circumstances that happen to obtain). In summary then, we individuate a propositional attitude, not by picking it out individually, but by giving its place in the network of propositional attitudes that form part of any given agent's mental armoury – so the fact that we *can't* pick out such states in isolation from others with which they are interdefined needn't give us too much cause for concern.

Perhaps a similar manoeuvre is available to us in the case of states of tacit knowledge. All the interdefinability shows is that tacit knowledge too has to be ascribed in a holistic fashion : we do not give an account of what tacit knowledge of e.g. a name "a" consists in apart from an account of what constitutes tacit knowledge of the complete axiomatic base – we give such an account by delineating the constraints which govern the ascription of tacit knowledge of the axiomatic base *as a whole* , and by examining the conceptual relationships that obtain between states of tacit knowledge of its various parts. And this is where the mirror constraint has a part to play : faced with a semantic theory T, we decide whether or not a speaker should be ascribed tacit knowledge of T by seeing whether the theory meets the mirror constraint with respect to that speaker, in other words, by seeing whether the derivational structure *of the whole theory* is isomorphic in the relevant sense to the causal structure found in that speakers overall linguistic competence. Just as we give an account of what beliefs are by giving their location within a wider network of propositional attitudes, we give an account of what tacit knowledge is by showing how the causal states in question are located in a wider causal structure : having given the structure, we need say nothing more about the composition of the individual states. The mental, unlike the metallurgical, is essentially holistic.

So my conclusion is thus that although Davies' move from Evans' dispositions to the mirror

constraint does after all provide the resources for a rebuttal of Wright's objection concerning interdefinability, that rebuttal is perhaps more complicated than Davies' paper would initially have us believe.

We now move on to look at Wright's second objection to the account of tacit knowledge suggested by Evans's proposal. It will be my contention that although the solution which Davies proposes *is* a good solution to a problem which Evans perhaps ought to have taken account of, it simply *fails* to engage whatsoever with Wright's objection in the form in which he originally raised it.

Davies summarises the objection thus :

"The second objection relates to the account of tacit knowledge as a certain kind of causal structure. Suppose that a subject knows (tacitly or in the ordinary sense) what the various sentences of L mean ; and suppose that underlying those pieces of knowledge there is indeed a causal structure of the kind which, on Evans's account, is required for tacit knowledge of T_2 – the articulated theory. Wright asks why such a subject would not be at least as well described by a two-part theory. The first part would be the *semantic* theory T_1 – the listiform theory ; the second part would be 'some appropriate hypotheses of a *non-semantic* sort, about the presumed causal substructure of the dispositions which T_1 describes'. Why, in short, does mere *causal* structure justify articulation in a *semantic* theory ?"⁸

The suggested problem, then, seems to be that not all attributions of causal explanatory structure will be pertinent to the ascription of tacit knowledge to the speaker concerned, and in order to bring this point home Davies provides an example of a case in which "a pattern of breakdown is intuitively misleading as to the attributability of tacit knowledge ...[but where] ... the evidence is not obviously misleading as to the presence of some kind of causal structure".⁹ Consider the following diagram (see over) :

F G H I J K L M N O

X _a	a	Fa	Ga	Ha	Ia	Ja	Ka	La	Ma	Na	Oa
X _b	b	Fb	Gb	Hb	Ib	Jb	Kb	Lb	Mb	Nb	Ob
X _c	c	Fc	Gc	Hc	Ic	Jc	Kc	Lc	Mc	Nc	Oc
X _d	d	Fd	Gd	Hd	Id	Jd	Kd	Ld	Md	Nd	Od.
X _e	e	Fe	Ge	He	Ie	Je	Ke	Le	Me	Ne	Oe.
X _f	f	Ff	Gf	Hf	If	Jf	Kf	Lf	Mf	Nf	Of.
X _g	g	Fg	Gg	Hg	Ig	Jg	Kg	Lg	Mg	Ng	Og.
X _h	h	Fh	Gh	Hh	Ih	Jh	Kh	Lh	Mh	Nh	Oh
X _i	i	Fi	Gi	Hi	Ii	Ji	Ki	Li	Mi	Ni	Oi
X _j	j	Fj	Gj	Hj	Ij	Jj	Kj	Lj	Mj	Nj	Oj
		Y _F	Y _G	Y _H	Y _I	Y _J	Y _K	Y _L	Y _M	Y _N	Y _O

and consider the speaker C who has a language system which performs derivations in an explicit representation of the semantic theory T_1 – the above matrix is supposed to show how the representations of the 100 axioms of T_1 are arranged, and each of its members represents an information storage unit. Now we elaborate the example somewhat and suppose that in order to function properly the individual units of the matrix have to be supplied with certain nutrients, which flow through the matrix in channels. Suppose also that there are two types of nutrient X and Y, that the X nutrient flows through the matrix in ten channels X_a, \dots, X_j "each of which serves the ten storage units for the ten sentences containing a single name", and that in similar fashion the Y nutrient flows in ten channels Y_F, \dots, Y_O "each of which serves the ten storage units for the ten sentences containing a single predicate". The supposition crucial for example is that each of the storage units will *fail* to function if it doesn't get its supply of *each* of the nutrient types, so that "failure of a unit prevents nutrient flow through at least one of the channels serving that unit".¹⁰

The difficulty should now be clear : we want to say that if the speaker has tacit knowledge of any theory then he must have tacit knowledge of T_1 , but the patterns of breakdown likely to occur in his linguistic behaviour will suggest the ascription of tacit knowledge of T_2 e.g. if his competence with Gg is knocked out, because of the implications this has for the channels of nutrient flow, his competences either with all other sentences containing G, or with all other sentences containing g,

or both, will be knocked out also.

Davies' first point here seems basically correct ; not all causal structure will be germane to the attribution of tacit knowledge, so we need some account which will enable us to discriminate between causal structure that is thus relevant, and causal structure that is not. Nutritional structure is causal structure, but is intuitively irrelevant to the ascription of tacit knowledge. Why is this so ? Davies suggests that this is because "the causal explanatory structure in the example is in no way *sensitive* to the *information* stored in the units", and that this lack of sensitivity is manifested in the fact that the patterns of *revision* of semantic beliefs is unlikely to follow the patterns of semantic decay : "we do not expect that *revision* of C's belief about the meaning of 'Fa' would go hand in hand with corresponding revisions of his beliefs about other sentences".¹¹ So, we need to revise our account of causal structure, and with it the constraint we extracted from Evans's proposal in the previous section, in such a way that the requisite sensitivity to informational content is introduced ; and this, in accordance with the above remarks, is introduced via the notion of revision, so that the earlier constraint thus becomes :

If, and only if, the operative states implicated in the causal explanation of a speaker's beliefs about the meanings of S_1, \dots, S_n are jointly sufficient for a causal explanation of his belief about the meaning of S ; *and* those first states together with the revision of the speaker's belief about the meaning of S provide an explanation of the speaker's corresponding revisions in his beliefs about the meanings of S_1, \dots, S_n – should the semantic resources sufficient for the canonical derivation of truth conditions specifications for S_1, \dots, S_n be sufficient for the canonical derivation of a truth condition specification for S .

But does the introduction of the notion of sensitivity to information really solve the difficulty which Wright raised ? I want to suggest that it does not, and that in fact Davies has misunderstood the character of Wright's objection here.

Wright's problem was *not* that some patterns of causal structure were irrelevant to tacit knowledge ascriptions, but that *even if* we could find a good constitutive account of tacit knowledge in which some form of causal structure *was* relevant to its ascription, this still would not justify articulation in the derivational structure of a semantic theory. Even if we *grant* the assumption that certain of the causal interrelations amongst speakers' competences *are* worth describing, the objection raised is that we needn't run the risk of having the structure of these

interrelations reflected indirectly via the derivational structure of a theory of meaning. Such structure would be equally well described (in the case of the language L) by a listiform theory like T_1 supplemented with a rider along the following lines : speakers are generally able to understand novel utterances provided they only involve familiar semantic primitives, and changes in their semantic beliefs about a sentence tend to be associated with changes in their semantic beliefs about all sentences containing one or more of the semantic primitives figuring in that sentence. More *detail* can then be obtained via the recursive syntax which was initially wedded to the listiform semantic theory. It might be worthwhile to pause briefly and investigate precisely how the relevant detail could, in the suggested way, be brought to light.

Wright suggests that the recursive syntax will provide this detail of the condition that it itself satisfies the mirror constraint. Now what does it mean to say that a *syntax* satisfies the mirror constraint ? – that constraint is framed in terms of truth conditions specifications, and a purely syntactical theory will presumably have *no* such specifications. I suggest the following reconstrual of the mirror constraint for syntactical theories :

If, and only if, the causal states implicated in the causal explanation of a speaker's beliefs about the meanings of S_1, \dots, S_n are jointly sufficient for a causal explanation of his belief about the meaning of S ; *and* those first states together with the revision of the speaker's belief about the meaning of S provide an explanation of the speaker's corresponding revisions in his beliefs about the meanings of S_1, \dots, S_n – should the syntactic resources sufficient for the canonical derivation of wff specifications for S_1, \dots, S_n be sufficient for the canonical derivation of a wff specification for S.

If a syntactical theory satisfies this then *its* derivational structure (the canonical routes from its axioms or base clauses to its specifications of well-formedness) will mirror with at least as much clarity the causal structure of speakers' competences which was initially mirrored in the derivational structure of the semantic theory. It is important to note that this objection holds good even where the causal structure *is* sensitive to the informational content stored in the units of the representation of a semantic theory. Davies betrays his misunderstanding of this point in the following passage :

"it may be that, by Wright's lights, no refinement of the notion of causal structure could ever justify the idea that a *semantic* theory should mirror that structure. But, to the extent that any refinement ensures that the salient causal

I disagree with this – even if we find that the salient causal structure can in fact be described as an information processing structure, the objection still stands. The problem is not one of answering the question : how does *causal* structure justify articulation in a semantic theory ?, but rather of answering the question : how does causal structure justify articulation in a *semantic* theory ?

Thus, notwithstanding the fact that Davies' revised account provides a useful sharpening up of the notion of causal structure considered as relevant to the attribution of tacit knowledge, I would suggest that that revision leaves Wright's objection, properly read, more or less completely untouched.

How then can we deal with Wright's objection ? The motivating thought behind that objection was simply that the causal structure we are concerned to reflect can be reflected by a purely syntactical i.e. non semantic theory, and that some additional reasons therefore have to be provided to ground the preference for reflection in theories of meaning. I think we can undercut Wright here by *denying* that he has in fact shown how to reflect the salient causal structure in a non-semantic theory. Let's look more closely at the conjunction : listiform semantic theory plus syntactical theory which meets the mirror constraint. The latter part of this conjunction will reflect precisely the same causal structure as any compositional semantic theory which satisfies the mirror constraint. But is it actually non-semantic in nature ? I would say that it is not – that if we stipulate that the syntactical theory must meet the mirror constraint (as modified to apply to such theories) then such a theory is no longer *purely* syntactical. Agreed, the theorems which form the output of this theory are concerned solely with the well-formedness or otherwise of the sentences of the language ; but what a theory is *about* , what *sort* of theory a given theory *is* , is determined not only by the content of the sentences which make up its output, but also by the constraints in accordance with which that output is generated. For example, suppose we have a theory A whose output consists of sentences detailing the amount of money possessed by a sample of 100 Scottish women, and a theory B whose output consists of sentences detailing the amount of money possessed by a sample of 100 people but which has been constructed in accordance with the constraint that the people included in the sample space should all be Scottish women. Then, even though the theory B's output is couched in pronouncements of the form "Person X has £x" which feature no mention of Scotland or of women, there is a clear sense in which that theory is still

about Scottish women. To get back to the linguistic case, we should note that the mirror constraint is a *semantic* constraint – it makes reference to beliefs about the *meanings* of sentences and to relations that obtain between beliefs about the *meanings* of sentences. So, any theory which is required to satisfy the mirror constraint will be a semantic theory in the sense that it will encode semantic information, including, despite appearances, any theory whose output mentions only the grammaticality of the language's sentences.

I would thus deny that Wright has shown how the appropriate causal structure can be reflected by a non-semantic theory : what he has given us is in fact an account of how that structure can be reflected in another *semantic* theory. Indeed, it is perhaps misleading even to speak of *another* semantic theory here : just as the theory B above seems to be little more than a *reformulation* of the theory A, given the extensional equivalence of T_1 and T_2 , the conjunction of T_1 with a recursive "syntax" which is really partially semantic seems to me to be little more than a reformulation of the explicitly semantic T_2 .

3.3 The 5% Difference

In this section I outline the third objection which Wright raises against Evans. We shall see that Davies' more sophisticated account involving the mirror constraint initially evades Wright's objection, and that although an argument which Davies himself provides has a conclusion similar to that of the objection, this does not have the destructive effects which Wright envisages.

According to Evans's original account the job of the theory of meaning is to describe the dispositions which speakers have corresponding to each of the expressions for which that theory provides a separate axiom – but, if this is *all* that the theory is meant to do then the 21st axiom of T_2 – the compositional axiom – ought to be redundant, because someone who has the dispositions described by the other 20 axioms will thereby be disposed to judge correctly of the truth conditions of the sentences of L (this is again a consequence of the fact that the dispositions connected with the names and predicates are inter-defined). However, without the compositional axiom the theory T_2 will be paralysed – it will be impossible to derive any truth conditions specifications for the

sentences of L. So, "the conclusion is that Evans has not shown how we are to construe an articulated semantic theory as a description of speaker's dispositions".¹

The suggested way out of this difficulty is again to switch from talk of descriptions of dispositions to talk of the states underlying those dispositions and of the reflection of causal explanatory structure via the satisfaction of the mirror constraint. The job of the theory of meaning is now viewed, not as the description of certain dispositions, but as the reflection of a certain sort of causal structure. There is then no obstacle preventing the theory T_2 from satisfying the mirror constraint as it was stated on p.45, and hence no obstacle to its reflecting the structure in question and thus doing its proper job.

Let me try to clarify this point with the aid of one of Davies' examples. Consider another semantic theory T_3 . This has the same axioms for the proper names of L as the theory T_2 , but differently styled axioms for its predicates. These axioms will instead be of the form :

a sentence coupling a name with the predicate 'F' is true if
and only if the object denoted by that name is bald.

As Davies puts it "What T_3 does is to parcel out the content of T_2 's compositional axiom among the ten predicates of the language"², so that T_3 is not open to Wright's objection even when we take the Evans line about the description of dispositions. But more importantly, when we take the line in terms of causally operative states and reflection of causal structure, it seems that on this account there is nothing to choose between the ascription of tacit knowledge of T_2 and the ascription of tacit knowledge of T_3 . To see this, define a relation of DS-equivalence (equivalence in point of derivational structure) on theories of meaning as follows :

Two (extensionally equivalent) theories T_k and T_m are said to be DS-equivalent whenever the following holds (for any sentences S_1, \dots, S_n, S) – the semantic resources in T_k which suffice for the canonical derivation of truth conditions specifications for S_1, \dots, S_n suffice also for the canonical derivation of a truth condition specification for S *if and only if* the semantic resources in T_m which suffice for the canonical derivation of truth conditions specifications for S_1, \dots, S_n suffice also for the canonical derivation of a truth condition specification for S .

Then it turns out that T_2 and T_3 are DS-equivalent, for "although T_2 has an extra axiom relative to T_3 , the use of this resource in T_2 is constant across all derivations of meaning specifications for whole sentences".³ Let's clarify this assertion by means of an example. In T_2 the semantic resources sufficient for the derivation of a truth condition specification for Ga and Fb are sufficient also for the derivation of a truth condition specification for Fa :

- | | | |
|----|--|---------------------------|
| Ga | (1) 'a' denotes Harry | (axiom for 'a') |
| | (2) an object satisfies G iff it is happy | (axiom for 'G') |
| | (3) a sentence coupling a name with a
predicate is true iff the object denoted
by the name satisfies the predicate | (compositional axiom) |
| | (4) 'Ga' is true iff Harry is happy | (from (1), (2), and (3)) |
| | | |
| Fb | (1) 'b' denotes John | (axiom for 'b') |
| | (2) an object satisfies F iff it is bald | (axiom for 'F') |
| | (3) compositional axiom (as above) | |
| | (4) 'Fb' is true iff John is bald. | (from (1), (2), and (3)) |
| | | |
| Fa | (1) 'a' denotes Harry | (axiom for 'a') |
| | (2) an object satisfies F iff it is bald | (axiom for 'F') |
| | (3) compositional axiom | |
| | (4) 'Fa' is true iff Harry is bald | (from (1), (2), and (3)) |

We can see then that the resources used in T_2 in the derivations of the specifications for Ga and Fb were : the axioms for 'a','b','F','G', and the compositional axiom. These give all we need in order to derive a specification for Fa . Now, because of the way the compositional axiom in T_2 is built into the axioms for the predicates in T_3 , the same thing holds in T_3 . Witness,

- | | | |
|----|-----------------------|-----------------|
| Ga | (1) 'a' denotes Harry | (axiom for 'a') |
|----|-----------------------|-----------------|

- (2) a sentence coupling a name with the predicate
 'G' is true iff the object denoted by the name is
 happy (axiom for 'G')
- (3) 'Ga' is true iff Harry is happy (from (1) and (2))

- Fb (1) 'b' denotes John (axiom for 'b')
- (2) axiom for 'F' (see p.55)
- (3) 'Fb' is true iff John is bald (from (1) and (2))

- Fa (1) 'a' denotes Harry (axiom for 'a')
- (2) axiom for 'F' (see p.55)
- (3) 'Fa' is true iff Harry is bald (from (1) and (2))

Here the resources used in the derivations of specifications for Ga and Fb were : the axioms for 'a','b','F', and 'G' – and again these give us all we need in order to derive a specification for 'Fa'. We could do this again for all of the appropriate sentences of L, and this would amount to a conclusive proof of Davies' assertion that T₂ and T₃ are DS-equivalent.

We thus find ourselves in the following position : because of the DS-equivalence of T₂ and T₃ either both theories satisfy the mirror constraint for a given speaker, or both theories fail to satisfy the constraint for that speaker, so that the 5% difference – the difference between the 20 axioms of T₃ and the 21 axioms of T₂ – *doesn't* in fact matter. Given that we are not concerned with the description of dispositions or of dispositional states apart from their position in a causal web whose structure we are concerned to reflect, the fact that we can describe that causal structure in theories which do not devote an individual axiom to name-predicate concatenation is innocuous : it does not vitiate its reflection in theories which *do* contain such an axiom.

However, having given Davies' reasons for rejecting Wright's objection, I now outline another argument which Davies provides which has the effect of undoing the response above.

Suppose we are trying to decide whether a given articulated semantic theory satisfies the mirror constraint with respect to a particular speaker of L. Suppose also that that speaker revises his belief

concerning 'Fb' from 'John is bald' to 'John is baldish'. Then, in accordance with the mirror constraint, we will look and see whether he revises his beliefs about 'Fa', 'Ga', and 'Gb' *correspondingly*. But what counts as the *corresponding* revision of, say, 'Gb' ? Is it the null revision, which leaves the speaker with the belief that 'Gb' means that John is happy, or is it the revision with the belief that it means that John is happyish ? No one of these answers seems to be uniquely correct, and according to Davies this latent indeterminacy is a time bomb which threatens the stability of his proposed account of tacit knowledge. But this threat seems to dissipate somewhat when we note that the indeterminacy can in fact be resolved "according as we look at the semantic properties of a language through the grid of one theory rather than the other".⁴ Precisely how this resolution is achieved can be seen from the following quotation from one of Davies' earlier papers – to say that the revision of the semantic belief concerned with S_i corresponds (considered from the viewpoint of a particular semantic theory) to the revision in the semantic belief concerned with S is to say that :

If A were to revise his belief about the meaning of S in that respect of the meaning which the semantic theorist discerns as a deductive consequence [in the semantic theory in question] of the presence in S of the syntactic item ∂ assigned the semantic property π , and if what A believed about the meaning of S as the result of this revision were to be the deductive consequence of a revision of π to π' , then A would revise his belief about the meaning of S_i (and the meanings of any other sentences containing ∂) in such a way as would be the deductive consequence (in the theory in question) of the assignment to ∂ of π' rather than π .⁵

Given this, we can now say which revisions merit the ascription of tacit knowledge of T_2 , and which merit the ascriptions of T_3 . And it seems that these will not coincide :

"If we consider T_2 , then there are two possible changes in the proper axioms, each of which would result in 'Fb' being assigned the meaning that John is baldish. A change in T_2 's axiom for 'F' has one pattern of consequences ; a change in the 21st axiom – the compositional axiom – has a different, and more extensive pattern of consequences. If we consider T_3 , on the other hand, then there is only one possible change to the proper axioms which would result in 'Fb' being assigned the meaning that John is baldish."⁶

So, our original account, in which the ascription of tacit knowledge of T_2 simply was ascription of tacit knowledge of T_3 , and vice-versa, will have to be revised in such a way as to take this into account. For, consider the following two speakers A and B :

"for whom, a form of the language of thought hypothesis is true. For these speakers, language comprehension – in particular, the assignment of meanings to sentences – is a matter of derivations in a semantic theory explicitly represented in a special purpose language processing system. Suppose that speaker A conducts on his inner blackboard derivations in theory T_2 , while speaker B conducts derivations in theory T_3for the purposes of tacit knowledge ascriptions speakers A and B are grouped together."⁷

But now :

"just as theory T_2 with its 21 axioms provides an extra locus of content sensitivity over theory T_3 with its 20 axioms, so the causal explanatory structure in speaker A provides an extra locus of systematic revision over the causal explanatory structure in speaker B. So, not altogether surprisingly, it is speaker A – conducting inner derivations in theory T_2 – who meets the condition for tacit knowledge if the indeterminacy is resolved by looking at the language through the grid of T_2 . And it is speaker B who meets the condition if the indeterminacy is instead resolved by looking at the language through the grid of theory T_3 ."⁸

What is going on here ? The suggestion is I think that *only if* we have a one-one correspondence between the axioms of a theory of meaning and the explanatory loci of systematic revision which go towards making up the causal explanatory structure in a speaker can we spell out satisfactorily the notion of systematic revision amongst the implicit semantic beliefs that constitute that speakers linguistic competence ; for, if the axioms for the language's expressions and the speaker's competences with those expressions correspond with each other one by one, then, since "For each axiom or rule, the required notion of *systematic revision* can be spelled out in a quite determinate way"⁹, similarly determinate sense can be made of the notion of systematic changes in the nature of the speaker's competences i.e. in his revisions of his semantic beliefs. If this is correct then the 5% difference *will* matter – A will be seen as having tacit knowledge of T_2 (and not T_3) while B will be viewed as having tacit knowledge of T_3 (and not T_2).

Although I am suspicious of the details of Davies' example of the speakers A and B – it is unclear for instance whether A really has the 21 causal explanatory states required for tacit knowledge of T_2 (what state now corresponds to the compositional axiom ?), and witness the hardly uncontroversial assumption concerning the language of thought hypothesis – I think we can accept that he has provided a good argument to the effect that a speaker with causal states underlying Evans' dispositions can only be ascribed tacit knowledge of T_3 and not of T_2 , because in the latter case we would have no means of resolving the indeterminacy which surrounds the possible revisions such a speaker could make in his semantic beliefs. However, as Davies himself would have us notice, given T_3 , this amounts only to a minor revision, and certainly not to a wholesale rejection, of the account of tacit knowledge which has its roots in Evans' suggestions.¹⁰

3.4 The mirror constraint and a perspective on semantic facts

Under the mirror constraint, a semantic theory reflects the causal explanatory structure to be found in speakers' competences with the language concerned. But we have not yet provided any reasons which would suggest that such causal structure is worth reflecting in the first place ; and, intuitively, what we really ought to be concerned with is *semantic* structure and the provision of a rationale for *its* reflection in a theory of meaning. If what we are after is semantic structure, why bother with the causal explanatory structure which underlies a competent speaker's capacities for use of the language ? Martin Davies makes what is, in effect, the same point, when he writes that

"The project which the mirror constraint governs is not the project of the semantic theorist. His project is, surely, that of providing the structurally most revealing theory of meaning for a particular language, and not that of providing the structurally most revealing description of a particular speaker's competence with that language."¹

My response to this problem will have, broadly speaking, two parts. Firstly, I will argue that the semantic structure of a language, on one influential view of what semantic facts *are* , is most naturally viewed as being isomorphic to a particular sort of causal structure underlying one of its competent speaker's abilities. Because I think that that view of what semantic facts are is the

correct one, I shall claim this as evidence in favour of my account of semantic structure ; but I shall not attempt to *defend* that view here. Secondly, in 3.5, I shall argue that viewing semantic structure in this way does provide us with a strong motive for reflecting structure in a semantic theory – for a semantic theory aims to give a representation of speakers' abilities, and since, as I shall argue, the structure of the ability constitutive of language mastery is essential to its identity, any representation which keeps silent on the structure of that ability is to that extent incomplete. This will also provide us with a response to one objection which might be levelled against our constitutive account.²

Why, then, should we equate semantic structure with the sort of causal explanatory structure we spoke of in the sections above ? For a believer in what Elizabeth Fricker³ has termed the "Metaphysical Perspective", "semanticalism", the thesis that semantic facts are true in virtue of autonomous semantic states of affairs, is false : if there are any determinate semantic facts then they must owe their existence to the obtaining of facts of a less mysterious nature, and for which we have something like an acceptable epistemology. That is to say, semantic facts must be *grounded* in other, non-semantic facts. What other facts ? Plausibly, facts about human behaviour, construed in physical, non-intentional, and non-semantic terms. Precisely *how* the facts about human behaviour, thus construed, serve to determine the semantic facts, has been the source of much recent discussion of radical interpretation, and the question as to how such determination is possible and the methods whereby it can be in principle effected is not one that I shall attempt to answer in detail here. Rather, for the purposes of what follows, I will accept as providing a satisfactory rough outline of an answer to that question, the picture presented by Christopher Peacocke.⁴ What we have, on this picture, is the following : human beings, described *both* as the bearers of beliefs and desires and other propositional attitudes, *and* as the sources of various speech acts and assorted pieces of behaviour described in intentional terms. Now according to Peacocke each of these components (we might say, a mental component and a semantic component) is linked to the other by a series of a priori principles which collectively exhaust the content of the semantic and mental concepts ascribed, and any acceptable theory of interpretation must exhibit those concepts as satisfying these principles. But of course, this is not enough : if these principles were the *only* constraints on interpretation then it would be possible, as Fricker says, that "...the same internally consistent theory could be imposed on all communities".⁵ There

is thus a need for another brand of a priori principle, this time delineating the conditions under which a given theory can plausibly be said to have satisfied the physically described behaviour. We thus have *two* sorts of principle at work in the determination of an interpretation : principles which link the semantic and mental concepts to each other, and principles which define what it is for a given scheme of interpretation to fit the behaviour construed as brute unintentional movement.

For someone who accepts the Metaphysical Perspective the question as to whether the semantic facts are determinate is really the question as to whether the two sorts of a priori principle alluded to above serve to determine, for a given theory of interpretation, whether or not it is correct. On this view, the correctness of a proposed interpretation is *constituted* by its satisfying the relevant range of a priori principles - this contrasts sharply with the semanticists' viewpoint, from which the correctness of a proposed interpretation is constituted rather by the obtaining of some set of *sui generis* semantic states of affairs, and from which the a priori principles have a status more akin to empirical generalisations than constitutive norms. We can perhaps sloganise the view from the Metaphysical Perspective (MP) as follows : *Semantic facts about a natural language have to be appropriately grounded in the behaviour of the speakers of that language.*

However, the following problem now arises for an advocate of MP : how well, if at all, do our pre theoretic semantic notions meet the above requirement ? The biggest such notion concerns the notion of sentence meaning itself - for meaning is an essentially *normative* notion, and the problem is to see how anything could satisfy the above requirement and yet leave scope for this notion of normativity. This, I take it, is the problem that most contemporary discussions of the Wittgensteinian critique of rule following find themselves facing, and I will not attempt to address this in the present essay. Rather, I want to concentrate on the question as to how well our pre-theoretic intuition that natural languages are *semantically structured* stands up to the requirement laid down by MP, in other words on answering the question : how is the semantic structure (if there is one) of a natural language grounded, in the appropriate way, in its speakers' behaviour ? If it turns out that *no* satisfactory answer can be given to this question it will follow that there are *no* determinate facts concerning the semantic structures of natural languages, a conclusion surely unpalatable to most, if not all, philosophers of language. So how *do* we answer it ? I propose the following answer :

The semantic structure of a natural language is *constituted* by the causal explanatory structure that underlies its

speaker's abilities to competently use its expressions, and its semantic primitives will be the expressions for which the speakers' understanding of sentences in which they are contained is the product, amongst other things, of *distinct* causal states associated with them.

It is the acceptance of this constitutive account of semantic structure that leads to the imposition of the mirror constraint : this constraint ensures that a semantic theory which satisfies it will provide a detailed description of the underlying causal explanatory structure to be found in the competence structure of one of the language's speakers, and therefore, on our present account, a detailed description of the semantic structure of the language.

It might be worthwhile to pause for a second here, and ask precisely how the answer given immediately above satisfies the demands of the metaphysical perspective. In order to see how this is the case, we need some information on precisely how the facts about causal-explanatory structure are themselves grounded, in appropriate ways, in facts about speakers' behaviour. How, for instance, is the truth of the claim "The operative states implicated in the causal explanation of the intentional state P are sufficient for a causal explanation of the intentional state Q", grounded in facts about behaviour ? I suspect that Evans and Davies both wish to view a claim such as this as a theoretical claim, issuing from a satisfactory "causal, presumably neurophysiologically based, explanation of comprehension". Presumably, such an explanation will be an empirical theory, based on empirical evidence – the evidence mentioned earlier relating to patterns of acquisition, loss, and revision of semantic beliefs. The idea is that we discover – empirically, by observation of behaviour – that certain causal relationships obtain between various components of a speaker's linguistic competence ; and these observations – facts about behaviour – provide us with evidence in favour of a causal theory of competence which attempts to explain these facts by the postulation of a system of causal states standing in specific causal relations to each other. This is why the structure in question is called causal *explanatory* structure. I think it is this point that Davies has in mind when he draws our attention to the distinction between the evidential and constitutive sides of his account of tacit knowledge. The causal structure possession of which constitutes tacit knowledge of a semantic theory is not *constituted* by the facts about acquisition of semantic beliefs etc – rather these facts provide *evidence* for the existence of such a structure, because the existence of such a structure is conceived of as contributing towards the causal explanation of those facts. Of course, it does not follow from this that the existence of the causal structure is independent of the

behavioural facts, because without those behavioural facts there would be no explanatory warrant for the postulation of the structure in the first place.

Anyone wishing to reject the imposition of the mirror constraint therefore has two options : he can either provide an argument against the metaphysical perspective, or he can provide another account of semantic structure, not invoking the mirror constraint and the notion of causal explanatory structure, which nevertheless shows how facts about that structure are grounded in facts about speakers' behaviour. To the extent that neither of these can be provided, our imposition of the mirror constraint will have been vindicated.

3.5 The doctrine of essential linguistic structure

Having shown that the adoption of MP provides a motivation for the identification of semantic structure with causal explanatory structure, we must now face up to a difficulty which seems to follow on the heels of such an identification : namely, that if semantic structure *is* causal explanatory structure, there is no motivation for requiring semantic theories to reflect semantic structure (i.e. no good reason for the imposition of the mirror constraint).

From the viewpoint of MP there is nothing more to meaning than can be abstracted from the behaviour of those who speak the language in question ; thus, a theory of meaning will be, from this viewpoint, a representation of what those speakers abilities consists in. Fricker, following Dummett, entitles such a theory a TRSA : a *theoretical representation of speakers' abilities* , and claims that a theory of meaning must be viewed as such is entailed directly by the acceptance of MP.¹ This fits in well with the conception of theories of meaning which I have been working with in this essay. On my view a semantic theory is also to be viewed as a TRSA, and it achieves the theoretical representation by specifying the content of the intentions constitutive of the recognitional abilities which make up sentential understanding.

However, disturbing results seem to issue from this conception of a semantic theory as a TRSA. If all a semantic theory is, is a representation of an ability, then the constraint that such a theory should specify the structure of that ability is theoretically gratuitous, *unless* it can be shown that the structure in question is "essential to its identity". Mark Sainsbury has argued that generally the

structure of an ability is not essential to its identity, and that consequently no good motivation has yet been provided for imposing the mirror constraint.²

Consider chicken sexing : the ability to associate a young chicken with a number in the range 0 to 1 representative of the probability of its being male. It is, claims Sainsbury, conceivable that there could be two people possessed of the same chicken sexing ability – people, that is, who when presented with the same input (this will consist of various visual and tactile properties of chickens), assign to them the same numerical probability representation – but in whom that ability is realised in different causal structures. One has learnt which number is associated with each type of input by training concerning that particular input, so that that the sub abilities connected with each type of input are causally independent ; while the other's sub abilities are causally interrelated : the causal states underlying some finite subgroup of his sub-abilities are sufficient to enable him to associate "a number with hitherto unperceived combinations of properties along lines reflected in the theory of probability".³ Sainsbury claims that here we have an example where the *same* ability is realised in different *causal* structures, and that this shows that the structure of an ability cannot be viewed as an essential ingredient in its characterisation. And his claim initially seems plausible enough : is it not true that the two chicken sexers in the example are able to do the *same* thing ?

A similar argument can be run in the case of the ability to speak a language. It seems plausible to say, given that the essential role of language is that of *communication* , and given that two people can plausibly be said to be in a position to communicate successfully when they attach the same meanings to the same sentences, regardless of *how* they have come to attach those meanings to the sentences, that a TRSA will be adequate if it gets, for each syntactically described input, the right semantic output : "Formally, it is a function from sentences specified as syntactic strings, to redescriptions of these as speech acts."⁴ But given this, it now seems to follow that a semantic theory, or TRSA, for a language L, needn't discern L's semantic structure – a semantic theory will be adequate if it specifies the meaning of each of L's sentences, and there is no requirement that it needs to do this in a way that reflects the structure of L – there are no further constraints on *how* the meaning specifications for L's sentences are obtained, a result (which Fricker terms LIST) curiously at odds with recent and contemporary semantic theorising, and with the wish to impose the mirror constraint.⁵

Can a defender of MP avoid LIST ? It seems to me that there are *two* distinct ways in which he

might do this, which ways I think Fricker unhelpfully conflates in her account, and either of which might be construed as adherence to a doctrine of "essential linguistic structure".

One way – what I shall call the "strong" doctrine of essential linguistic structure (ELS) – would simply be to deny the premiss that what is criterial for understanding – for membership of a language – is "attaching the right meanings to the right sentences". Specifically, we could deny this by arguing that in addition the structure of the competences underlying those attachments is similarly criterial. But would two people with differently structured competences be unable to *communicate*, even though they attached the same meanings to the same sentences? The prospects for a negative answer to this question at first sight appear bleak, and I shall return to it in due course.

Another, I think distinct, way, would be to *accept* – even if only for the sake of argument – that what is criterial for communication is simply getting the right meanings in the right places, but to argue that if a semantic theory is to successfully capture the appropriate meaning specifications then it must discern semantic structure. This – and I shall call this the "weak" ELS doctrine – would be to deny that the *correctness* of the theory's meaning specifications was independent of facts pertaining to L's semantic structure. This is different to the approach above, because here we are not necessarily committed to denying that what is criterial for successful communication is a common stock of meanings attached to a common stock of sentences : it's just that the interpretation of their language will be indeterminate if there are no constraints involving semantic structure.

I shall return to the strong ELS doctrine later in this section ; before that I want to investigate Fricker's argument for weak ELS, and attempt to counter some objections which Davies has raised against it.

Fricker states her case, tersely and briefly, as follows :

"Insofar as interpretation is determinate, the primary feature ensuring this is the fact that the assignment of a meaning to one sentence constrains and is constrained by the assignments made to other sentences....this interdependence obtains through the presence, amongst the a priori principles, of some form of structural postulate requiring that sentences be seen to be composed out of a stock of semantically primitive elements which make the same characteristic contribution to sentence meanings in all their occurrences....Some principle for discerning structure in sentences is essential to interpreting a language, and the correctness of an interpretation cannot be considered to be

Let's attempt to flesh out this argument with a bit more detail. Suppose, as usual, that we are out to interpret sentences and propositional attitudes of some previously unmet Quincan tribe, and that are assumed to have complete knowledge of all of the physically described behaviour of the natives. Suppose also that we observe one of the natives pointing towards a nearby mountain and uttering "Lo monnadi kel guro". What does he mean? As interpreters we will no doubt initially take the native to be saying something about the nearby peak, and let's suppose, for the sake of argument, that this is in fact what he is trying to do. He might have been saying any number of things about it, maybe that it is icy, that it would be exciting climbing it, that it would command a good view of the neighbouring terrain, and so on. What enables us to decide which of these interpretations is the correct one is the contribution the syntactically distinguished predicate "guro" makes to the meanings of the other sentences in which it appears. For example, suppose at some later time after his utterance of the sentence above we observe the native picking up a glass of milk containing ice cubes and uttering "Lo pazo kel guro"; then the most plausible interpretation of the earlier sentence will be that he was claiming that the mountain in view was icy. (This is on the assumption that the native doesn't believe that glasses of ice-cube filled milk make exciting climbing material etc - this is where the principle of charity comes into play: we certainly don't believe that sort of thing so we won't normally settle for an interpretation that attributes that sort of belief to the native). It might appear that there are alternative interpretations of the sentence possible (e.g. we might think he could be interpreted as saying that the mountain was white) but the method whereby such alternative explanations can be eliminated is clear. But note that crucially we can only employ this method on the assumption that the predicate "guro" makes the same systematic contribution to the meanings (or truth conditions) of the sentences in which it figures, and to say that a language contains expressions that make *recurring* and *systematic* such contributions is just to say that it is semantically structured. The interpretation of one sentence constrains, and is constrained by the interpretation of others *via* the presence of such expressions, and so *via* the presence of semantic structure.

The conclusion of Fricker's argument then, is that the *correctness* of the ascriptions of meanings to sentences made by a semantic theory is *not* independent of the semantic structure discernible in the language for which the theory is being constructed, because if there were no such structure the interpretation would in the end be massively indeterminate with respect to the assignment of

meanings to sentences. If a language has no semantic structure it will not be possible to give it a determinate interpretation.⁷

How good is this argument? Martin Davies has suggested that although it does establish that discernment of semantic structure is a *practical* requirement on the process of interpretation, this practical requirement does not amount to the *absolute* claim that Fricker appears to want to argue for. As he puts it himself:

"...semantic structure in a language...crucially facilitates radical interpretation. But it does not quite demonstrate that radical interpretation depends upon there being structure in the speaker's ability. The conclusions about prima facie practical requirements are consistent with the possibility of a language that is partially or wholly lacking in semantic structure."⁸

If a language had no discernible structure, the idea seems to be, then *we* could never obtain a determinate interpretation of its sentences. But this does *not* show that having structure is part of what *constitutes* its being determinately interpretable. Davies' thought seems to be that the threatened indeterminacy is in fact epistemological in nature rather than ontological: *we*, in the absence of structure, will be unable to *find out* which interpretation is the correct one, but this does not as yet amount to a demonstration that there *is* no correct interpretation. Davies' idea – if we interpret him as being charitably disposed towards the form of Fricker's argument – is that what Fricker needs for this conclusion is not so much an argument concerning the practicalities of radical interpretation considered from the viewpoint of rather epistemically limited creatures like ourselves, but instead an argument concerning the viewpoint of a would-be interpreter with *unlimited* access to the facts about the natives' brute behaviour, or less dramatically an argument to the effect that *only if* the a priori principles constitutive of interpretation contain structural postulates will those principles effectively serve to pick out a correct scheme of interpretation *no matter from whose viewpoint the actual process of interpretation eventually takes place*.

However, Davies fails to see that this is *precisely* how Fricker sees the argument from MP as proceeding, because in her argument we are in fact conceived of as having access to *all* the relevant behavioural facts. She makes it quite plain that this is her intention when she says that "the scenario of radical interpretation...is only a heuristic device for thinking about the question, which is not an epistemic one: How do *the facts* (about behaviour) determine *the facts*, (about meanings)?"⁹, and when she points out that from MP the interpretative principles involved are not

to be viewed as mere empirical generalisations which facilitate the would-be interpreter's task, but as a priori principles the satisfaction of which is in a strong sense constitutive of an interpretation's correctness. We can perhaps get clearer on where Davies errs here, by schematising Fricker's argument thus :

(1) Suppose we have knowledge of all of the facts about the natives' physically described behaviour.

(2) Without a structural postulate – some means of discerning semantic structure in the sentences of the native language – we could not arrive at a determinate interpretation of their sentences and propositional attitudes.

So, (3) Without a structural postulate the facts about behaviour do not serve to render the semantic facts determinate.

So, (4) If a language has no semantic structure, then it has no determinate interpretation.

We can see now that Davies has simply misunderstood the role which the allusion to epistemology plays in Fricker's argument – his criticism parallels the criticism of Kripke which McGinn makes when he claims that there are two sorts of scepticism in play in the Kripkean "bizarre sceptic's" attack on the factuality of the notion of meaning and its cognates, namely, an epistemological scepticism and a "constitutive or metaphysical" scepticism.¹⁰ He fails to realise that because the epistemology has been idealised, there is no metaphysical or constitutive problem *distinct* from the epistemological problem : epistemology is for Kripke (and, in the context of her argument, for Fricker) a route *into* metaphysics, so there simply is no room, given that the "practical requirements" are married to an idealised epistemology, for the sort of gap between practical and absolute requirements which Davies envisages.

So I think that Fricker's argument is in fact sufficient to block the conclusion LIST. A semantic theory must be structure reflecting, otherwise it will not even be able to assign a determinate meaning to each sentence of the language under scrutiny. But there is of course the residual problem stemming from the argument to LIST itself : granted, a language must have a semantic

structure to be determinately interpretable, but if semantic structure simply is the causal explanatory structure underlying a speaker's abilities, and if, as the argument to LIST suggests, the same ability can be realised in different causal structures, the semantic theorist whose project is governed by the mirror constraint is in danger of being led into oblivion : a language will have as many semantic structures as there are different causal structures underlying its speakers' abilities, so that all talk of the reflection of *the* semantic structure of the language could simply drop out of the picture.

Fortunately, however, it is not too difficult to see how the argument to LIST is flawed.¹¹ To see this we have to reflect a little on the use of the notion of *communication* in the argument to the LIST result. That argument, as presented by Fricker and Sainsbury, explicitly takes as the criterion for a shared language the possibility of speakers successfully communicating with each other : A speaks the same language as B iff A can successfully communicate with B, and vice versa. But this now seems much too simple minded : what is required for the existence of a shared language is not so much the possibility of communication at a given point in time, but rather *the possibility of communication over an extended period of time*. Thus, what will be relevant in determining whether or not A and B speak the same language will not only be the meanings they assign to sentences at a given time but also the meanings they *will* assign to sentences in the future given similar subsequent exposure to new linguistic phenomena : the sort of communication required for the existence of a shared language is essentially *cross temporal* .

Thus, in deciding whether or not A and B speak the same language we will have to take into consideration the conditionals true of them that lay down what ascriptions of sentence meaning result from any further linguistic training and exposure (we might call these conditionals the *exposure conditionals*) ; and my contention is that if the exposure conditionals differ to such a degree that cross temporal communication is damaged then A and B can not in fact be counted as speaking the same language even though they *presently* assign the same meanings to the same sentences. And now of course different competence structure will start to affect the question as to whether two speakers are speaking the same language because difference in competence structure effectively implies differences in the exposure conditionals.¹²

For example, consider a speaker A who has tacit knowledge of an articulated theory and a speaker B who has tacit knowledge of a listiform theory. For the sake of argument, suppose that A and B are both, at present, competent with exactly the same sentences. Does this mean that they

have the same ability? No – the causal structure underlying A's linguistic competence is such that he can understand a sentence provided that that sentence only contains familiar semantic primitives and modes of syntactic construction. This is because all the operative states implicated in the causal explanation of his semantic belief about such a novel utterance are already present in the structure of causal states which constitutes his tacit knowledge. So, presented with a sentence containing only familiar semantic primitives, A will be able to understand that sentence. B, however, will not: there is no explanatory route from the causal states *he* already possesses, to an intentional state constitutive of understanding the novel utterance.

A good, but by no means perfect, way of helping to make to this point is via the remark which Wittgenstein makes in §18 of the *Philosophical Investigations*: "Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of new boroughs with straight regular streets and uniform houses." The point Wittgenstein is making here is that a language, like the architecture of a city, is in constant flux. Imagine two apprentices, each provided with a vast (but finite) stock of building material with which they are each to construct a city. Let us suppose that in order to get them underway and to train them in the basic techniques of building, a friendly master-builder builds the initial part of the city for them. He builds the same initial segment in each city. After this initial period of training, the two apprentices are left on their own to get on with the job of construction. Now I would claim that the identity of the city which each apprentice is engaged in constructing is determined not only by the state of the cities at the point in time at which the master builder departs, but also by how each apprentice subsequently goes on to utilise the building materials that have been left at his disposal; if they differ in this respect – if one of them transforms the initial core of the city into a replica of Paris, while the other transforms the initial core into a replica of London – then I think there is a clear sense in which the two apprentices can be said to be living in qualitatively different cities even during their initial period of training. Cities, like languages, are essentially dynamic entities. In order to individuate a city (architecturally at least), we need more than a description of that city in the initial stages of its construction – we need an account of how the initial core is going to be transformed (if at all) in times to come. The linguistic analogue of this account is semantic structure.¹³

Thus, my claim is that the justification for the imposition of the mirror constraint turns on the essentially dynamic, as opposed to static, nature of natural language. A consequence of this may be

that a justification for the imposition of the compositionality constraint may still be lacking in the case of formal languages, but this does not vitiate my account – I hold no brief for the provision of such a justification : I said at the beginning of the essay that I was interested in justifying compositionality in semantic theories for *natural* languages. Not everything that can be shown to be true about a natural language can be shown to be true about a purely formal language, so we should not expect *a priori* that every constraint that can be imposed on theorising about languages of the former sort can be legitimately imposed on theories concerning languages in the latter category.¹⁴

A moral to be drawn from this is that while concentration on artificially constructed formal languages – such as the one that takes centre stage in Evans' and Davies' discussion – may be a useful heuristic device for bringing some of the issues facing philosophers of language more sharply into focus, we should not allow this to blind us to the facts that our reasons for looking at such languages lies in our hope that some light might thereby be shed on the construction of semantic theories for *natural* languages, and that where we ignore the disanalogies that undoubtedly exist between the two sorts of language – the primary one being that people *speak* natural and not formal languages – darkness, rather than illumination, is likely to result.¹⁵

3.6 Causal structure, novel utterances and learnability.

Having attempted to justify the imposition of the mirror constraint on the grounds that it ensures that an essential ingredient in the description of a speaker's competence – an account of underlying causal-explanatory structure – is included in our theoretical representation of his linguistic abilities, I now move on to look at the implications our account has for the provision of answers to the questions (2) and (3) which we took note of at the outset of this chapter.

Can the imposition of the mirror constraint help us to understand speakers' capacities to understand novel utterances, provided they include only familiar semantic primitives and modes of construction ? Suppose that a speaker can come to understand S after exposure to the sentences

S_1, \dots, S_n , that loss of competence with S occasions loss of competence with S_1, \dots, S_n , and that a revision of a speaker's belief concerning the meaning of S occasions corresponding revisions in his beliefs about the meanings of each of S_1, \dots, S_n . Then, on the basis of this evidence we shall claim that : the operative states implicated in the causal explanation of the speaker's beliefs about S_1, \dots, S_n are jointly sufficient for a causal explanation of his belief about the meaning of S – and those first states together with his revision of the belief about the meaning of S are sufficient for an explanation of the corresponding revisions in his beliefs about the meanings of S_1, \dots, S_n . Of course, to make such a claim is not yet to *provide* the causal explanations alluded to, so that we are still left with the question : how is it possible that the operative states implicated could be thus sufficient ? I take it that it is precisely this question that is thought of as being answered by showing how the semantic resources sufficient for the canonical derivation of meaning specifications for S_1, \dots, S_n are sufficient for the canonical derivation of a meaning specification for S , by the provision of a semantic theory in which they are so sufficient. That is, the move from the operative states to the belief about the content of the novel utterance is viewed as being explained by the provision of a theory which generates, on the basis of the information represented by the relevant operative states – represented by the axioms of the theory – a meaning specifying theorem which gives the content of the novel utterance in question. So, the provision of a compositional theory of meaning is supposed to make explicable the notion that a speaker could come to form an implicit belief about the content of a novel utterance on the basis of a pre-existing set of causally operative states, because the route from the operative states to the belief about content is reflected by the derivational route from the axiomatic base of the semantic theory to the appropriate meaning specification. Thus, the construction of a compositional semantic theory, and the notion that speakers might have, in the requisite sense, tacit knowledge of such a theory, is supposed to provide an explanation of how those speakers could understand previously unencountered sentences.¹

As a consequence of this, it follows that the ascription of such tacit knowledge can also provide us with an answer to question (2). To see this, we have to get clear on what we mean by learnability. Following Davies, I shall say that a language L is learnable when it is possible for one of its speakers to come "to come to know the meanings of all the sentences of L by way of

exposure and projection".² Thus, we can say that a language is learnable when one needs explicit training with only a relatively small number of sentences in order to secure competence with a possibly very large number of sentences outwith that set – that is, when it is possible to understand the vast majority of sentences simply on the basis of competence with a distinct set of sentences. So we can now see that on this characterisation of learnability, to say that a language is learnable is just to say that its speakers can understand novel utterances, without explicit training in their use : questions (2) and (3) are, therefore, for natural languages at any rate, the same question.

Given the above, it follows that if our account can explain the capacity to understand novel utterances, it can thereby render explicable the learnability of language, as we have conceived it. In fact, we can show that if a language is learnable, then a theory of meaning for that language which satisfies the mirror constraint can have only finitely many proper axioms.³

In order to prove this we have to introduce some terminology. Suppose we have a theory of meaning $M\Omega$ for a language L_1 which contains the following axioms :

(M1a) S_1 means (in L_1) that snow is white

(M1b) S_2 means (in L_2) that the earth moves

(M2) $(\partial)(f)(p)(q)[((\partial$ means (in L_1) that p) & (f means (in L_1) that q)] \rightarrow (" ∂ & f " means (in L_1) that p & q)]

and from which we can derive the following theorem

(T1) " S_1 & S_2 " means (in L_1) that snow is white and the earth moves

We shall say that an axiom or theorem deals non-semanticly with a sentence if and only if it explicitly specifies, in non-semantic terms, what the sentence means ; and that an axiom or theorem deals non-semanticly with a subsentential expression if and only if that axiom or theorem explicitly specifies a semantic property of that expression by specifying, in non-semantic terms and in full generality, the contribution that expression makes to the meanings of sentences in which it

occurs.

Thus, each of the clauses above deals non-semantically with S_1 , S_2 , "&", and " $S_1 \& S_2$ " respectively. This is in contrast, say, to the theorem

$$(T2) (p) (q) [((S_1 \text{ means (in } L_1) \text{ that } p) \& (S_2 \text{ means (in } L_1) \text{ that } q)) \rightarrow ("S_1 \& S_2" \text{ means (in } L_1) \text{ that } p \& q)],$$

which specifies the meaning of e.g. " $S_1 \& S_2$ " only in terms of the meanings of S_1 and S_2 , and, hence, semantically.

Another piece of terminology we shall need is this : we shall say that the meaning specifications for S_1, \dots, S_n d-determine the meaning specification for S if and only if the semantic resources sufficient for the (canonical) derivation of the meaning specifications for S_1, \dots, S_n are sufficient for a (canonical) derivation of the meaning specification for S .

We can now state our first claim :

CLAIM 1 : (a) If a theorem which is not an axiom deals non-semantically with an expression B then there is a sentence S containing B and sentences S_1, \dots, S_n not containing B such that the meaning specification for S is d-determined by the meaning specifications for S_1, \dots, S_n .
(b) In contrast, if an axiom deals non-semantically with B then there are no such sentences S and S_1, \dots, S_n .

I will not attempt to prove this claim in its full generality here, but will look at an example which should render it plausible.

The theorem T1 deals non-semantically with the expression " $S_1 \& S_2$ ", and it is derived using the axioms for S_1 , S_2 , and "&". Using these axioms we can derive meaning specifications for S_1 and S_2 , neither of which contain the expression " $S_1 \& S_2$ ". So the meaning specifications for S_1 and S_2 jointly d-determine the meaning specification for " $S_1 \& S_2$ ". So for the case in hand there is a sentence S containing B (" $S_1 \& S_2$ "), and sentences S_1 , S_2 , such that the meaning specification for S is d-determined by the meaning specifications for S_1 and S_2 .

Why doesn't this work when we're dealing with an axiom? Let's try it for the axiom S_1 . Clearly there are no sentences not containing S_1 whose meaning specifications could d-determine the meaning specification for S_1 , because using the axiom for S_1 alone we cannot derive meaning specifications for sentences other than S_1 . So the meaning specification for S_1 is not d-determined by the meaning specifications for sentences not containing it.

Having (hopefully) rendered Claim 1 at least plausible, we now go on to make a couple of definitions.

Definition 1 : β is a semantic primitive according to a theory Ω if and only if an axiom of Ω deals non-semantically with β .

Definition 2 : β is a semantic primitive of a language L if and only if β is a semantic primitive according to some theory of meaning for L which meets the mirror constraint.

Using these we can now go on to state our second claim :

CLAIM 2 : β is a semantic primitive if and only if it is not possible to come to know the meaning of any sentence of L which contains β simply by exposure to sentences not containing β .

The proof of this falls into two parts :

"If" part : Suppose that β is a semantic primitive of the language L . Then, by definition 2 above, there is a theory Ω for L which meets the mirror constraint, and of which β is a semantic primitive.

By the previous claim 1, there are no sentences S containing β and S_1, \dots, S_n not containing β such that the meaning specification for S is d-determined by the meaning specifications for S_1, \dots, S_n . Since our theory of meaning Ω was assumed to satisfy the mirror constraint, it now follows, from the statement of that constraint, that there is no causally explicable route from the semantic beliefs about the meanings of sentences of L not containing β to a semantic belief about the meaning of any sentence of L which does not in fact contain it.

"Only if" part : Suppose that there is no causally explicable route from beliefs about the meanings of sentences of L not containing β to a belief about the meaning of a sentence of L which does contain β . Then, from the statement of the mirror constraint, there are no sentences S containing β and S_1, \dots, S_n not containing β such that the meaning specifications for S is d-determined by the

meaning specifications for S_1, \dots, S_n . Now in order to reach the conclusion that there is an axiom of Ω which deals non-semanticly with β , we need what is in effect the converse of the claim 1(b) above. Is such a claim actually plausible? We need to show that if there are no sentences S containing β and S_1, \dots, S_n not containing β such that the meaning specifications for S_1, \dots, S_n d-determine the meaning specification for S , then there is an axiom which deals non-semanticly with β . We can prove this by proving the contrapositive. Suppose that there is no axiom which deals non-semanticly with β . Then, on the assumption that ^{of} the language's expressions must be dealt with semantically somewhere (otherwise there will be expressions on whose semantics Ω will keep silent), there must be a *theorem* which deals non-semanticly with β . It then follows by claim 1(a) that there is a sentence S containing β and sentences S_1, \dots, S_n not containing β such that the meaning specification for S is d-determined by the meaning specifications for S_1, \dots, S_n . This proves the contrapositive, and the desired result now follows immediately.

Before moving on to our next claim we need another definition.

Definition 3 : β is a semantic constituent of μ according to Ω if and only if the derivation in Ω of the theorem which deals non-semanticly with μ employs an axiom or theorem which deals non-semanticly with β .

Given this we are now in a position to state

CLAIM 3 : If R_1 and R_2 are sentences and R_1 is a semantic constituent of R_2 in L , then knowing what R_2 means is sufficient for knowing what R_1 means (or equivalently, knowing what R_1 means is necessary for knowing what R_2 means).

The proof of this is as follows. Suppose that R_1 is a semantic constituent of R_2 , and that a speaker knows what R_2 means. Suppose also that we have a theory Ω which satisfies the mirror constraint. Then, by definition 3, the derivation in Ω of the theorem which deals non-semanticly with R_2 employs an axiom or theorem which deals non-semanticly with R_1 . This means, in effect, that the meaning specification for R_2 d-determines the meaning specification for R_1 . Hence, from the statement of the mirror constraint, there is a causally explicable route from the speaker's

knowledge of the meaning of R_2 to knowledge of the meaning of R_1 .

Our next claim follows immediately from claim 3 :

CLAIM 4 : If a speaker knows what a sentence S means he must have had some training with sentences containing each of the semantic primitives which are semantic constituents of S .

We now need our final definition :

Definition 4 : A language is said to be learnable just in case there is a fixed and finite subset of its sentences from which there is a causally explicable route to knowledge of the meanings of every one of its sentences,

which enables us to make

CLAIM 5 : If a language L is learnable then there must be a finite set of sentences of L containing as semantic constituents all the semantic primitives of L .

This is proved as follows. Suppose that language L is learnable. Then a speaker can come to know the meanings of all the sentences of L by exposure to a subset of the overall stock of sentences. By claim 4, the speaker must have had some training with a set of sentences containing *all* the semantic primitives of L . Claim 5 follows immediately.

It now follows trivially that :

CLAIM 6 : since each sentence contains only finitely many semantic primitives, a learnable language can have only finitely many semantic primitives.

Before we can finally reach our desired conclusion, we need :

CLAIM 7 : If a theory of meaning Ω for L meets the mirror constraint, then, in effect, a single proper axiom of Ω deals non-semantically with each unambiguous semantic primitive for L .

We prove this as follows. Suppose that Ω has two proper axioms which deal non-semantically with an expression β . Either the two axioms are always employed together in the derivations of meaning specifications for sentences containing β as a semantic constituent, in which case the two proper axioms should be combined into a single axiom.⁴ Or, on the other hand, there is an axiom which is employed in the derivations of meaning specifications for some, but not all, sentences

containing β . Then, since Ω satisfies the mirror constraint, there are two batches of sentences containing β such that a further course of exposure is needed to proceed from the one to the other. To see this, suppose that A and B are two batches of sentences containing β , and that the meaning specifications for the sentences in A are derived using an axiom which is not employed in the derivations of meaning specifications for the sentences in B. Thus, the meaning specifications for batch B do not d-determine the meaning specifications for batch A. Hence, from the statement of the mirror constraint, there is no causally explicable route from knowledge of the meanings of the sentences in batch B to knowledge of the meanings of the sentences in batch A. So further exposure will be required. In this latter case, β is then really an instance of lexical ambiguity, and so we should really regard the occurrences of β in each batch as the occurrence of two distinct semantic primitives.

It now follows that :

CLAIM 8 : A theory of meaning for a learnable language which meets the mirror constraint will have only finitely many proper axioms,

which is immediate, given claim 5, and the fact that the language has only finitely many unambiguous semantic primitives, which follows straightforwardly from claim 6.

3.7 The structural constraint and some thoughts on computer programming.

In this section I move on to look briefly at another possible constraint which might be imposed on semantic theorising. This is what Davies has called the "structural constraint", and it can be stated as follows :

If, but only if, it would be possible for a speaker of L to proceed by rational inductive and deductive means from knowing what S_1, \dots, S_n mean to knowing what S means, should the semantic resources sufficient for the canonical derivation of meaning specifications for S_1, \dots, S_n be sufficient for the canonical derivation of the meaning specification for S.¹

Under this constraint, a structurally revealing semantic theory is not conceived of as reflecting the

causal explanatory structure which underlies actual speakers' competences with the language, but as reflecting the rational – inductive and deductive – logical interrelationships that obtain between states of knowledge of the meanings of the the language's sentences.

However, it seems to me that this constraint is really only another version of Davies' mirror constraint. To see this, we need only note that there is an a priori argument which suggests that causal explanatory structure and rational inferential structure are themselves always isomorphic.

Suppose that a speaker can proceed from knowledge of the meanings of S_1, \dots, S_n to knowledge of the meaning of S by rational inductive and deductive inference. Then, in order for him to actually make this inferential move, the operative states implicated in the causal explanation of his beliefs about the meanings of S_1, \dots, S_n must suffice for a causal explanation of his belief about the meaning of S – if they didn't how could he make such a move in the first place? A necessary condition for my inferring, say, knowledge of P from knowledge of Q , is surely that my knowledge of Q provides a causal basis for my knowledge of P .

Suppose, conversely, that the operative states implicated in the causal explanation of the speaker's beliefs about the meanings of S_1, \dots, S_n are indeed sufficient for a causal explanation of his belief about the meaning of S . What we mean by this is that the operative states which determine the content of S_1, \dots, S_n already suffice to settle the content of S : and it is difficult to see how this could be possible if it were not possible to move, via a rational inductive and deductive route, from the information codified in the former set of states to that codified in the latter. A necessary condition of viewing knowledge of P by itself as causally *explaining* knowledge of Q is that it at least be *possible* to rationally infer knowledge of Q from knowledge of P , otherwise the notion of explanation in play would seem to have a very poor pedigree indeed.

Thus, given that Davies wishes to view the notion of causal-explanatory structure in play here as an information processing structure, I would suggest that there is no significant difference between the structure reflected by the ^{structural} ~~mirror~~ constraint and that reflected by the mirror constraint.

There are also close connections between our account, and the suggestion which Wright makes at the end of section III of "Theories of meaning and speakers' knowledge". Wright complains that whereas the account of tacit knowledge given by Evans – and, he might have said, the account which uses Davies' mirror constraint – requires the semantic theorist to pay attention to

empirical facts about language acquisition, loss, and revision, *actual* semantic theorising seems to have proceeded in happy ignorance of such facts. Loth to conclude that the right account of the semantic theorists task "has not greatly impinged upon the consciousness of workers in the field", Wright suggests that the following account of what theorists of meaning ought to be doing perhaps harmonises rather better with what they have in fact been content to do.

Speakers are able to understand novel utterances, provided they include only familiar semantic vocabulary, and the semantic theorist's task is to give an answer to the question "How do they do it?". This question is viewed as dividing into two sub-questions. (a) How is it possible for there to be a device which could, when fed with information concerning the visible or audible structure of a sentence, process that information in such a way that a theorem about the meaning of the sentence is generated?, and (b) How is it possible for such a device actually to be embodied in normal, competent speakers of a natural language? Question (a) is conceived of as being answerable a priori, independently of the empirical evidence upon which Davies and Evans lay so much stress – it will have been answered when a suitable "computer program" has been written, the writing of which will not, as Wright puts it, demand elevation from the armchair. He thus thinks that if we view semantic theorists as attempting to answer question (a) in the manner he describes, the discrepancies between the account of what their project is, and what they actually do, will vanish; and he views the provision of an answer to (a) as an essential prerequisite for any attempted answer to (b):

"The sort of understanding of the actual capacities of speakers which is called for would be achieved exactly when enough was known about them to enable us to understand how in detail they embody such a device. And, of course, there can be no such understanding before we have formed the appropriate theoretical conception of the powers which the device must have. Doing that requires writing the computer program."²

However, it seems to me that Wright has overlooked something crucial here. Given that we are interested in answering the question "How do *they* do it?", even after (a) and (b) have been answered there remains the further question: (c) Do they, in reality, embody the device whose theoretical powers the computer program describes? An affirmative answer to this question is crucial if the answers given to (a) and (b) are to have any value whatsoever – an account of the theoretical powers of the device and an account of how it might be *possible* for speakers to embody it will have no explanatory value at all as an answer to the question "How do they do it?",

if they do not actually embody the device in question. (a) and (b) together tell us how it *could* be done, but what we are really after is an account of how it *is* done, and for this we need an affirmative answer to (c). And, of course, the provision of an answer to (c) is an empirical matter, dependent on the sort of empirical evidence concerning language acquisition and loss which we have repeatedly mentioned above : whether or not a speaker actually does embody a particular information-processing mechanism is a matter which is amenable to empirical investigation. We will not view a particular speaker as embodying a particular information processing mechanism unless the derivational structure of that mechanism is isomorphic to the causal explanatory structure the existence of which is suggested by the appropriate sorts of empirical data.

So I would say that if semantic theorists really are out to answer the question "how do they do it?" of our capacity to understand novel utterances, they are, at some point, going to have to pay attention to matters of empirical detail. The question of course is "Where?" Two broad answers to this question suggest themselves. The semantic theorist could begin to pay attention to such detail at the final stage of his enterprise : having written the appropriate computer programme, and having shown -- somehow or other -- that it is possible for it to be embodied in speakers, he could then *go on* to ask whether or not they actually do embody it, and it is at this last stage that the usual sources of empirical evidence will be crucial. The semantic theorist would thus be attempting to answer the questions (a), (b), and (c) in a simple linear fashion. But it is clear that this could be an extremely inefficient way of going about things : once answers to (a) and (b) have been provided there is no guarantee that the required affirmative answer to (c) will be forthcoming. The semantic theorist could write out his program, argue that it is possible for the information-processing device to be embodied in actual speakers, and then go on to discover -- to his horror -- that it is in fact *not* thus actually embodied. And this could happen time and time again. (Of course, he might get an affirmative answer to (c) on one of his early attempts after all, but it is clear that if he is proceeding in the linear fashion described, this could only have the status of a happy accident.)

The alternative, and much more efficient way of proceeding would be to write the computer program out as before, but this time in accordance with empirical constraints which ensure that, once (a) and (b) have been answered, no obstacle remains to the production of an affirmative answer to (c). In effect, this involves writing the computer program -- constructing the theory of meaning -- in accordance with some constraint along the lines of the mirror constraint. The semantic theorist is thus still viewed as attempting to answer the questions (a), (b), and (c), but no

longer in the simple linear manner described above. We answer (a) and (b) subject to the constraint that there is an affirmative answer to (c) via the imposition of something like the mirror constraint, and we are justified in so doing because without an affirmative answer to (c) the answers to (a) and (b), no matter how elaborate, will have no explanatory value whatsoever as answers to the question "How do they do it?". There is no point in writing a program describing the theoretical powers of an information processing device which we do not embody if the whole aim of the enterprise is to achieve an understanding of how we can understand novel utterances.

My conclusion is thus that far from providing the bones of an account of the semantic theorist's task which does not involve the imposition of the mirror constraint, Wright's suggestions in the end only serve to highlight the need for the imposition of that very constraint.³

CONCLUSION

In this essay I have attempted to sketch a defence of the claim that a semantic theory for a natural language should in some way reflect the semantic structure of that language. I took as my starting point (1.1 and 1.2) the Wittgensteinian idea that understanding a sentence is akin to the possession of a type of ability ; I then suggested (1.3) that what distinguishes the exercises of an ability from the manifestations of a disposition, is that the former are intentional : sentential understanding was thus to be construed as an intentional state, whose content is specified by the appropriate meaning-specifying theorem of the semantic theory. This account was then defended (1.4) against objections made by Appiah, Baker and Hacker, and McDowell ; and (2.3) against the claim that it violates a very plausible constraint on the ascription of propositional attitudes. In chapter 2 I asked (2.1 and 2.2) whether a similar account could be given of the nature of the relationship between speakers and the semantic theory's axiomatic base, and came to the conclusion that although it could not, we could nevertheless view states of tacit knowledge of semantic axioms as "subdoxastic states" – cognitive states which are not genuine intentional states, but which nevertheless have some kind of informational content. In chapter 3 I attempted to find out whether the account sketched in the first two chapters could provide the resources for a response to Wright's demand (3.1) for a justification of the imposition of the compositionality constraint, and suggested that this could be achieved via the imposition of Davies' mirror constraint, which we saw (3.1 and 3.2) to have its roots in suggestions made by Gareth Evans. The mirror constraint was then shown (3.2 and 3.3) to contain the resources for a rebuttal of three objections which Crispin Wright has raised against Evans ; to be suggested (3.4) by a plausible thesis concerning the nature of semantic facts ; and defended (3.5) against the claims that its imposition is unnecessary for the completion of the semantic theorist's task, and that there was no guarantee that it would serve to pick out a unitary semantic structure. It was then suggested (3.6) that the imposition of the mirror constraint could help explain speakers' capacities to understand novel utterances and the fact that natural languages are typically learnable, and a proof was provided to the effect that a theory of meaning for a learnable language satisfying the mirror constraint would have only finitely many axioms. Finally, I looked briefly (3.7) at the relationships between the mirror constraint, the

structural constraint, and Wright's conception of the semantic theorist as engaged in writing a computer program.

However, even in the unlikely event of everything I have said above turning out to be correct, there would still be at least one major problem outstanding which anyone wishing to embrace my account would have to face up to. I cannot do more than mention this problem here, as a proper discussion of it and its implications would require another essay at least as long as the present one. It stems from the fact that our account attempts to give an explanation of language mastery – in particular of the capacity to understand novel utterances – in cognitive–psychological terms. In Wright's words, it is an ability we are conceived to have

"...because we are appropriately related to a finite body of *information* which may be inferentially manipulated in such a way as to entail, for each novel string on which we can exercise our 'linguistic-creative' power, appropriate theorems concerning its grammaticalness and content",

which commits us to

"...the picture of language as a kind of syntactico-semantic mechanism, our largely unconscious knowledge of which enables us to compute the content which, independently and in advance of any response of ours, it bestows on each ingredient sentence... [and in which] the mechanism does the generating and the competent adult keeps track of what (and how) it generates."¹

Prima facie, this seems like an accurate description of our account : the information codified in the causal states corresponding to semantic axioms is conceived of as settling in advance, and independently of anything we go on to say or do, the content of the totality of admissible sentences in the language.

But why should this be a problem ? What is wrong with the picture of language as a syntactico-semantic mechanism, whose output competent speakers are able to track ? Wright's suggestion is that it is precisely this sort of picture of the ability constitutive of language mastery which is the ultimate target of Wittgenstein's remarks in the *Investigations* and elsewhere on the nature of rule following : that if we try to construe language mastery as an ability to track states of affairs constituted independently and in advance of what we say and do, we shall find ourselves unable to give any coherent account of the epistemology of the tracking accomplishment. Outlining Wittgenstein's arguments to this effect, let alone attempting to evaluate them, is outwith the scope of this essay ; accordingly, I shall end by recording my opinion that there is an issue to be resolved here, and that (at least) the following two questions require to be answered before we can rest

content with our account of tacit knowledge of theories of meaning :

(1) Does the account of tacit knowledge which uses the mirror constraint actually commit us, as Wright would suggest, to conceptions of meaning and linguistic understanding which the rule-following considerations would counsel us against ?²

(2) If the answer to (a) is affirmative, are the arguments which stem from the rule-following considerations actually successful in showing that the conception of linguistic understanding as an ability to track independently constituted states of affairs, is mistaken ? —

NOTES

Introduction

1. We should distinguish between two senses of the phrase "Theory of meaning". This could be taken to signify, on the one hand, a theory relating to a single language which attempts to state the meaning of every sentence in that language ; or, on the other, a theory relating to language in general, which "attempts to analyze, elucidate, or determine the empirical content of, the concept of meaning in general" (Sainsbury (1), p.127). In this essay we will be concerned exclusively with theories of meaning in the former sense.
2. See e.g. Platts (1), p.10, and Wright (2), pages 237 and 238.
3. Wittgenstein (1), §112.

Chapter 1

1.1

1. I use the theoretical / practical knowledge distinction in setting the scene here despite Dummett's recent claim that it is precisely when we come to look at sentential understanding that the distinction begins to break down (see Dummett (2), p.262). The reason Dummett gives for this claim strikes me as unsatisfactory : he says that in the case of an "ordinary" practical capacity, such as swimming, I can try to swim even if I have not learned what swimming is, whereas with language "there is no such thing as trying to speak a language that one does not know", and "it is only by learning to speak and understand a language that one comes to know what it is to speak it". These seem trivially false : I can try to speak German even though I don't know it, and I can know what it is to speak German – in the sense that I can quite easily identify speakers of German as such – without myself being able to speak and understand its sentences. Of course Dummett might be speaking of language *as such* here, as opposed to any particular language, but even when the claim is interpreted in this sense the first of Dummett's two claims above is still clearly false – witness the infant's first attempts to mouth sentences of his native tongue ; and the second claim, while true – a one year old child doesn't yet know what it is to speak a language – doesn't seem to say anything about language in particular : such a child doesn't yet know what it is to swim either.
2. The distinction between the two sorts of knowledge is drawn in Chap.2 of Ryle (1). For expositions of Wittgenstein's arguments here, see McGinn (1), Budd (1), and Baker and Hacker (2). The essence of Wittgenstein's argument is that meaning construed as mental baggage sets no standards for the *correct use* of an expression – thus construed it fails to be normative, and normativity is of the essence of meaning. Unlike McGinn, I do not want to claim that the *knowledge how* / capacity construal of understanding *by itself* can furnish us with a satisfactory solution to the sceptical argument adumbrated in Kripke (1) : see Wright (7) for the reasons why.
3. Quoted in Appiah (1) p.23.
4. Wright (1) p.53. For a similar view, see also Fricker (1), p.56.

1.2

1. Dummett (1) p.217.
2. I can see no reason why these two conditions should not come apart – I might very well be able to provide paraphrases without being in a position to give a statement of the underlying rules, thus leaving room for two notions of explicit knowledge – but I let this pass for the moment.
3. Dummett (1) p.224.
4. I am thus essentially in agreement with the points made in Appiah (1) pp.4 - 9, in defence of the notion of unconscious belief – I deny only that it is in virtue of possession of unconscious beliefs in the sense distinguished by Appiah that we can be said to possess sentential understanding. See section 1.4.
5. The use of the intentionalistic sounding "implicitly knows that" may seem strange given that all we are doing in ascribing it is ascribing an ability, but this strangeness should hopefully evaporate when we reach section 1.3
6. I think Dummett gets himself unnecessarily into trouble because he attempts to *combine*

these two philosophically innocuous senses of implicit *knowledge that*, in that he attempts to construe a speaker's possession of a practical ability (and hence of implicit *knowledge that* in the latter sense) as being *in virtue of* his possession of implicit knowledge in the former sense. A speaker would thus be conceived as unconsciously knowing a proposition (one which stated the rule for the linguistic expressions' correct use), and being able to use that expression in the appropriate ways *because* of the knowledge he has of this proposition, but which he is unable to bring to consciousness or explicitly state. His possession of the implicit *knowledge that* would in this case be viewed as potentially providing an *explanation* of his possession of the ability. (See e.g. Dummett (3), *passim*.) Dummett is here in grave danger of violating the insights provided by the Wittgensteinian arguments which prompted the construal of sentential understanding as a recognitional ability in the first place: on such a conception it would still be possible for a speaker to be apprised of that propositional knowledge and yet consistently fail to use it in the correct ways, and so we would be no nearer an account of in what his understanding consists – construing the propositional knowledge in question as unavailable to consciousness obscures rather than solves the difficulty here.

7. Dummett (1), p.224.

8. See for instance McDowell (1), Davidson (1), and Foster (1). Also Platts (2), p.231.

1.3

1. In Quine (1).

2. Wright (1), p.53.

3. Of course, the problem now is this: how can I ascribe to a speaker an intention to uphold a regularity when an explication of what is involved in that regularity involves concepts which the speaker quite manifestly does not possess? In the present case our speaker typically has no notion of what T-theorems are, nor of satisfaction by an infinite sequence and the rest of the Tarskian paraphernalia – so how can we say that he intends to uphold regularities which require use of these concepts and ideas for their proper description? A satisfactory answer to this question is absolutely necessary before we can be content with our notion of implicit knowledge – see 2.3 of the present essay for a discussion of this fundamental problem.

1.4

1. Appiah (1) p 5 - 6.

2. Ibid. p.8.

3. Baker and Hacker (1) p.359.

4. To the extent that Dummett writes as if explicit propositional knowledge can at least sometimes be constitutive of linguistic understanding Baker and Hacker's characterisation of his position is perhaps not as far off track as I suggest in the text – but I think Dummett's writing in this way is probably due to carelessness on his part, and I think that the account that I gave in section 1.2 captures what Dummett would have said had he been consistent. It is clear that as far as *my* "Dummettian" account is concerned, the characterisation which Baker and Hacker give *is* way off track.

5. Baker and Hacker *ibid.* p.367.

6. In the same paper (McDowell (2)) he raises another objection: that it will be radically indeterminate from the speakers' behaviour which rules they are actually implicitly following; because the pool of their total behaviour is finite we can always devise "bent rule" hypotheses the following of which would equally well account for that behaviour. This is in essence the argument employed by the meaning-sceptic in Kripke (1). I do not discuss this here because I am at least partially satisfied that those sceptical arguments have been dealt with satisfactorily elsewhere – for instance, see Wright (4), (5), and (6).

7. McDowell *ibid.* p.66. The explanation of Dummett's which McDowell alludes to here is quoted on p.62 *ibid.*

8. Wittgenstein (1), § 201.

9. Given the frequency with which this result has appeared in recent discussions of the rule-following considerations one would have expected an attempt at a principled rejection of Wittgenstein's insight, or at least a reference to it, from McDowell; but it is clear from, e.g., McDowell (3), that he actually embraces that insight – McDowell in his attack on Dummett seems then to make the very same mistake that in his (3) he accuses Kripke of making. Of course, the interesting and most pressing question is: how can this be a way of following a

rule at all? I cannot attempt to address this in the present essay.

10. See Wright (2), pp. 219-220.

11. This was made in a lecture on implicit knowledge, St. Andrews, 16th Dec. 1987.

12. see section 3.7 of the present essay.

13. In any case, it is difficult to see how Wright's distinction could actually save him from his own objection here: if a rule is genuinely normative – capable of settling disputes – shouldn't a statement of that rule equally well furnish us with a means of settling them?

Chapter 2 :

2.1

1. This appears in Wright (2), section III and Evans (1), section § III
2. One wonders in the case of the rat whether the failure to satisfy the constraint arises trivially from the fact that the rat has no beliefs or projects as such *at all*. A more interesting example would have been a case of an intentional agent's rigid disposition to behave in a certain way – this would leave scope for the possibility of holistic interaction which as far as I can see is ruled out in the case of the rat because there is nothing for the disposition to interact *with*. In any case, the example just about to be considered provides us with just such a case.
3. Evans *op.cit.* p.133.
4. Wright (2) p.227 and p.228 respectively.
5. See Stich (1).
6. Wright (2) p.227 and pp.237-238 respectively. Wright's reason for wanting to preserve an intentionalistic construal of the relationship between the speakers and the *theorems* is that such a construal seems necessary if the anti-realist attack on the realist's account of understanding is to be couched in the usual Dummettian way – as placing "impossible demands on the range of what speakers may reasonably be regarded as (implicitly) knowing" (Wright *ibid.* p.209). Our move to talk of intention is conservative in this respect – the standard anti-realist challenges concerning e.g. acquisition and manifestation can be stated as being about ascriptions of intention.
7. It might be thought that this move is unnecessary on the grounds that we can do the same work more simply, by adding the following to our initial constraint :

...and if P can itself come about as the result of interaction between many other propositional attitudes of the agent W.

This captures Stich's claim that it is another feature of genuine beliefs that the routes *to* them are, as he puts it, inferentially promiscuous. The claim would then be that states of tacit knowledge of axioms could not come about as the result of interaction between other genuine intentional states of the agent. I find it difficult to come to any opinions concerning the merits or otherwise of this claim, hence the strategy adopted in the text. In any case it's no easier to see how tacit knowledge of *theorems* could come about in this way.

8. Stich (1) pp.508-509.

9. *Ibid.* p.507.

10. This is Stich's own way of describing what he's doing ((1), p.500), and coming from him it sounds a little ironic, to say the least, given the line taken in Stich (2).

11. Hess took two photographs of the same girl and enlarged the pupils of her eyes in only one of them and then showed the two pictures to a range of male subjects who were unaware of the change; Hess found that the males consistently described the girl as looking more attractive in the altered picture, although they were unable to say in what the difference consisted. The idea is that there must have been causal states of the males containing information about relative pupil size which played a part in the causal production of the explicit beliefs about relative attractiveness, although those causal states are not to be counted as genuine beliefs themselves. For a fuller discussion of the Hess experiment, see Stich *op.cit.* p.503, pp.505-506, and p.511.

12. Martin Davies in his (2) comes close to making the same sort of point when he says that "it is an impressive feature of propositional attitude explanation that no particular pattern of behaviour is guaranteed by the presence of any particular belief" (p.84) and "the explanatory role of the implicit beliefs is exhausted by their yielding conscious beliefs" (pages 85 and 86). But his account of the matter is still somewhat confused since he (a) makes no explicit

reference to any notion similar to that of my "directness", (b) does not distinguish clearly between the two sorts of explanation, and (c) does not seem to realise that the crucial intentional notion at issue is intention, as opposed to belief.

It may seem that I have mistakenly overestimated the role of sentential understanding in rationalistic explanation of behaviour. The received wisdom in the philosophy of mind currently has it that "an action is intentional only if it is done with reasons in the light of which the agent's behaviour can be seen to be comprehensible", where "reasons for acting are appropriate combinations of desires and beliefs" (See Smith and Jones (1), p.132). Given this, and taking the stated condition as sufficient as well as necessary, it seems that adverting to beliefs and desires not constitutive of sentential understanding will be enough to secure the correctness of a description of a particular linguistic act as intentional – e.g. why did she say "Pass me the salt" ? – because she had a desire for some salt on her food, together with a belief that you would pass it if asked etc. No mention will generally be made in such an explanation to the intentions the speaker has concerning the future use of the linguistic expressions in question – does this not suggest that understanding cannot be an intentional state distinct from the non-linguistic propositional attitudes the speaker possesses, on the grounds that rationalistic explanation of behaviour can proceed without reference to it ?

This objection is far from conclusive, for the following two reasons. Firstly, we can see that the intentional state allegedly constitutive of sentential understanding is necessary for rationalistic explanation, by explicitly supposing it absent and examining the consequences : given that a speaker does not understand a sentence S we cannot even begin to give an explanation, in terms of his beliefs and desires, of why that speaker uttered S on a particular occasion. Secondly, the reason that sentential understanding is not usually referred to in rationalistic explanation is, I suggest, that it is usually assumed as part of the background against which explanation (of linguistic acts) in terms of beliefs and desires can take place. So if we were giving a *full* rationalistic explanation of a speaker's utterance, we would have to make some reference to the intentional state constitutive of his understanding of that sentence. This ties in with my diagnosis of Appiah's mistake in Chap. 1 : it's not that rationalistic explanation and giving an account of what understanding consists in are *separate* enterprises – rather, it's that the latter has to be assumed before the former can proceed with reference to linguistic acts. Appiah's mistake was to take the absence of any mention of intentions constitutive of understanding as signalling a genuine absence, as opposed to a genuinely held background assumption. Thus, whereas in Stich (2), Stephen Stich answers the question "What does every speaker of a natural language know ?", and gives the answer "Nothing interesting", I would ask the question "What does every speaker of a natural language intend ?", and give the answer "Quite a lot !". And this answer is justified because of the role the intentions in question play in the rationalistic explanation of speakers' linguistic behaviour.

2.2

1. See Davies (1).
2. Evans (2) pp.100-105.
3. This is stated eloquently in Davidson (2), p.200 :

"Beliefs are identified and described only within a dense pattern of beliefs. I can believe a cloud is passing before the sun, but only because I believe there is a sun, that clouds are made of water vapour, that water can exist in liquid or gaseous form ; and so on, without end. No particular list of further beliefs is required to give substance to my belief that a cloud is passing before the sun ; but some appropriate set of related beliefs must be there."

2.3

1. See Wright (2), pp. 217-214.
2. Wright makes essentially the same point at *ibid.* p.238.
3. Anscombe (1), §19.
4. Initially, that in which the answer, if positive, gives a reason for acting. Anscombe goes on to qualify this, and to discuss it in detail ; but these are matters I cannot follow up here. See Anscombe *ibid.*, §5, 6.

5. Wright (8), p.22. Wright's discussion proceeds on the assumption that 1st person avowals *do* possess truth-values, a thesis which some philosophers have called into question (see Hacker (1), chapter 9). I will not attempt to discuss this here, and will accept the assumption that they do in fact have truth-values.
6. See Wright (6), section IV.
7. This is even more apparent in Wright's earlier presentation of the account, where we are invited to reflect on the following biconditional (Wright (5), p.401) :

X intends that P if and only if X is disposed to avow the intention that P, and would be sincere in so doing, and fully grasps the content of that intention, and is prey to no material self-deception, and ... and so on.

The satisfaction of the conditions on the RHS is again viewed as being constitutive of the state of affairs described by the LIIS – but it's clear that if Wright wants to avoid prejudging the issue on tacit knowledge he is going to have to drop the biconditional, and settle for the weaker conditional leading from the RHS to the LHS.

8. In the former case would the modification in question be a merely *ad hoc* way of saving the agnostic stance on the implicit knowledge issue ? I think not, for there are other reasons why the particular C-condition under attack should be modified, independent of the details of the tacit knowledge debate. These concern the satisfaction of condition (3). In the case of avowals this amounts to the condition that the satisfaction of the C-conditions has to be, in any particular instance, logically independent of the details of the extension of psychological predicates involving the notion of intention. But the C-condition concerning possession of concepts seems to violate this : Is the truth of the claim that Jones has grasped a concept 'X' logically independent of considerations pertaining to the matter of what his intentions are as regards the future use of sentences involving that concept ? As I argued in 1.3 we have to assume that he possesses certain such intentions, if we are to assume that he has indeed a grasp of the appropriate concepts – in particular an intention to uphold the regularity prescribed by the truth (or assertibility) theorem corresponding to the sentences which feature that concept in a correct semantic theory for his language. Ascription of such intentions is necessary if we are to be able to distinguish between genuine linguistic *abilities*, and mere *dispositions* to utter sequences of noises in response to given sensory stimulations ; without intention, the former simply collapses into the latter – a collapse which I suspect would not be welcomed by Wright. So I think that this provides some independent motivation for the alteration to the c-conditions which Wright requires if he is to avoid simply prejudging the issue on tacit knowledge. (See Miller (1).)
9. Wright rests his case for this claim on the grounds that we find ourselves unable to give an account of the content of the concept which the dog might be alleged to possess.
10. Wright (2), p.222.
11. Wright himself explicitly acknowledges this view of concepts in the passage quoted on page 3 of the present essay.
12. Geach (1), p.13.
13. Or rather, there is no deep problem here which does not also present itself when we consider explicit knowledge, explicitly held intentions, or intentional states of whose content the possessor is, in the requisite sense, "conscious" – the problems that surround the notion of following a rule do not disappear if we suppose the putative rule-followers to be "explicitly aware" of the rules.

The account of sentential understanding which underlies this chapter, and which will figure again in the next chapter, could be called "functionalist". To the extent that functionalist theories are perhaps the most widely held such theories in today's philosophy of mind, I can claim this as an advantage of my account. But such a label might seem strange for an account which claims to have its roots in Wittgenstein's reflections on these matters – doesn't functionalism, after all, characterise e.g. an intention as a mental state, something which Wittgenstein himself appeared to wish to repudiate ? (See e.g. the footnote to p.59 of the *Investigations*). But it is not clear to me that there need be any genuine conflict between Wittgenstein's views and the functionalist conception here :when we remember that the mental states in question are individuated by their causal (functional) role in the production of behaviour – in the production of *use* – the functionalist story begins to have a distinctive Wittgensteinian ring to it. And the main problem for functionalism – at least for functionalist accounts of the propositional attitudes, as opposed to such accounts of *occurrent* mental phenomena – seems to be its inability to account for the intuitive 1st person epistemology

associated with such states : but isn't a very similar problem noted by Wittgenstein when he sees that there is a tension between his dictum that "meaning is use" and the idea that the meaning of an expression can be grasped "in a flash" ? (See Wittgenstein (1), §138). Perhaps a Wittgensteinian solution (if there is one) to the latter problem might provide the resources for a satisfactory solution to the functionalist problem concerning 1st person epistemology ; unfortunately I cannot attempt to follow up this suggestion here.

Chapter 3 :

3.1

1. See Wright (2) pp.205-206, and the references which he provides there to Davidson's *Inquiries* op.cit.
2. See Evans (1), pp.122-123.
3. For a clear statement of the challenge, see Wright (10). This seems to me to be somewhat weaker than the challenge of Quine's we mentioned in the first chapter – the suggestion of which was that even where there was no empirically respectable evidence which would discriminate between two sets of "rules" there is still theoretical room for the distinction between those rules "guiding" and "fitting" the behaviour of the creatures concerned. However, it may be that I have put a stronger interpretation on the Quinean challenge than the one Quine wished to give himself, which may well coincide with the challenge laid down by Wright to the semantic theorist..
4. Evans (1), p.124.
5. Evans *ibid.*, pp.124-125.
6. *ibid.* p.125.
7. *ibid.* p.127.

3.2

1. For Wright's criticisms see Wright (2) (section III) and Wright (9) (section II). For Davies discussion of Wright see Davies (3).
2. Wright (1), p.233.
3. *ibid.*
4. Evans *op.cit.* p.125.
5. I would frame the constraint slightly differently from Davies, replacing his talk of beliefs about meaning with talk of intentions constitutive of understanding. For ease of reference I have stuck to Davies way of talking in the text, but it should be remembered that I take what he refers to as semantic beliefs to be intentions.
6. What follows clearly owes a lot to the discussion in Fricker (1), pp.50-52.
7. See Lewis (1), pp.113-114.
8. Davies (3) pp.443-444. We should perhaps note that Wright himself describes this objection as the "most serious" objection to the mirror constraint. See Wright (2), p.213.
9. *ibid.* p.451.
10. *ibid.* p.452.
11. *ibid.* p.453.
12. *ibid.* p.454.

3.3

1. *ibid.* p.444.
2. *ibid.* p.445.
3. *ibid.* pp.446-447.
4. *ibid.* p.459.
5. Davies (4), p.149.
6. Davies (3) p.459.
7. *ibid.* pp.447-448.
8. *ibid.* p.459.
9. *ibid.* p.460.
10. So if we want to use the notion of derivational equivalence to pick out the class of semantic

theories that are all tacitly known by a given speaker we shall have to amend our definition to :

Two (extensionally equivalent) theories T_k and T_m are said to be DS-equivalent whenever the following holds (for any sentences S_1, \dots, S_n, S) – the semantic resources in T_k which suffice for the canonical derivation of truth conditions specifications for S_1, \dots, S_n suffice also for the canonical derivation of a truth condition specification for S *if and only if* the semantic resources in T_m which suffice for the canonical derivation of truth conditions specifications for S_1, \dots, S_n suffice also for the canonical derivation of a truth condition specification for S – *and* the number of proper axioms of T_k is equal to the number of proper axioms of T_m .

3.4

1. Davies (2), p.54.
2. The objection is that our account leaves open the unattractive possibility that a language could have a multitude of different semantic structures.
3. See Fricker (1), section I.
4. See Peacocke (1).
5. Fricker op.cit. p.51.

3.5

1. See Fricker op.cit. pp.55-56, and Dummett (4) pp.69-70.
2. See Sainsbury (1), section 1 on "Understanding and structure".
3. *ibid.* p.131.
4. Fricker op.cit. p.56.
5. So, if the argument to LIST is correct, Wright's argument (Wright (2), p.212) that "...a full description of the competence possessed by speakers should not merely characterize its ingredients but ought also to reflect their causal interrelations", is off track to the extent that it erroneously assumes that the causal interrelations in question play an indispensable role in the characterisation of the ability.
6. Fricker op.cit. p.59. My interpretation of Fricker's argument here follows closely the one to be found on pp.149-152 of Davies (1).
7. This is in effect the argument which is used in Evans (3) to rule out Quine's suggested alternative interpretations of "gavagai".
8. Davies (1), p.151.
9. Fricker op.cit. p.52. The underlining is mine.
10. See McGinn (1), pp.149-50, and Wright (7) for some able criticisms.
11. We could say that the argument for the weak ELS doctrine shows that the argument to the list result *must* be flawed, while the argument for the strong ELS doctrine which I am about to give shows *where* the argument to LIST is flawed.
12. This argument is hinted at, but not explicitly developed by, Fricker at her op.cit., p.64. It means, in effect, that Davies definition of a language as an ordered pair consisting of a set of sentences and a set of meaning specifications will have to be enriched to include some mention of exposure conditionals – or better still, of semantic structure.
13. Of course, the analogy is imperfect in the sense that it leaves open the possibility that if the two apprentices have the same "inner city" and the same architectural plan, they will be deemed to be in the same city, even though the cities they are engaged in building are spatially distinct. But all analogies have to be imperfect in *some* respect, and I don't think this imperfection detracts at all from the main point I am trying to establish.
14. This should dispose of the example which Davies gives in an attempt to prove that there could be a language with semantic structure none of whose complexity is reflected in the causal structure of its speakers' competences (Davies (1), pp.151-152) : its speakers are born endowed with an innate correlation between form and meaning which is "wired in" not word by word, but sentence by sentence. This example does not prove that there could be a *natural* language with semantic structure outrunning the causal structure infusing speakers competences, because the language it speaks of is not a natural language ! I said in the introduction to this essay that I mean by a natural language one which is learned by training and

- projection from that training – any language "learned" by "magic" is not a natural language. This might seem like moving the goalposts, and in a sense I am : back to where they should have been at the start of the game. We are simply not in the business of providing theories of meaning for magical languages. (These remarks apply also to Sainsbury op. cit., pp.131-132).
15. What, then, of Sainsbury's chicken sexers? I think I might be willing to concede that the case described is of a single ability realized in different causal structures. But imagine a tribe where just about everything its members do involves the exercise of the chicken sexing ability, and who thus exercise that ability in just about every hour of their waking lives : I think structure would again begin to play a crucial role in the characterisation of the overall ability, just as in the linguistic case.

3.6

1. This claim differs from the one made by Evans on behalf of his theory when he says that "the notion of tacit knowledge of a structure-reflecting theory of meaning, explained as I have explained it, cannot be used to explain the capacity to understand new sentences" (Evans (1), p.134). Evans' argument to this effect is opaque – see section IV of Evans (1), *passim* – and in any case I am willing to take the remarks I make in the text as showing that the same point cannot be made concerning the account based on the mirror constraint.
2. Davies (2), p.60.
3. This is in fact a rerun – with a few minor gaps filled in – of the proof given by Davies (see Davies (2), pp.57-60), which purports to show that if a language is learnable then a theory of meaning for that language which satisfies the structural constraint (see next section) can have only finitely many proper axioms.
4. Using the plausible principle that a semantic theory should deal with an expression using the most economical means possible.

3.7

1. See Davies (2), pp.54-57. This constraint is also discussed in Wright (2), pp.214-216.
2. Wright (2), p.236.
3. So although I do not want to go so far as to claim that "the right account has not greatly impinged upon the consciousness of workers in the field", I would suggest that the most efficient way of implementing that account seems to have alluded many of its practitioners.

Conclusion

1. See Wright (6), section I, for a very clear exposition of the problems here.
2. It seems to me that there are perhaps ways of attempting to keep the mirror constraint in play, and yet avoid commitment to the views of meaning and understanding which the rule-following conditions suggest to be objectionable. One way would be to simply drop all talk of causal explanation and opt for a purely descriptive version of the constraint, along the lines of

If, and only if, a speaker is able to come to understand sentence S without any further exposure or training after having been taught what each of the sentences S_1, \dots, S_n mean, and if any revision in his belief as to what S means occasions a corresponding revision in his belief as to the meanings of S_1, \dots, S_n , should the semantic resources sufficient for the canonical derivation of a meaning specification for S be sufficient also for the canonical derivation of meaning specifications for S_1, \dots, S_n

But I suspect that for most philosophers, the price paid here will be too high : we want to say that linguistic behaviour, like the other sorts of behaviour exhibited by an agent, is caused, and that if there is causation then there must be some states which do the causing. So perhaps a more attractive suggestion would be that we keep Davies' version of the mirror constraint as it stands, but drop the claim that the causal-explanatory states corresponding to semantic axioms have informational content – that they represent information – which somehow settles, in

advance, the content of the intentional states constitutive of sentential understanding. It is difficult to assess this suggestion in the absence of an account of what the relationship between the causal states and the axioms now is, and of when causal states which are not genuinely intentional should be claimed to have informational content. In Campbell (1) (p.24), Evans is taken to be suggesting that the causal states underlying language mastery have no such content :

"[Evans suggests that] the discernment of structure by a description of understanding is in effect a discernment of non-psychoic, purely neural structure.... [and that] there is no *psychological* machinery which typically explains a speaker's perception of the meaning of the heard sentence",

while Davies is explicit in his desire to take the opposing view (Davies (1), p.140) :

"...we ought to explore the differences between propositional attitudes and other information-containing cognitive states" (my emphasis).

Deciding who is right will require a full account of the sort of content possessed by Stich's subdoxastic states together with an account of the conditions under which such content should be ascribed – this question should perhaps be the starting point for any future discussion of the notion of tacit knowledge : all I want to suggest here is that if Campbell's interpretation is the correct one then there might still be a way of holding on to the mirror constraint, while avoiding the conceptions of meaning and understanding which Wittgenstein would have us call into question.

BIBLIOGRAPHY

(1)

- Anscombe, G.E.M. *Intention* (Oxford 1957).
- Appiah, A. (1) *For Truth in Semantics* (Blackwell 1986)
- Baker G.P. and Hacker P.M.S. (1) *Language, Sense, and Nonsense* (Blackwell 1984).
- Baker G.P. and Hacker P.M.S. (2) *Meaning and Understanding : an analytical commentary on the Philosophical Investigations* (Blackwell, 1980).
- Budd, M. (1) "Wittgenstein on meaning, interpretation, and rules", in *Synthese* 58 (1984), pp.303-323.
- Campbell, J. (1) "Knowledge and understanding", *Philosophical Quarterly* (1982), pp.17-34.
- Davidson, D. (1) "Reply to Foster", in *Inquiries into Truth and Interpretation* (OUP 1984), pp.171-179.
- Davidson, D (2) "The method of truth in metaphysics", in the *Inquiries* op.cit., pp.199-214.
- Davies, M. (1) "Tacit knowledge and the structure of thought and language", in Travis (ed.) *Meaning and Interpretation* (Blackwell 1986), pp.127-158.
- Davies, M (2) *Meaning, Quantification, Necessity : themes in philosophical logic* (Routledge & Kegan Paul, 1981).
- Davies, M (3) "Tacit knowledge and semantic theory : can a 5% difference matter ?", in *Mind* 1987, pp.441-462.
- Davies, M (4) "Meaning, structure, and understanding", in *Synthese* 48, 1981, pp.135-161.
- Davies (5) "Meaning and structure", in *Philosophia* 13 (1983), pp.13-33.
- Dummett, M.A.E. (1) "The Philosophical Basis of Intuitionistic Logic", in *Truth and Other Enigmas* (Duckworth 1977), pp.215-247.
- Dummett, M.A.E. (2) "Reply to John McDowell", in Taylor (1), op.cit., pp.253-268.
- Dummett, M.A.E. (3) "What is a theory of meaning ?", in Guttenplan, S (ed.) *Mind and Language* (Oxford 1975), pp.97-138.
- Dummett, M.A.E. (4) "What is a theory of meaning ? (II)", in Evans and McDowell op.cit., pp.67-137.
- Evans, G. (1) "Semantic theory and tacit knowledge", in Leich and Holtzmann (eds.) *Wittgenstein : To Follow a Rule* (RKP1981), pp.118-137.
- Evans, G (2) *The Varieties of Reference* , (OUP 1982).
- Evans (3) "Identity and predication", *Journal of Philosophy* 72 (1975), pp.343-363.
- Foster, J. (1) "Meaning and Truth Theory", in Evans and McDowell (eds.) *Truth and Meaning* (Oxford 1976), pp.1-32.
- Fricker, E (1) "Semantic structure and speakers' understanding", in *Proceedings of the Aristotelian Society* 1981-82, pp.49-66.
- Geach, P (1) *Mental Acts* (RKP 1957).
- Hacker, P.M.S (1) *Insight and Illusion* (Oxford 1972).
- Kripke, S. (1) *Wittgenstein on Rules and Private Language* (Blackwell 1982).

- Lewis, D.K. (1) "Radical interpretation", in *Collected Papers* vol.1 (Oxford University Press 1983), pp. 108-121.
- Miller, A. (1) "An objection to Wright's treatment of intention", forthcoming in *Analysis*.
- McDowell, J. (1) "On the sense and reference of a proper name", in *Mind* 1977, pp.159-185.
- McDowell, J. (2) "In defence of Modesty", in Taylor (1) op.cit, pp.59-80.
- McDowell, J. (3) "Wittgenstein on following a rule", in *Synthese* 58 (1984), pp.325-363.
- McGinn, C. (1) *Wittgenstein on Meaning* (Blackwell 1984).
- Peacocke, C. (1) *Holistic Explanation* (Oxford, Clarendon Press, 1979).
- Platts, M. (1) Introduction to *Reference, Truth, and Reality* (RKP 1980).
- Platts, M. (2) *Ways of Meaning* (RKP 1979).
- Quine, W.V.O. (1) "Methodological Reflections on Current Linguistic Theory", in Harman, G (ed.) *On Noam Chomsky : Critical Essays* (Anchor Books, New York, 1974), pp. 104-117.
- Ryle, G (1) *The Concept of Mind* (Hutchinson 1949).
- Sainsbury, R.M. (1) "Understanding and theories of meaning", *Proceedings of the Aristotelian Society* 1979-80, pp.127-144.
- Smith, P and Jones. O.R. (1) *The Philosophy of Mind : an introduction* (Cambridge 1986).
- Stich, S (1) "Beliefs and Subdoxastic states", in *Philosophy of Science* 45 (1978), pp.499-518.
- Stich, S. (2) "What every speaker knows", *Philosophical Review* (1971), pp.476-496.
- Taylor, B.(ed.) (1) *Michael Dummett : Contributions to Philosophy* (Reidel, Dordrecht, 1987).
- Wittgenstein, L. (1) *Philosophical Investigations* (Blackwell 1953).
- Wright, C.J.G. (1) "Truth conditions and criteria" , in Wright (3), pp. 47-69.
- Wright, C.J.G. (2) "Theories of meaning and speakers' knowledge", in Wright (3), pp. 204-238.
- Wright, C.J.G. (3) *Realism, Meaning, and Truth* (Blackwell 1986).
- Wright, C.J.G. (4) "Kripke's Account of the Argument Against Private Language", in *The Journal of Philosophy*, 81 (1984), pp.759-778.
- Wright, C.J.G. (5) "On Making Up One's Mind : Wittgenstein on Intention", in *Logic, Philosophy of Science and Epistemology* , (Proceedings of the 11th international Wittgenstein Symposium at Kirchberg, Austria) eds. Weingartner & Schurz, (Holder-Pichler-Temsky, Vienna, 1987), pp.391-404.
- Wright, C.J.G. (6) "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics", forthcoming in George, A, (ed.) *Reflections on Chomsky* (Blackwell 1989).
- Wright, C.J.G. (7) Review of McGinn (1), forthcoming in *Mind*.
- Wright, C.J.G. (8) "Moral values, projection, and secondary qualities", in *Proceedings of the Aristotelian Society* supp.vol LXII (1988), 1-26.
- Wright, C.J.G. (9) "How can the theory of meaning be a philosophical project ?", in *Mind and Language* Vol. I no.I 1988, pp.31-44.
- Wright, C.J.G. (10) "Rule-following, objectivity, and the theory of meaning", in Leich and Holtzmann op.cit., pp. 99-117.