

Rule-following, Compositionality and the Normativity of Meaning

Peter Pagin

The principle of compositionality says:

- (C) The meaning of a complex expression is determined by the meanings of its parts and its mode of composition.

Is (C) a valid principle for natural languages? One traditional reason for believing so is that it can explain our ability to understand new sentences, i.e. sentences which we have never encountered before. The explanation says that, for this ability to be possible, the meaning of complex expressions must be determined by rules, so that, by knowing the meaning of the parts of a new sentence, and knowing the rules by which they are combined, I can also know the meaning of the new sentence.

However, if Wittgenstein's so called *rule-following considerations* are correct, then this reason for believing in the validity of (C), is mistaken. The conclusion of those considerations is that we must reject the idea that rules are things which determine possible cases of application *before* those cases are actually encountered and decided by speakers. If this is right, then there is no rule which determines the meanings of new sentences, i.e. *before* those sentences have actually been used. Therefore, it might *seem* that (C) is not valid for natural languages.

My view is that these conclusions are mistaken. I think that if more attention is paid to how the principle of compositionality can properly be justified, the result will be that the rule following considerations are actually *irrelevant* to the validity of (C).

This conclusion goes against the common view of these matters in the literature, in works e.g. by Saul Kripke, Noam Chomsky and Crispin Wright.¹ However, it does so, in a sense, only on the surface. For I think that there is an underlying conflict between the view on meaning that

1. See section 4.

is associated with the rule-following considerations and that associated with the justification of compositionality. That conflict concerns the *normativity* of meaning. The assumption of normativity is a driving force behind the rule-following considerations, and explicitly rejected in the other view. Moreover, I think that the rule-following considerations provide a *reductio ad absurdum* of the normativity assumption itself.

The four sections of this paper are devoted to

- i) (my reconstruction of) the rule-following considerations
- ii) the justification of (C), and
- iii) its relation to the rule-following considerations
- iv) the issue of normativity, and its bearing on i), ii) and iii).

1. Rule-following

Wittgenstein's so called rule-following considerations comprise a number of reflections on the nature of rule-following and, partly, on the nature of the mental and of linguistic meaning. Usually included under this title are §§ 136-242 of the *Philosophical Investigations*², and a large number of paragraphs in the *Remarks on the Foundations of Mathematics*, especially in sections VI and VII (of the third edition). In the primary sense of the expression "the rule-following considerations" I shall intend this body of material, even if only roughly characterized.

The literature on the rule-following considerations has grown to a considerable size, especially since the mid- and late seventies, and a number of different interpretations of Wittgenstein have been proposed. This paper is not intended as a paper on Wittgenstein. It has rather limited exegetical ambitions. The presentation, or reconstruction, is not standard, but I don't think it is controversial either, except perhaps insofar as leaving things out which some regard as essential.³ I shall be concerned with conclusions about rule-following, which may and may not be Wittgenstein's, but which I consider both interesting and important and *prima facie* rele-

2. There is consensus about the end point of the rule-following segment (i.e. #242) but not about its starting point. My choice is #136, since this is where the distinction between "fitting" and "belonging to" is introduced.

vant to the status of the compositionality principle.

I shall present parts of Wittgenstein's discussion, the parts I consider directly relevant, as a *reductio* argument. The assumption from which absurdity is to be derived is this:

1. Assumption: A rule is *non-contingently* (internally) related to the correctness or incorrectness of every possible application, (whether or not considered by any rule follower).

On this assumption, a rule is thought of as something which *determines*, in advance of being actually applied to a particular case, what its correct application is in that case.⁴ Thus, the conclusion of the argument will be that rules do not have this predetermining property. Of course, this might be true for the reason that rules and rule-following do not exist, but with the extra premiss that they do, as Wittgenstein of course held, the conclusion is that a certain view of existing rule-following practice is erroneous. Wittgenstein illustrates this view with the picture of infinitely long rails: all you have to do as rule follower is to follow the rails. Where they take you is already determined.⁵

To avoid misunderstanding about assumption 1, it should be pointed out that it is not concerned with differences between rules, or with differences between applications of the same rule, or different rules. For instance, it is not specially concerned with unclear cases, i.e. it is not

3. My presentation is primarily concerned with the *negative* part of the rule-following considerations, not with the accompanying positive characterizations of rule-following. Moreover, it is directly concerned only with what may be called the metaphysics of rule-following, i.e. the question of *determination* of what is right or wrong in new cases, or possible cases, of rule application. Wittgenstein discusses this, and also discusses what may be called the *epistemology* of rule-following, i.e. questions about *knowledge*, or the possibility of knowledge, of what is right or wrong in new cases. There are negative results concerning the epistemology (thus, we don't really *know* what is right or wrong in a new case, but make "a stab in the dark", or even "decide" what is to be right), but these are at bottom consequences of the negative result concerning the metaphysics: it isn't determined what is right or wrong in the new case, and therefore it can't be known. Because of this I don't think anything essential is lost by leaving out the epistemological aspect.

an assumption that all cases are well determined. There may be a rule that at a certain time of the day you should perform ceremony C_1 if you are indoors, and ceremony C_2 if you are outdoors. But what if you are at a place where there are walls but no roof, or a roof but no walls? Unless there are special clauses for this, the rule doesn't say. Such things occur. Assumption 1 is not an assumption that they cannot occur, or do not occur. It is an assumption about all applications of rules to new cases, not just some that happen to be less clear than others.

Also, rejecting assumption 1 does not amount to rejecting the distinction between rule statements, or formulas, that do and those that don't determine a uniquely correct result. Wittgenstein himself gives as example the two expressions " $y = x^2$ " and " $y \approx x^2$ ": the first does determine a unique value for y , given a numerical argument, and the second doesn't.⁶ This distinction between determining and non-determining is not denied by rejecting the assumption. It is a distinction made *within* a rule-following practice. The assumption, on the other hand, is of a metaphysical nature. It concerns the *general* relation between rules and correctness of rule applications. It implies that the relation between a rule and its correct application in a particular case is *internal*, which is to say – at least – non-contingent. These are the important aspects: the rule is fixed at a certain time, and what is a correct application of the rule, in any case, depends only on the rule itself, as initially fixed.⁷

The next step in the argument is the stipulation that

4. In the case of mathematics it may be said that a rule determines all the correct applications, in the sense e.g. of determining the extension of a function, i.e. a set of ordered pairs or argument and value. Empirical concepts or rules cannot be thought of in this way. Rather, we must say something like that for any given object, a rule governing the term 'cat' provides the conditions under which the term applies to that object. And then it is an empirical matter, not a semantic one, whether the object meets the conditions.

5. *Philosophical Investigations*, # 218.

6. *Philosophical Investigations*, # 189.

7. For many rules, like traffic rules, there are also essential circumstances of application (as opposed to mere mathematical arguments), and the correctness of a rule application of course also depends on what the circumstances are. However, the rule and the circumstances are mutually independent. So given (independent) circumstances, the correct application is determined only by the rule.

2. Rules are expressed by our rule formulations.

That is to say that there is a rule expressed by each of our meaningful rule formulations, and that any rule is expressible by some rule formulation.⁸ In order to make the argument well articulated, this premiss needs to be entered. The whole argument concerns two relations of determination: the determinations *of* rules or meaning (by something or other) and the determination *by* rules or meanings of the correctness of applications:

BASE _L MEANING / RULE _R APPLICATION

This is the schema of determination of correctness of applications. There is a left determination relation, _L, between some base, which may be the rule formulation, perhaps together with the expression of explanations of how the formulation should be understood, and the meaning of the expression, or rule expressed by the rule expression. And there is a right determination relation, _R, between the rule, or concept, and the instances of application. The whole point of the discussion is this: by composing the left and right relations, there would be a determination of correctness of applications from the base, but no such determination is possible (the relation is contingent). So either the left or the right determination must be given up.

Wittgenstein is not explicit about this, nor does he take care to distinguish between rule and rule formulation. Rather, in line with his outlook, he just keeps these two together. Premiss 2 is questionable on the basis of doubt about the expressive capacity of language: there might be rules of the questionable kind, even though our formulations could not be assigned meaning well enough to express them. This doubt can be eliminated either by promoting the expressive capacity, or by demoting rules. Wittgenstein takes the latter option.⁹

Next comes a crucial point:

8. There is a further question of conditions under which different formulations express the same rule, but that does not enter into the present argument.

3. The relation between public rule expressions and new rule applications is contingent.

This is the thesis that is usually illustrated by familiar examples. Wittgenstein gave the example of a schoolboy who is to write down the sequence "+2", i.e. write down the number you get by adding 2 to the previous number, and continue in that way.¹⁰ All goes well up to "1000", but then he continues "1004", "1008", "1012". Our task is to explain to him why that is wrong.

As this example is normally developed, there is nothing we can do that can bring the schoolboy to regard his new applications as mistaken and ours as the right ones. We propose an explicit definition ($x+2 =_{\text{def}} (x+1)+1$) but he may insist that he is following that definition ($1000+1 = 1002$), and when we say he isn't doing the same thing as before he may reply that he is the one who is doing the same (he uses different criteria for what counts as "doing the same" in this case). And when we try to explain to him by means of examples of application, that will not work either, for no matter how many examples we give, when it is time for him to continue on his own, his course of action will sooner or later deviate from ours, and he will persist in it.

The main point is, or at least seems to be, that whatever expression we use, be it explicit verbal definitions, contextual definitions, series of examples, or whatever, it is always possible to

9. On this point I think that Saul Kripke, in Kripke (1982), does misinterpret Wittgenstein. For Kripke, the whole problem is a problem about the expressing relation. Thus, whereas Wittgenstein keeps BASE (rule formulation) and RULE together, thus concluding that rules don't determine new applications, Kripke keeps RULE and APPLICATION together, thus concluding that it remains indeterminate which rule is expressed. For the way Kripke sets up the problem is this: no facts determine *which* function or rule I intend. But this isn't an issue for Wittgenstein. His question concerns the nature of the rule intended.

Furthermore, I think this misinterpretation is encouraged, but not necessitated, by the fact that in the presentation of the problem Kripke considers the application of an expression at a *later time*, together with a question about what was really meant at an earlier time. The real issue is the application of an expression to a *new case*, a function to a new argument, not that the rule follower performs at a later time.

It is, clearly, a fact about us that we need a new time to consider a new case, for by definition no application I have already performed *now*, is new. But Wittgenstein's question was whether the rule expression, with its *present* meaning, predetermines new cases, not whether later applications of the same expression are applications of the same rule.

interpret them differently from what is intended. These expressions work as explanations of the rule only for someone who understands the expressions in the right way. But if our most basic ways of expressing ourselves are understood in a deviant way, we cannot use them for explaining less basic things. And there is no logical or necessary connection between the expression itself and the way it is supposed to be understood. Because the relation between public rule expressions and new applications is contingent, any proposed explanation can be understood in a deviant way.¹¹ When the schoolboy continues on his own, he *can* always use the explanation for arriving at deviant results. The results will still be correct, however, with respect to the *different* rule expressed by the same formulation, as *he* understands it.

Now, because of this, from 1, 2 and 3 together, we conclude

4. The *expressing relation*, holding between a public rule formulation/explanation and the rule it expresses, is contingent. Therefore a rule cannot be publicly explained.

If there were some means to convey to another person precisely how a particular expression of a rule should be used for applying the rule in new cases, then those means would be used, and we would not have the problem of deviant understanding. But every attempt at formulation or explanation is ambiguous, and can therefore be interpreted in several ways. So the most we can hope for is that we shall hit upon a way of expressing ourselves that *in fact* the schoolboy in his mind will associate with our preferred way of continuing the series. What that way is, is something we cannot make explicit. If it could be made explicit, then the schoolboy's deviation would be a cognitive *mistake*, a misapplication of his knowledge of the rule as explained by us.

10. *Philosophical Investigations*, # 185.

11. In this paper I speak both of rule explanations, which are instructions for proceeding in accordance with a rule, and also of theoretical explanations, in particular of the success of instances of linguistic communication. I hope that context will easily determine which of these two senses of 'explanation' is employed in each particular occurrence.

But no deviation need be a cognitive mistake.

The association of a rule formulation with a rule is something that each rule follower must effect by him/herself. Each rule follower is able to grasp a rule, and to select a formulation as expressing it, but cannot unambiguously convey to any other rule follower which rule the formulation expresses.

Therefore, we have

5. A rule formulation is associated with a rule by way of a (private) state of understanding.

The understanding of a rule cannot (ultimately¹²) be publicly articulated. Therefore each rule follower has an understanding that is private, because not conveyable to others. The state of understanding is then what will assume the position of BASE in the determination schema.

The next step in the argument is to attack this conclusion, i.e. 5. The reasoning, in reconstruction, goes like this.

First, it is observed that the understanding of a rule, as a private phenomenon, or mental state, must determine what the correct applications of the rule are. For if it would be left undetermined by that the mental state e.g. what the right result of applying "+2" to "1000" is, and there is a predetermined correct result, then knowing or understanding the rule cannot consist in being in that mental state, for the knowledge would include more than is determined by that state. So we have

6. The state of understanding a rule is non-contingently related to the correctness or incor-

12. This connects with #201 of the *Investigations*, where Wittgenstein speaks of putting interpretations behind interpretations. You can use expression B to explain how expression A is to be understood, but then the question arises how expression B is to be understood, and so on. It is not clear, however, that Wittgenstein in #201 speaks of interpretations of expressions used in interpersonal explanation. Presumably, he does not.

rectness of every possible application, (whether or not considered by any rule follower).

The state of understanding therefore predetermines all correct possible applications that haven't been performed. That is, there is a determination of APPLICATION from BASE (the state of understanding).

Next there is the question of how that can be *accomplished*. There are two possibilities. Either the mental state in some way *contains* all possible correct applications of the rule. Or else it does not contain them, but determines them in some indirect way.

The first possibility would obtain e.g. if the subject has already thought of every possible application. But this cannot be. If the rule has only finitely many applications, then it is in principle possible to have thought of them all, but this case is irrelevant, since we were asking for the possibility of applying the rule to new cases. This is what we take rule-following to consist in. And if the rule has infinitely many applications, then it is just not possible to have thought of each of them individually.

Therefore, the state of understanding must contain all possible applications in some other way. But what that would be, is wholly mysterious. We would, in some way, grasp each of the infinitely many applications in a moment, without having thought of them. But we don't have anything as a model for understanding such a phenomenon.¹³ This possibility must therefore be rejected as incomprehensible.

The alternative is indirect determination. There is then some identifiable state of understanding, such that the properties of that state in some way determine correct applications. But the state alone cannot, of course, accomplish this. In itself it is just an isolated state, just as a linguistic expression, without interpretation, is just a syntactic object. We cannot reasonably conclude *anything*, whether right or wrong, about meaning from a state of understanding, whether it is a mental image or mental process (understood as Wittgenstein understood mental processes)

13. Wittgenstein rhetorically asks for one in PI # 191.

or a disposition, of whatever kind. We also need some principle of *correlation*.

But the need of a principle of correlation is itself a problem, and indeed the same problem again. For there are more than one possible principle of correlation, and if meaning is to be determined by a state of understanding, then the state of understanding must also select the principle of correlation, and this brings us back to the first problem again.¹⁴

In the case of dispositions, as discussed by Kripke¹⁵, a principle of correlation is in fact provided. The principle is this: an expression *E*, as used by a speaker *S*, is true of an object *O* iff *S* is disposed to apply *E* to *O*. In this simple way the disposition to apply an expression is thought to determine the extension of the expression. Kripke's objection is that this makes the rule follower immune to error (whatever the person having those dispositions would do, as a result of having them, would count as correctly applying the rule), which in turn means that there is no rule he is following. However, other principles of correlation, which do not have this consequence, are possible. After all, there are indefinitely many ways of correlating dispositions with meanings. But then we have the question which principle to choose. And so the basic problem with the idea that meaning is determined by dispositions is not that we are disposed to make mistakes, or that it automatically would make us immune to mistakes, but that, as with other BASE candidates, it remains to be settled how dispositions should be *correlated* with meanings. And this again means that meanings cannot be fixed by dispositions.

From these considerations we can conclude that:

7. A state of understanding a rule cannot be non-contingently related to the correctness or incorrectness of every possible application.

14. This circumstance is considered by Wittgenstein in the example of the cube projection (#139-41). And again it connects with #201. For if a principle of correlation is to be selected by the state of understanding, another principle of correlation is needed for determining which one, and so on.

15. Wittgenstein on Rules and Private Language, pp 22-37.

But 7 contradicts the conjunction of 5 and 6. Therefore we do have a contradiction in the argument. The mistake is diagnosed to lie in the assumption, 1, i.e. in the statement that a rule is something which *determines*, in advance of being actually applied to a particular case, what its correct application is in that case. So, 1 is rejected (the right determination relation doesn't hold). That is the negative conclusion of the rule-following considerations. At least, for the purposes of the present discussion, I shall regard it as the conclusion.¹⁶

8. Contradiction. Hence the assumption, 1, is false.

Notice that the problems really are avoided by giving up 1. For if rules do not determine correctness of applications in advance, then two rule followers are in agreement about what the rule is as long as they are in agreement about all applications already made. They need not agree about future applications, for those are not determined by the rule. Once they come to a new application where they disagree, they will no longer be following the same rule, but that will only mean that they have *extended* one and the same rule, i.e. one and the same practice, in different ways. We are not allowed to conclude from this that there was a difference in understanding *before* the disagreement. Therefore, the claim that rules cannot be publicly explained, does not follow, and is indeed false: by giving a full description of how the rule is applied in all cases so far determined, I can provide a full public explanation of the rule, for there isn't more to know about it.

However, we could not, on Wittgenstein's view, I think, be said to share a rule-following *practice*, unless we also were agreeing, over and over, on new cases. In this sense, an explanation "works" if it can achieve the purpose of agreement in new cases, but strictly speaking it

16. As far as Wittgenstein exegesis is concerned, I would not say that Wittgenstein asserted, or would assert, the negation of 1. Rather, I think that the proper Wittgensteinian reaction would be to reject both assertions, in favour of the conclusion that both assertions employ the term "determination" in an unacceptable philosophical sense.

doesn't determine in advance what is correct in the new cases.¹⁷

This concludes the presentation of the argument.

2. The validity of the principle of compositionality

To repeat, the principle of compositionality says:

- (C) The meaning of a complex expression is determined by the meanings of its parts and its mode of composition.

In this formulation we again find the word "determine", that was of such a crucial importance in the rule-following considerations. If that word is used in the same sense in (C) as in the rule-following considerations, then, apparently these considerations have a direct bearing on the validity of the principle of compositionality. In short, if the rule-following conclusion is correct, then (C) is false (or possibly nonsensical). That is, this is the conclusion, provided we have to do with the same notions of determination and meaning.

Of course, the wording of (C) leaves it open, not only what notions of determination and meaning are involved, but also other things, such as the relevant notion of expression, part and mode of composition. Without further qualifications, all we have on determination is that the meaning of the whole is to be a *function* from the meanings of the parts and the mode of composition. We are not told whether this function need be anything else than a function in the set theoretic sense (a set of ordered pairs), whether it must be finitely expressible, or computable.

17. In fact, this seems to be a way to understand Wittgenstein's remark in #202 that rule-following is a practice. Going by a rule is acting in relation to a rule explanation. The explanation doesn't determine new cases, but its relation to new cases must still be non-arbitrary, since otherwise it isn't rule-following. Agreement between rule followers is then what non-arbitrariness can amount to.

As long as the left determination relation is retained, the practice of going by a rule can be regarded as a practice of *extending* the rule, or the concept, to cover new cases. This is in accord with Wittgenstein's view of what is done in mathematics when a new proof is accepted. However, this picture is not available to someone like Kripke, who rejects \mathcal{L} and keeps \mathcal{R} . So Kripke's picture of rule-following must look quite different, and so it does, as it is presented in his so called sceptical solution.

Neither are we told anything about the underlying reality, if there is one, in virtue of which the meaning of the whole is the value of such a function.

However, we can arrive at the relevant interpretation, or at least at a range of acceptable interpretations, by inquiring into the *justification* of the principle. If some particular interpretation of ‘determine’ renders the resulting principle unjustifiable, then that interpretation should be rejected. The only relevant interpretations are those allowing for a justification. I shall argue that if ‘determine’ is taken precisely in the sense of determination by rules, i.e. in a normative sense, then the principle cannot be justified. It can be justified only if understood descriptively, as relating to the ways people in fact speak and interpret each other. (C) would then be said to be true of a given natural language provided there is a compositional semantic description of the language that correctly enough describes and predicts how the speakers interpret the sentences.

Below I am going to consider the two, closely related, traditional arguments for the principle of compositionality, according to which it can be understood in the normative way. I shall argue that they fail. Then I shall present the justification that, in my view, works, the argument from communication. There are some other, related arguments, like Chomsky’s argument from productivity, or Fodor’s argument from systematicity, but my view is that either they are inadequate, or else they rely on the communication argument. So I shall not here devote space to them, since the central contrast can be made without.¹⁸

The argument from learnability

Donald Davidson writes:

When we can regard the meaning of each sentence as a function of a finite number of features of the sentence, we have an insight not only into what there is to be learned; we also understand how an infinite aptitude can be encompassed by finite accomplishments. For suppose that a language lacks this feature; then no matter how many sentences a would-be speaker learns to produce and understand, there will remain others whose meanings are

18. For Chomsky’s argument, see e.g. Chomsky (1980), pp 76-78. For Fodor’s argument, see Fodor (1987), pp 147-153.

not given by the rules already mastered. It is natural to say that such a language is *unlearnable*. This argument depends, of course, on a number of empirical assumptions: for example, that we do not at some point suddenly acquire an ability to intuit the meanings of sentences on no rule at all; that each new item of vocabulary, or new grammatical rule, takes some finite time to be learned; that man is mortal.¹⁹

In short, the argument is that unless a language is compositional, it would not be learnable (by a finite being). As Davidson himself says, this argument does build on some assumptions. One is that the language does have an infinite number of sentences, and another (unless this is included in the notion of a sentence) that infinitely many of those sentences do have a definite meaning.²⁰ Let's call such a language an *infinitely rich* language.

If we are allowed to assume that natural languages are infinitely rich, then we do have a good argument for compositionality. That such languages have a compositional structure is (part of) a good explanation of how a speaker can learn it. It is probably the best explanation, and possibly the only reasonable explanation, for no person, whether or not mortal, can learn all the meaningful sentences, one at a time, in any finite time (provided that the time needed for learning a new sentence does not rapidly approach zero). The problem is that we cannot just make the assumption. The claim that natural languages are infinitely rich is a strong claim about natural languages. It would be question begging to simply assume that it is correct. And it is not something we get directly from observations of natural language speakers. It needs a more theoretical justification.

Now clearly, if we could justify the assumption that speakers speak an infinitely rich language *without* compositional structure, then we would know in advance that the learnability

19. Davidson (1967) pp 8-9.

20. Actually, if infinity is required, we would have to add that infinitely many different meanings are expressed by the infinitely many meaningful sentences. It would not be difficult to learn a language with infinitely many meaningful sentences if they all mean the same, or if e.g. sentences with an even number of letters all mean that *p* and those with an odd number all mean that *q*. So we would have to require an infinite number of equivalence classes of synonymous sentences.

But perhaps infinity is not needed. A huge finite number of sentences to be learned might serve as well as a basis for the argument. This has been pointed e.g. by Richard Grandy.

argument is flawed, because then, if speakers learn such a language from each other we know that it cannot be correctly explained by means of compositionality, simply because we know that the language is not compositional. On the other hand, if we can justify directly the assumption that speakers do speak infinitely rich *compositional* languages, then we already have an argument for compositionality, and need not add any extra considerations about learnability.

The only reason for turning to learnability as an argument would be that we can justify the assumption that natural languages are infinitely rich, *without* being able, directly, i.e. from the same premisses, to justify the belief that they are compositional. But I cannot see at all how we could end up in that situation, for I don't think there could be *any* good reason for ascribing to a speaker the property of speaking an infinitely rich language *unless* on the basis of believing that a compositional theory gives the best account of the finitely many utterances he has actually made.²¹

If this is correct, then the learnability argument fails, in the sense that we would *already* have good reasons for compositionality, if we had good reasons for the *premisses* of the learnability argument. We would not then need any further considerations about learnability.

The argument from new sentences

Another traditional argument is concerned with our capacity to understand new sentences. It is said that a language must be compositional in order that it be possible to understand new sentences, sentences you have never used or encountered before. This argument, however, can be understood in two different ways.

One way to understand it is to think of a language as an independent entity, containing many sentences. Only some of its sentences have actually been used by its speakers. Frequently, how-

21. Similarly (disregarding the semantic issues), I don't think there can be any good reason for believing someone to speak a language with infinitely many sentences unless on the basis of believing that there is, in the sentences already uttered, a pattern of combination of recurring elements, and that the application of that pattern can be extended in order to generate new sentences (but not non-sentences).

ever, speakers are confronted with sentences never used before, and they manage to understand them. Compositionality is then invoked for explaining this. It may even be claimed that this sort of understanding would not be possible unless the language were compositional. It is possible to understand in this way the famous passage from Frege:

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of the sentence, so that the structure of the sentence serves as an image of the structure of the thought.²²

Again, this may seem like convincing argument. But again the argument employs a questionable premiss, e.g. the premiss that there are sentences which never have been used or encountered before, but which nevertheless *already* have a definite meaning. And again we must ask for the reasons for believing this. As far as I can see, the situation is exactly analogous, and even virtually the same, as in the argument from learnability. For it is difficult to see how there could be a reason for believing in semantically predetermined sentences which did not either involve the principle of compositionality itself, or else provided a direct argument for that principle. If we have a direct argument, then again, we would not need the further consideration about the understanding of new sentences.

In both these cases there is an assumption about the existence of predetermined facts, explained to be predetermined results of rule applications. If we have good reasons to believe in those predetermined facts, then we have also good reasons to believe that they are determined by rules, despite the rule-following considerations. The problem is that we don't have the independent reasons needed for believing in the predetermined facts.

22. Frege (1923), p 55.

The argument from communication

The only good reason we have for believing in compositionality is, I think, that there is so much of successful communication, the success of which cannot reasonably be explained by anything else than compositionality. Suppose you and I agree about the sentence

- (1) Let's meet exactly eight months from now at the intersection of 26th St and Broadway on Manhattan

And suppose we meet eight months later at the intersection of 26th St and Broadway on Manhattan, because we both expected to meet the other there and then. Could there a better explanation of this than that each of us was successful in grasping the thought expressed by the other? We may further assume that none of us had any previous experience with (1). I used it for the first time, and you heard it for the first time. No one had explained it to us, and we had not by any other means managed to agree about its interpretation. How come then, that by using it I managed to convey my thought? Could it be hints in the context, or charity of interpretation, or empathy? That any one of them, or any combination, could provide the solution, save under very special circumstances, is wildly implausible. The only workable explanation is the compositionality explanation, or, more cautiously, an explanation which involves compositionality. You and I assign the same meaning to the parts of (1), and we attach the same semantic significance to the structure of (1), and this leads us to interpret (1) as a whole in the same way. And this does provide a good explanation of how we could be successful in coordinating our future whereabouts.

Most of us have conversations every day, in very shifting contexts, about very shifting subjects – some of which have no direct relation to the contexts in which the conversations are taking place – and often with people of which we have virtually no personally knowledge. By ordinary standards most of these conversations are successful. Moreover, success is immediate. Speakers don't have to ponder for months about what a new utterance means. They get it right

instantaneously, without effort, mostly without knowing what making an effort would amount to. Misunderstandings do occur, but the rate of success, in cases where sentences used are new to both speaker and hearer, is astonishingly high. It cannot be adequately explained by random interpretation, charity, contextual clues or empathy, or any combination of these. But we can explain much of it by assuming compositionality. This is the basic reason for believing that the principle of compositionality is valid for natural languages.

Our belief is then reinforced by noticing the systematicity of our sentences, and even more by the fact that for some fragments of language we can actually devise semantic theories which do allow us to successfully predict how speakers will interpret sentences they have never heard before. But the frequency of communicational success, with the use of new sentences, is the basic reason.

3. Relevance of the rule-following considerations

We have settled for one particular way of arguing for the principle of compositionality. Now it is time to ask for the *way* in which this principle can be said to be valid for natural languages. What mode of validity is justified by the selected argument from successful communication?

Well, the principle is justified to the extent that its validity serves to explain successful communication. Any idea about its validity, or interpretation of the principle itself, which does not serve this explanatory purpose, lacks justification, and should not be accepted.

Suppose, for instance, that we think of a natural language as an ideal entity of which people have *knowledge*. You and I, it is said, both *know* the meaning of (1), and that is why we manage to understand each other. The idea that we know it must then involve the further conception that the meaning of (1) is *determined* by some factors independent of you and me, perhaps the rules of the language. Our knowledge of these rules, or these other independent factors, is what makes it possible for both of us to know the meaning of (1) itself. This is the story.

The problem with this story is that it contains two crucial elements which, for the purpose of

explanation, are redundant. Those are the assumption about independent *determination* of meaning, and the assumption about the *knowledge* of meaning on the part of the speakers. For you and I would understand *each other* equally well if we both made systematic mistakes about what the real meanings of sentences are, provided we made precisely the *same* mistakes. What explains the success of our communication is the assumption that we understand sentences in the *same* way. That it is also the *right* way, and that we *know* that it is the right way, does not add any extra explanatory force. Therefore, these assumptions are precisely *not* justified by the argument from successful communication.

If the mistake of this example shall be avoided, then we must take the principle of compositionality as a principle that is valid in this sense: it adequately *describes* a very general feature of the way speakers interpret, or assign meaning to, natural language sentences. The principle of compositionality is therefore to be understood as part of a *theory* about speakers, *not* as some very general rule about what sentences really mean, apart from what speakers actually mean by them.

But then, the principle *seems* to be concerned with meanings of linguistic expressions, not with speech and interpretation habits of speakers. There is no conflict, however. Semantic theories themselves are only justified to the extent that they serve to explain linguistic communication. Meanings can justifiably be assigned to individual expressions, and semantic significance to different syntactic and morphological constructions, only to the extent that these fit together into a semantic system that, as a theory of speakers of the language, does serve to explain communication. The principle of compositionality is just the overarching structural principle of such a system.

And this way of understanding the role of the principle is also what must guide us in getting the relevant sense of the word “determine” in the formulation of the principle:

- (C) The meaning of a complex expression is determined by the meanings of its parts and its mode of composition.

Since the principle itself describes very general features or properties of speakers, the same goes for the word “determine”. Speakers are so constituted that they arrive at certain interpretations of whole sentences on the basis of the component words and the modes of composition. As theorists, or self reflecting speakers, we can discern *patterns* in our assigning meaning this way, based on composition. And we assume further that there are more patterns in our speech habits than we have so far become fully aware of. A meaning that is *determined*, in the relevant sense, is just the meaning you arrive at by instantiating this *pattern* of interpretation. “Determine” is part of a description of *the structure of our interpretational dispositions*.

Once this position is reached, it is clear that the conclusion of the rule-following considerations is just *irrelevant* to the principle of compositionality. It would be relevant only if compositionality were understood as a kind of determination by rules, where the rules themselves, once established, settled what sentences *shall* or *do* mean, even if never used, and even if speakers would understand the sentences differently from what was determined by the rules. But this conception of what compositionality amounts to is of no interest, because it cannot provide any justification of the principle, and this, in turn, because compositionality, on this conception, does not have any explanatory value.

Compositionality is part of a theory about speakers, and as such the rule-following considerations are of no immediate relevance to it. This conclusion, however, does not just concern the principle of compositionality itself, but has implications for the general view of linguistic meaning. In particular this concerns the issue of *normativity* of linguistic meaning.

4. On rule-following and the normativity of meaning

My conclusion that the rule-following considerations are irrelevant to compositionality goes against virtually all of what has been said, that I know of, about this issue, even though they are fairly well in line with some of Noam Chomsky’s remarks. But writers such as Kripke, Crispin Wright, Gordon Baker, and indeed Chomsky, have thought that there is a genuine conflict.²³

However, saying that there is no conflict, although true, is misleading, since it conceals an underlying conflict about the nature of linguistic meaning. This conflict involves differences over the possibility of explaining communication and agreement in applications.

Explaining why people agree on applying terms, or why they manage to grasp each others' beliefs and desires on the basis of utterances, is an activity that perhaps can be carried on without in any way taking *meanings* into account. One might give causal, neurophysiological and other explanations of behavior, and of how one person's behavior can cause changes in the other person's beliefs. Semantics need not enter into it, and as long as it doesn't, there is no conflict between the possibility of such explanations and the rule-following considerations.

However, as soon as you bring in meanings in such explanations, there is, or at least seems to be, a conflict. And appealing to meanings in such explanations is the basic idea of the approach I favour. How come George gave the answer "125" to the addition problem "68+57", which he had never encountered before, and how come he applied the term "bicycle" to this physical object, in fact a bicycle, which he had never seen before? Intuitively, the explanation would involve an hypothesis about the meaning George assigns to the numerals, the plus and the identity signs, and to the word "bicycle". It appears that the rule-following considerations yields the result that this cannot be done.

Both Kripke and Wright are fairly explicit about this. Kripke notes²⁴ that:

We cannot say that we all respond as we do to '68+57' *because* we all grasp the concept of addition in the same way, that we share common responses to particular addition problems *because* we share a common concept of addition. ... Rather, our licence to say of each other that we mean addition by '+' is part of a 'language game' that sustains itself only because of the brute fact that we generally agree.²⁵

The same point, or virtually the same point is made by Wright:

23. See Kripke (1982), Wright (1980) and (1989), Baker (1980), Baker and Hacker (1985), Chomsky (1986). The conflict line is also taken in Niles (1992) and Shogenji (1995).

24. Kripke (1982), p 97.

25. Kripke (1982), p 97 (Kripke's emphases).

It is a basic fact about us that ordinary forms of explanation and training do succeed in perpetuating practices of various kinds – that there is a shared uptake, a disposition to concur in novel judgements involving the concepts in question. The rules-as-rails mythology attempts an explanation of this fact. But the reality is the other way about it: it is the agreement which sustains all rules and rule-governed institutions. The requirements which our rules impose upon us owe their existence to it.²⁶

The main point concerns new cases. New cases aren't pre-determined. So when two speakers agree on a new case they do not simply apply the rule, or the concept, they extend it. Therefore their agreement about the new case cannot be explained by appeal to their having the same concept, since that concept just *didn't* cover the new case. The fact that they applied the term in the same way isn't explained by the fact that they assigned the same meaning to it, since how this term should be applied to the new case wasn't determined by the meaning assigned to it by the speakers. We might indeed be able to appeal to linguistic structures for explaining speaker agreement in application of terms and acceptance of sentences, but we would not be justified in claiming that those structures are *semantic* structures.

Since the approach I have associated with the principle of compositionality involves the very idea that meaning is to be introduced in connection with the explanation of communication, there is clearly an underlying conflict over the concept of linguistic meaning. And I think that at the heart of this conflict we have a disagreement over the principle that

(N) An utterance can express a belief or other attitude only if the meaning of the expression uttered has been fixed (settled, nailed down) in advance.

I think of (N) as a principle of *normativity* of meaning, despite the fact that the formulation does not contain any overtly normative vocabulary. There is nothing about how words *should* be used, or about acting correctly or incorrectly. But the idea that *linguistic meaning* is normative isn't captured anyway by the idea that words in any reasonable sense *should* be used this way or that way, since semantics is one thing and how people should speak another. There are many

26. Wright (1989), p 244. Cf p 245.

factors, moral and prudential and others, determining how people should speak, which have nothing to do with questions of linguistic meaning.

You can also say that someone applies a predicate *correctly* if he applies it to something of which it is true, but that need not mean that meaning itself is normative in any sense, since what is correct and incorrect in this sense simply *depends* on the meaning of the predicate. The distinction comes out clearly by considering the determination schema again:

BASE L MEANING / RULE R APPLICATION

There are two separate questions, whether L is normative, and whether R is normative. An application of a predicate can be said to be correct if it is applied to something which satisfies the predicate. If you think that this by itself is what normativity of meaning amounts to, then you think of the R as normative. But it would be so, I think, only in an uninteresting sense.²⁷

The interesting question about the normativity of meaning is whether L is normative, and thus whether *having* such and such a meaning is a normative matter. But exactly what the normativity of meaning, in the sense of the normativity of L , really boils down to, is difficult to say. If there is a substantial issue of normativity, it should make a difference, storable in *descriptive* terms, whether it is true or not. It is unclear to me whether the very idea of prescriptive force can be given any theoretical content or significance. Still I think there are two theoretical ingre-

27. Kripke's remark that meaning is normative ((1982) p 37) seems not to amount to anything more than that. Similarly, Robert Brandom, who in Brandom (1994) repeatedly emphasizes the importance of the "normative" in semantics, does not distinguish between correctness as depending on truth and falsity, and normativity of truth conditions themselves. His only contrast is rather that between the normative and the *descriptive*, in the sense of natural or causal (see e.g. p 12). That is, R would be classed as normative, as long as it isn't causal. He ascribes the normativist position to Frege, but overlooks the fact that both Frege and Husserl objected to *both* the psychologists and the normativists of their time. Both Frege and Husserl claimed that the laws of logic and mathematics are *ideal* laws, and that to the extent that there are prescriptions for thinking, these can only be *derived* from the ideal laws that hold independently. The normativists thought that the norms were basic, and tended to be relativists.

The issue of the normativity of meaning is extensively discussed in Glüer (1997), largely in agreement with the views expressed here.

dients to be found in the normativity conception. Both concern determination of meaning, and both are meant to be captured by (N).

The first ingredient is *temporality*. If the meanings of words depend on rules, norms or prescriptions of some kind, then clearly a word cannot be meaningfully used unless its meaning has been determined in advance. And conversely, if it is true that meaning must be determined *before* meaningful use, then there must be something else than meaningful use of a word, and preceding it, that determines the meaning of the word, and that is precisely what a semantic norm should do.

The second ingredient is *sufficiency*. Intuitively, if e.g. an act of decision determines meaning, it isn't an example of normative determination if that act by itself isn't sufficient to determine meaning, but requires some extra principle of correlation to connect the act with the meaning. For in that case we would have to ask why that principle is the right principle, and how *that* is determined. Thus the requirement of sufficiency is a requirement of an internal, non-contingent relation between those facts, properties, states, events or acts that determine meaning and the meaning that is determined, for if the relation is contingent, we will need some extra principle.²⁸

Closely connected with this, I believe, is the idea that the very *concept* of meaning is somehow internally related to the meaning determining acts etc. For it is not just a particular meaning *entity* that is non-contingently related to such an act, but rather the fact *that x means y*. Somehow a certain pairing of an expression and a meaning must be determined, and it must be determined that the elements of the pair are related by the *meaning* relation, rather than simply related by some relation or other. So the concept of meaning must be involved. To the extent that the meaning determining facts are to be found in the mental life of the speakers, the concept

28. In Gampel (1997), p 226, normativity of meaning is said to consist in this: that meaning determining facts essentially have the property of justifying use. That seems to me closely related to the present requirement of sufficiency, or internality. Unfortunately it is not clear whether 'justifying' in Gampel means *correctness determining* or *citable as a reason*.

of meaning must, in a sense, be a *practical* concept. It must be a concept that is at least tacitly grasped by the speakers themselves, rather than something belonging squarely to theorizing about language.

Those two ingredients, the temporality and the sufficiency, are intended to be captured by the wording of (N). If what I have said is right, then (N) captures the descriptive aspects of the view that meaning is normative: if it is, then (N) is true, and if it isn't, then (N) is false. Since it is unclear whether prescriptiveness itself is theoretically significant, I shall simply regard (N) as the principle that meaning is normative. In a sense you could characterize the normativity conception of meaning as the view that *meaning precedes use*, temporally and conceptually, and the *non-normative* position as the position that *use precedes meaning*, at least conceptually, and to some extent also temporally.^{29 30}

Furthermore, (N) is related in the following way to the issue of applying rules or concepts to new cases. If you express a belief by means of the sentence 'Fido is faithful', i.e. the belief that Fido is faithful, then you are expressing a judgement which is correct if, and only, if the predicate '... is faithful' is correctly applicable to the referent of 'Fido'. By (N), you could not express a belief by this sentence unless its meaning had been defined in advance. But the meaning would not be defined unless the predicate '... is faithful' was defined for the argument Fido, i.e. the referent of 'Fido'. And the predicate would not be defined for Fido unless it was somehow determined whether it is correctly applicable or incorrectly applicable to Fido. So that must be determined in advance of the possibility of expressing a belief by the sentence. If the correctness of applying the predicate to Fido was not determined, then you could not *apply* it to Fido. You could *define* it for Fido, thus extending the meaning of the predicate, but you could not define it and apply it at the same time. If you could, those applications could not be mistaken, and applications which are immune to error are not applications of anything meaningful (which

29. In the Wittgensteinian tradition, and the tradition of ordinary language philosophy, the term 'use' usually meant *ways of using*, whereas 'usage' was employed to denote actual acts or utterances. Here I employ 'use' in the second sense.

of course is central to the private language argument).

The assumption of (N) is a driving force behind the rule-following considerations. The briefest rendering of the rule-following *reductio* argument would be as follows:

- (1) Rules determine correct applications in advance. (assumption 1)
- (2) What determines meaning is internally related to meaning.
- (3) (1) and (2) are incompatible.
- (4) Hence, (1) is false.

And here it is to be noted that (2) is a consequence of the requirement of normativity. The assumption of normativity is well illustrated in the crucial inference expressed in thesis 4 of the rule-following considerations:

- (4) The *expressing relation*, holding between a public rule formulation/explanation and the rule it expresses, is contingent. Therefore a rule cannot be publicly explained.

Intuitively, we *could* look at explanations like this: if you get what I am driving at thanks to my attempts at explanation, then I have indeed managed to explain, publicly, what I intended. If I

30. Given that several ideas are involved in the normative conception it is possible to find intermediate positions. One example is Ruth Garrett Millikan, whose philosophy largely is directed at giving a theoretical account of the normative character of meaning and other intentional phenomena. In Millikan (1990) this idea is applied to the rule-following discussions. Her position is that the meaning of an expression as used by a speaker is determined by facts about both the speaker's learning history of that word and the evolutionary history of his species (p 230). The determining depends on general principles about explanation in the theory of evolution. Reliance on such principles means that the sufficiency ingredient in normativity isn't present. On the other hand, if Millikan consistently counts only historical facts (and thus not facts about use) as meaning determining, then the temporality ingredient in normativity is present.

Another example is Michael Dummett. On the one hand Dummett gives a theoretical account of meaning, and on the other he requires (e.g. in Dummett (1978), last paragraph) that any speaker be able, by reflection, to recognize the truth of the theory as applied to her language. The speaker's knowledge of meaning therefore implicitly involves the theoretical concept of meaning. One might therefore characterize Dummett as requiring the sufficiency ingredient but not the temporality ingredient.

try to explain what it is to follow the rule +2, and I do it in such a way that you really get the rule I have in mind, then the attempt at explanation is successful.

But this is not how explanation is viewed in the rule-following considerations. Rather, the fact that no attempt at explanation is *unambiguous* means that there is no public explanation, because the meaning of the rule expression '+2' cannot be unambiguously *fixed* by any attempt at explanation. If you *do* get the right rule, then there must be a definite fixing of the meaning, but that fixing must take place in you, in your state of understanding. So it doesn't matter whether we are good or bad at making each other grasp the same meanings. It wouldn't even matter if we were virtually infallible in conveying the intended rules and the intended meanings to each other. We could still not be said to explain them, for the means employed in the explanations do not fix the meanings, and therefore do not make the expressions meaningful.³¹

Given this conclusion, (N) requires that meaning be fixed by private states of understanding, and this requires a non-contingent relation between the state of understanding and the rule or concept expressed. And the final steps in these considerations were to show that this is impossible. No finite state or entity, whether physical or mental, could be non-contingently related to an infinite rule, or even to unconsidered applications of a finite rule, and hence could not be what the understanding consists in.

Now, this view on linguistic meaning is to be contrasted with non-normative views. A typical representative of the non-normative position is Donald Davidson. On Davidson's view, what meaning an expression has is determined by the attitudes the speaker has to her own sentences. In this case those attitudes, like holding a sentence true, or wishing a sentence true, fulfils the role of use, which is reasonable, since it is precisely such attitudes that explain sincere speech acts. There is no question here of demanding that the meanings have been fixed by the speaker in advance of having or evincing the attitudes. On the contrary, and in this respect my views

31. This is not, of course, Wittgenstein's own view of explanation. For Wittgenstein, explanation is public (cf. ##208-210), but that depends on the rejection of assumption 1. This is especially evident in #209.

agree with Davidson's: semantic notions used in theories of natural language are *theoretical* and applicable to expressions of the language only on the *basis* of the use of the expressions.³² If this is right, then it cannot be required of the speaker to define or fix the meaning of her expressions, since that would force the speaker into a vicious regress.

On my view, non-normative as well, the explanation of successful communication is the point of departure. Just as compositionality can only be justified by appeal to the explanation of the understanding of new sentences in communication, I think, as indicated in section 3, that linguistic meaning in general should be viewed from the same explanatory perspective. A particular word *W* has a particular meaning *M* only if having this meaning uniformly contributes to meanings of sentences in a particular way, namely this: successful communication by means of these sentences can be well enough, and perhaps best, explained by a number of *semantic* assumptions, including the assumption that *W* means *M* in the language used. That is the general idea.

The data to be explained would consist of a number of linguistic transactions, together with data about the surrounding and other activities of the speakers. In some cases we could e.g. come to the conclusion that one speaker had gained knowledge of some particular state of affairs which he had not observed, but which his interlocutor had. This would be explained by the assumption that he had learned about this state of affairs by way of linguistic communication. And the assumed fact that this and other events of linguistic communication had been successful, would in turn be explained by certain semantic hypotheses, i.e. by hypotheses about the meanings of the words and sentences in the mouths of the speakers.

On such an explanation, we would already *assume*, but not without good reasons, that the speakers did express attitudes by means of their utterances. We would assume in general that speakers sometimes manage to convey to each other beliefs and desires of various kinds, often

32. The difference between my views and Davidson's concern the justification of compositionality. I claim that it can't be justified from considerations about charity and radical interpretation. See Pagin 1999.

highly context bound. We would not assume that for an occasion of successful linguistic communication to be possible the meanings of expressions used would have to be fixed in advance. That is, (N), the normativity claim would be rejected. One reason for this is the possibility, at least in special cases, to achieve understanding spontaneously, without the need of having the meaning of expressions determined in advance at all. Another reason is that the *concept* of meaning is taken to depend on the concept of a communicatively successful utterance.

So there is a fundamental difference over the validity of (N). If this difference cannot be settled, then we seem simply to have two different conceptions of linguistic meaning, not two opposing views about the same thing. On the other hand, if (N) is true, then the rule-following considerations are correct, and the approach to linguistic meaning I favour is mistaken. If, on the other hand, (N) is false, then my approach seems possible, and the rule-following considerations aren't fully motivated, since the requirement of a non-contingent relation between meaning and what determines meaning is not, in that case, adequate.

I think the difference pretty clearly can be settled. (N) is false, and the simple reason is that the conclusion of the rule-following considerations is in conflict with the conjunction of (N) itself and the true statement that we do express beliefs. For when I am applying an expression to a new case, that expression, by the rule-following considerations, is not defined for the case. I am free to choose how to apply it in the new case, even though steered by my own dispositions. But this means that although the meaning of the expression *was* fixed in advance, it wasn't fixed to the extent of covering the new case. And so, by the reasoning that motivated (N), I cannot express any *belief* by applying the expression to the new case. I cannot express any belief by the new sentence 'Fido is faithful', since it wasn't fixed in advance how '... is faithful' is to be applicable to the referent of 'Fido'. My applying the term, or asserting the sentence, would have the status of a definition, and therefore could *not* express a belief.

But that conclusion is plainly absurd. We do in fact express beliefs. As a result the normativity assumption must be given up, for we have this argument, connecting with the rendering

above of the rule-following argument:

- (N) A belief can be expressed by a sentence only if it is defined in advance.
- (6) If (1) is false, the meaning of sentences are not defined in advance.
- (7) We don't express beliefs (from (4), (N) and (6))
- (8) We do express beliefs. (fact)
- (9) Contradiction. Hence, either (4) is false (thus (1) true), or (N) is false.
- (10) But, if (1) is true, then (N) is false (from (3), since (2) follows from (N))
- (11) Hence, (N) is false. (from (9) and (10))

As far as I can see, this is a *reductio ad absurdum* of the normative conception of linguistic meaning. The conclusion crucially depends on (3), which is an intermediate conclusion of the rule-following considerations, and so in a sense these considerations themselves provide the *reductio* of the normativity view.

There is a final point to consider. I have here taken a stand on semantics that is pretty analogous to at least many things Chomsky says about Kripke in his discussion of rules of syntax. Chomsky is oscillating between finding a contradiction between his views and Kripke's, and seeing no conflict. In the latter case this is precisely because the grammatical rules Chomsky is discussing are descriptive, non-normative. Chomsky says e.g. that

Our theory of Jones – a grammar G of his language, a theory UG of the initial state of his language faculty – is a descriptive theory of his attained or initial competence, his knowledge system and its origins. Our theory is not a causal or dispositional theory concerning his behavior.

The rules of Jones's language are sometimes said to have "prescriptive force", but the term should be used with caution. These rules are not like the normative rules of ethics, for example. They entail nothing about what Jones ought to do (perhaps he should not observe the rules for one reason or another; they would still be his rules). And the question of the norm in some community is irrelevant for reasons already discussed- But whatever we conclude about the status of the rules, our theories about them are descriptive. ³³

Wright, however, objects that such a view

"threatens, like the dispositional account, to make a total mystery of the phenomenon of non-inferential, first-person knowledge of past and present meanings, rules and intentions".³⁴

In Wright's view there is a conflict between such a descriptive approach to rules as Chomsky's and the nature of first-person knowledge of one's language. The same should hold of every non-normative account of meaning, for a speaker's knowledge of the meaning of her own words seems to be a knowledge she has to derive from knowledge of her own use and some theoretical principles.³⁵

Is there such a conflict? Unsurprisingly, it depends on what you mean by "knowledge of meaning". If by "knowledge of meaning" you intend knowledge of the truth of statements such as

(2) 'chair' means *chair*

where "means" has the theoretical sense of the *meaning theory*, then clearly there is a conflict. No ordinary speaker (and at most few others) has any theoretical concept of meaning, and wouldn't even *understand* (2), let alone know that it is true, since that would require knowledge of both the theory and of the facts of use on which the truth of (2) depends. But there is nothing strange about that. If, on the other hand "means" as it occurs in (2) has an ordinary pre-theoretical sense, then it is just a disquotational truth, and there is no problem about knowing it.

However, by "knowledge of meaning" you can also intend something like this: a speaker knows the meaning of 'chair' if she associates the (right) meaning with the word and this association is generated and maintained by a reliable process. In this case there is no conflict. The

33. Chomsky (1986), pp 240-41.

34. Wright (1989) p 236.

35. In general it is analogous to, and to some extent overlapping with, questions of semantic externalism and first-person knowledge. And I think the answer to those questions are the same, too.

meaning theory will assign the concept of a chair as part of the contents of thoughts of mine expressed by sentences containing the word 'chair'. And in virtue of that I can be said to associate the concept of a chair with the word. But for this to be possible I need not have any *concept* of meaning whatsoever. I can know meanings without knowing that I do, for I don't need the concept of meaning for knowing the meanings themselves. And this knowledge of the meanings of my words is surely non-inferential. In fact, it should be treated as such by the meaning theory. As far as I can see, the alleged problem about first-person knowledge of meaning is just spurious.³⁶

Department of Philosophy
Stockholm University

References

- Baker, G 1981, 'Following Wittgenstein: some signposts for Philosophical Investigations ## 143-242.', in Holzman, S and Leich C (eds), *Wittgenstein: To Follow a Rule*, Routledge and Kegan Paul, London.
- Baker, G and Hacker, P, 1985, *Language, Sense and Nonsense*, Blackwell, Oxford.
- Brandom, R 1994, *Making it Explicit*, Harvard University Press, Cambridge, Mass.
- Chomsky, N 1980, *Rules and Representations*, Basil Blackwell, Oxford
- 1986, *Knowledge of Language*, Praeger, New York.

36. The first draft of this paper was written during a fellowship at SCASSS, The Swedish Collegium for Advanced Study in the Social Sciences, Uppsala, fall 1996. It was presented at SCASSS, in Stockholm at a seminar organized within the interdisciplinary project *Meaning and Interpretation*, in Prague, and in Lund. I am grateful to many participants at those occasions for comments, especially to Lars Bergström, Kathrin Glüer, Petr Kotatko, Sten Lindström, Per Martin-Löf, Paul Needham, Dag Prawitz, Gunnar Svensson and Dag Westerståhl and Åsa Wikforss. The last section of the paper was added after those seminars. I am especially grateful to Kathrin Glüer for comments on a penultimate version.

- Dummett, M 1978, 'What do I know when I know a language?', University of Stockholm pamphlet. Reprinted in Dummett, *The Seas of Language*, Duckworth, London 1993.
- Davidson, D 1967, 'Theories of meaning and learnable languages', in Davidson, *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford, 1984.
- Fodor, J 1987, *Psychosemantics*, The MIT Press, Cambridge, Massachusetts
- Frege, G 1923, 'Compound thoughts', in Frege, *Logical Investigations*, Basil Blackwell, Oxford 1977.
- Gampel, 1997, 'The normativity of meaning', *Philosophical Studies* 86, 221-242.
- Grandy, R 1990, 'Understanding and the principle of compositionality', *Philosophical Perspectives* 4.
- Glüer, K 1997, *Ist Bedeutung normativ? Versuch über eine Sprachphilosophische These*, Doctoral dissertation, Humboldt University, Berlin
- Kripke, S 1982, *Wittgenstein on Rules and Private Language*, Basil Blackwell, Oxford
- Millikan, R 1990, 'Truth-rules, hoverflies and the Kripke-Wittgenstein paradox', *The Philosophical Review* 99, pp 323-353.
- Niles, I 1992, 'Wittgenstein and infinite linguistic competence', in French, P.A., Uehling, T.E. and Wettstein, H.K. (eds), *Midwest Studies in Philosophy. The Wittgenstein Legacy*, Notre Dame, Indiana.
- Pagin, P 1999, 'Radical interpretation and compositional structure', in Zeglen (ed), *Discussions with Donald Davidson; Truth, Meaning and Knowledge*, Routledge, London.
- Shogenji, T 1995, 'The problem of rule-following in compositional semantics', *The Southern Journal of Philosophy*, Vol XXXIII, pp 97-108.
- Wittgenstein, L 1971, *Philosophische Untersuchungen*, Suhrkamp, Frankfurt am Main.

- Wright, C 1981, 'Rule-following, objectivity and the theory of meaning', in Holzman, S and Leich C (eds), *Wittgenstein: To Follow a Rule*, Routledge and Kegan Paul, London.
- Wright, C 1989, 'Wittgenstein's rule-following considerations and the central project of theoretical linguistics', in A. George (ed), *Reflections on Chomsky*, Basil Blackwell, Oxford.