THE UNIVERSITY OF CALGARY

SCEPTICISM ABOUT MEANING

An Examination of Quine and Kripke's Wittgenstein

by

Ólafur Páll Jónsson

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF ARTS

DEPARTMENT OF PHILOSOPHY

CALGARY, ALBERTA

AUGUST, 1997

To Anna

Table of Contents

Introduction

In 1960 Willard Van Quine published *Word and Object* and has ever since been among the most controversial philosophers of the century. But, despite a number of articles and three major conferences on Quine's philosophy followed by extensive publications including Quine's own responses, there is still a widespread misunderstanding of Quine's thesis of indeterminacy of translation. This may, in part, be due to the fact that Quine has changed his mind on a number of issues related to his thesis since 1960 but not given any comprehensive account of it.

In 1982 Saul Kripke published an exposition of the central arguments of Wittgenstein's *Philosophical Investigations*. This was *Wittgenstein on Rules and Private Language*, a book that was received with mixed feelings; it was welcomed with uncommon praise but it also sparked fierce objections. Much of the discussion has been concerned with the question whether Kripke's interpretation is fair to Wittgenstein's own views - a question that I will *not* consider - but it has also been recognized for its own merits. Kripke was credited with uncommon clarity; nonetheless interpretations of his work have differed greatly and unlike Quine, who has patiently responded to his critics, Kripke has not published any responses to any criticism of his thesis apart from what is already in *Wittgenstein on Rules and Private Language*.

Both these approaches have been labeled 'scepticism about meaning' or 'scepticism about semantic facts' or 'non-factualism about meaning'. And, perhaps, for obvious reasons. Kripke uses the expressions 'sceptical paradox' and 'sceptical solution' and concludes that there is no fact about me that determines what I mean, and Quine argues that there are no facts that can determine which of certain incompatible translations are right. But labels

seldom make things clearer - neither Quine nor Kripke's Wittgenstein subscribe to scepticism nor do they deny that there are certain facts about meaning. The facts are just not the facts that one might expect. But there are other striking similarities between Quine and Kripke's Wittgenstein; both reject any mentalistic account of meaning, and both emphasize communal agreement as fundamental to an account of meaning. They are also congenial in their use of considerations about meaning as a means to answer questions in philosophy of mind.

But despite these striking similarities and the fact that the works of Quine and Kripke have been widely discussed in recent years, a lengthy comparison of their approaches is painfully lacking. Kripke does mention a few points of contact and a few points where his interpretation of Wittgenstein's differs from Quine's thesis, but these are only comments and not detailed comparisons. The only attempt at a serious comparison that I am aware of is Dirk Koppelberg's paper from the San Marino conference on Quine's philosophy in 1995. But even if his discussion is interesting, it suffers from a serious misunderstanding which leads him to conclude that, despite certain similarities, Quine and Kripke's Wittgenstein are at distant poles in the philosophy of language.

My aim, in the first three parts of this thesis, is to give an exposition and a comparison of the approaches of Quine and Kripke's Wittgenstein, and I hope that I will succeed in providing something in the way of clarification rather than just add to the confusion. I conclude that far from being incompatible, these approaches could be seen as complementary. This may seem surprising since the claim that the notion of meaning is normative is central to Kripke's interpretation but does not play any role in Quine's approach. This difference is, however, explained by the fact that Kripke's Wittgenstein is concerned with the attribution of meaning to an individual's utterance - his conclusion is that such an attribution does not make any sense outside of a community - but Quine is concerned with interpretations of alien languages.

In the last part of the thesis I consider the plausibility of these approaches and find them too strong. As for Quine's thesis, I conclude that translation of ontological discourse will be determined even if translation of various other discourse will still be indeterminate. In other words, I reject his thesis about ontological relativity and, hence, the indeterminacy that goes with it. In the

case of Kripke's Wittgenstein I want also to weaken the conclusion. Instead of the claim that a community of speakers is necessary to make sense of rule following I claim that all we need is an external reality. This goes against the notorious private language argument, at least as it is interpreted by Kripke. But my conclusion is compatible with the claim that it is impossible to follow rules in the mind only. So, I can say that it is not possible to follow a rule privately, in the sense that doing something in the mind is to do it privately.

I have benefited from the help of various people. I want to thank Ali Kazmi for his patience and enormous help, C.B. Martin, Yosh Kobasigava and Atli Harðarson an Einar Logi Vignisson for discussing these things with me and the latter three for reading over a part of the thesis in a draft. I also want to thank C.B. Martin, Scott Soames and José Zalabardo for allowing me to use unpublished material og theirs. But all this would not have come to an end if it were not for the constant help and encouragement of my wife Anna Sveinsdóttir.

PART I

Quine's Philosophy of Language

One obstacle to overcome when studying a natural language like English is that we already speak English and various questions about English do not make much sense to us. Thus, for example, if I am walking in the woods with my friend and we see a rabbit scurrying by, my friend may say: 'That was a rabbit'. Now, did my friend mean that it was a *rabbit* or a *rabbit stage* or an *undetached rabbit part*? Of course he meant *rabbit*, we want to say, but is that due to the way the world is rather than the way the world is represented in English?

In order to answer this question Quine proposes a thought experiment which he calls radical translation. Quine tells a story similar to the one of me and my friend walking in the woods, except that in Quine's story my friend belongs to a previously unknown tribe which speaks a completely unknown language called Jungle, and I am a linguist trying to learn this language and set up a manual for translating it into English. Now, as we walk in the woods, a rabbit scurries by and my friend says 'Gavagai'. Now all the same questions arise. Did he mean that it was a *rabbit* or a *rabbit stage* or an *undetached rabbit part*? Sure enough, at this stage 'Rabbit' seems the most natural translation of the Jungle sentence, but only because that is how *we* talk about the world in *English*. Not because it was a rabbit rather than, say, a rabbit stage. And if there are no behavioral facts in virtue of which we can rule out the alternative translations, one linguist can set up a translation manual where 'Gavagai' is translated as 'Rabbit' and someone else can set up another translation manual where it is translated as 'Rabbit stage', and there will be no behavioral fact as to which is the right manual, or which is better. Which manual is better may be

indeterminate. This is Quine's thesis about indeterminacy of translation, or to use his own words:

> Manuals for translating one language into another can be set up in divergent way, all compatible with the totality of speech dispositions, yet incompatible with one another.[1]

But this formulation will probably give rise to as many questions as it answers; what exactly are manuals for translation, what is the totality of speech dispositions, and in what sense can these translation manuals be incompatible with one another? To answer these questions we must carefully explore the details of Quine's discussion.

Before we look into these details, it will be useful to bear in mind Quine's overall conception of philosophy as a part of the natural sciences – only the most abstract part. In the spirit of this conception is the version of the indeterminacy thesis in the closing paragraph of "Speaking of Objects".

> The whole truth about the most outlandish linguistic behavior is just as accessible to us, in our current Western conceptual schema, as are other chapters of zoology. The obstacle is only that any one intercultural correlation of words and phrases, and hence of theories, will be just one among various empirically admissible correlations . . . there is nothing for such a correlation to be uniquely right or wrong about.[2]

In giving an account of the behavior of some formerly unknown tribe, a scientist can come up with theories, or general descriptions of their behavior, including their linguistic behavior. These theories, or descriptions, will be factual; they are about facts, possible or actual. But if the scientist takes the further step of trying to translate the language of this tribe into, say, English he leaves the realm of facts. The correlation of words and sentences of the foreign language with some English words and sentences is not about anything. What

---

1 *Word and Object*, Cambridge Ma. 1960, p. 27. I will refer to this book as *WO*.

2 *Ontological Relativity and Other Essays*, New York 1969, p. 25. I will refer to this book as *OR*.

is factual in this project is just the fluency of conversation and effectiveness of negotiation that the translation serves to induce.[3]

In this chapter I will do four things: (i) Distinguish indeterminacy of translation from underdetermination of scientific theory and inscrutability of reference; (ii) clarify the reasons for the indeterminacy of translation; (iii) clarify the kinds and scope of the thesis; and (iv) spell out the difference between two forms of the thesis, an epistemological and an ontological form, and answer the question whether Quine saw these forms as just two sides of the same thesis or as two distinct theses.

My purpose at this stage is not to evaluate Quine's arguments, but to get clear about what these arguments are, what the premisses are and what conclusions Quine draws from those premisses. I will also identify some points about which Quine has changed his mind from the early views of *From a Logical Point of View, Word and Object* and *Ontological Relativity* to the later writings, especially *Philosophy of Logic, The Roots of Reference* and *Pursuit of Truth*. My main focus will, however, be on *Word and Object* since it contains the most detailed approach to the thought experiment of radical translation, and is presupposed to greater or lesser degree in the later writings.

Quine's illustration of his thesis is in terms of radical translation, a translation from a previously unknown language into English. The thesis has, however, wider application; it challenges the notions of proposition and meaning in general by casting doubt on a cross-language synonymy relation. If such a relation can be shown implausible, it will not make much sense to say that expressions of different languages have the same meaning, and, hence, it will not make much sense to say that they express the same proposition. But the thesis also casts doubts on certain questions about content of thought or belief. (*ORIT* 181) If it is right, we might give different, and even incompatible accounts of the beliefs of some individual of a different culture, and there need not be any way of determining which account is better as they may be equally good. This kind of consideration will, however, not be my concern here.

---

3 See for example Quine's *Pursuit of Truth*, revised edition, Cambridge Ma. 1992, p. 43. I will refer to this book as *PT*.

I. Radical Translation

Radical translation is, according to Quine, the limiting case of actual translation; it is actual translation where there are no cultural bonds to go by and no phonetic resemblances or resemblances of cognate word forms to guide the translation. (*WO* 28) If we remove this kind of evidence from ordinary translation, then, Quine argues, there will be nothing to go by for a linguist trying to break into a foreign language but behavioral evidence.

The first move in a radical translation is a correlation of stimulations and verbal behavior. All available evidence for the linguist, who carries out the translation, is verbal behavior and the circumstances in which this behavior takes place. The objective is to correlate bits and pieces of the verbal behavior to bits and pieces of the circumstances.

> All the objective data [the linguist] has to go on are the forces that he sees impinging on the native's surface and the observable behavior, vocal and otherwise, of the native. (*WO* 28)

But even when the linguist has determined what these forces are, he must make hypotheses about what forces constitute stimulations for the native. Before that has been done, the linguist can not so much as try to correlate stimulations and verbal behavior.

What then is the starting point of radical translation? One of the problems for the linguist is that circumstances do not come as bits and pieces – they appear as wholes, and, moreover, uninterpreted wholes. But some circumstances are, or at least appear to be, more easily divided into bits and pieces than others. The starting point will be short isolated utterances which seem to be responses to certain obvious changes in the circumstances. If a rabbit scurries by and the native says 'Gavagai', the linguist may, hypothetically translate 'Gavagai' as 'Rabbit'. But that is not to say that the rabbit is the reference of the native's utterance, only that the only relevant change in the circumstances, *as far as the linguist can tell*, seems to be the sudden scurrying by of the rabbit, and, assuming that things are similar form the point of view of

the native, his seeing of the rabbit will be the most obvious explanation of his utterance *for the linguist.*

Stimulations

Stimulations hold a central role in Quine's treatment of language learning and translation. But what exactly are stimulations? A stimulation is the evolving pattern of the triggering of sensory receptors during some convenient period of time. In *Word and Object* Quine gave the following definition:

> [Stimulations are] evolving irradiation patterns of all durations up to some convenient limit or *modulus*. (*WO* 32)

It is important not to equate stimulations and sense data. Sense-data are more closely related to awareness or mental acts than to surface irradiation; the identity of a sense-datum is given in terms of its qualitative character, while the identity of a stimulation is given in terms of numerical identity of triggered sense receptors. In short, while sense-data are usually related to the mind, stimulations are related to the surface. And Quine does, in fact, make very little use of sense-data; when concerned with linguistics or conceptualization he gives primacy to ordinary things, but, when doing epistemology he gives primacy to stimulations.[4]

Even though the terminology has changed in recent years, Quine has retained his understanding of stimulations.[5] But his emphasis on stimulations has sometimes been seen to be at odds with his insistence of the public character of language.[6] I will not go into these matters right now, Quine's view on the issue will emerge in due course. However, a few words explaining why

---

4 *Words and Objections: Essays on the Work of W.V. Quine*, D. Davidson and J. Hintikka eds., Dordrecht 1969, p. 298.

5 In 1992 Quine adopted the phrase 'neural intake' instead of 'stimulation' because of "inappropriate connotations or pointless debate on the part of readers who use the word 'stimulation' differently", but his definition of neural intake is in line with his earlier definition of stimulations: "let us define the subject's *neural intake* on a given occasion as the temporally ordered set of all firings of his exteroceptors on that occasion". See *On Quine: New Essays*, P. Leonardi and M. Santambrogio eds., New York 1995, p. 349.

6 See for example Dagfinn Føllesdal's paper "In What Sense Is Language Public?" in *On Quine.*

Quine sees the turn to stimulations necessary will be helpful. In *Pursuit of Truth* he says:

> I remain unswerved in locating stimulation at the neural input, for my interest is epistemological, however naturalized. I am interested in the flow of evidence from the triggering of the senses to the pronouncements of science. (*PT* 41)

Or, to put things in somewhat simplistic terms, we can say that Quine's concerns are inputs and outputs; the output is linguistic behavior, but the input, Quine claims, is best seen as the stimulation or, in his later terms, neural intake.

But why stimulation and not the external cause? Why is it the neural intake that prompts the natives utterance of 'Gavagai' and not the rabbit? The native may assent to 'Gavagai?', even if there is no rabbit, but, say a gopher, if the native mistakes the gopher for a rabbit. And, conversely, a rabbit may run by and the native may dissent from 'Gavagai?' since he may mistake it for something else, or not notice anything at all. That is why it is important to see what prompts the native's utterance of 'Gavagai' or his assent to the question 'Gavagai?' not as an external object, but certain stimulation; it is not the rabbit which is important but the native's seeing it. (*WO* 31) It is worth emphasizing that Quine's reasons for locating stimulations at the surface of the body is not that he thinks that external objects are somehow vague or less real.

> My naturalism does allow me free reference to nerve endings, rabbits, and other physical objects . . .. [failing a rabbit or other body to the purpose, perhaps the stimulus would be a shared situation] But I am put off by the vagueness of shared situations. (*PT* 41 and 42)

Now that we have Quine's definition of stimulation, or neural intake, and some motivation for turning to these, we focus on their role in translation.

Not all the forces impinging on the native's surface are relevant, since not all forces are relevant to stimulations. But not even all stimulations are relevant. Relevant stimulations are not just stimulations under which a native assents to or dissents from a particular sentence, but stimulations that *prompt* assent or dissent to that sentence.

> For suppose the queried sentence were one rather to the effect that someone is away tracking a giraffe. All day long the native will assent to it whenever asked, under all manner of irrelevant attendant stimulations; and on another day he will dissent from it under the same irrelevant stimulations. (*WO* 30)

Since what the linguist is after is a causal relation between stimulations and verbal behavior, it is not enough that certain verbal behavior takes place under certain stimulations, but that the stimulation actually causes that verbal behavior.

At this point, Hume's scepticism about causation will pose certain problems for the linguist, but, the linguist shares these problems with other scientists. Scepticism about causation is in no way, a more severe threat for the linguist than it is for other scientists.

Another problem is that of identifying stimulations; where does one stimulation begin and another end. What is a current stimulation and what is past stimulation? In practice, the linguist will have some standard of what to count as present and what to count as past stimulation. This standard is what Quine calls the *modulus of stimulation*. (*WO* 28) There is no ready made answer as to what modulus to adopt, not any more than there is an answer as to how long an occasion lasts, but, again, the problem is in no way essentially more difficult for the linguist than, say, for a doctor who is trying to give an account of what environmental factors increase the risk of cancer. If a hypothesis does not work, one of various ways of amending it is to change the modulus.

A third thing to notice about stimulations is their universal character. In classical physics, what enters into causal explanations are not particular objects, but abstractions from particular objects; not the third planet from the Sun, but a body with certain location, mass, velocity, etc. Prompting stimulations are, likewise, not particular events but a universal, repeatable event form. "We are to say not that two like stimulations have occurred, but that the same stimulation has recurred." (*WO* 34)

It is, however, not obvious how stimulations can be so universal. In *Word and Object* Quine wrote: "A visual stimulation is perhaps best identified . . . with the pattern of chromatic irradiation of the eye". (*WO* 31) But, on this

account, different individuals cannot have the same stimulation; *my* visual stimulation will be a pattern of chromatic irradiation of *my* eye, *your* visual stimulation a pattern of chromatic irradiation of *your* eye. Quine was well aware of this; of course the stimulations are not numerically identical, but, as long as the *patterns* are similar enough, we might say that the stimulations are of the same kind. And isn't that all we need? In this spirit is his remark in the first chapter of *Word and Object* about learning of terms.

> . . . if a term is to be learned by induction from observed instances where it is applied, the instances have to resemble one another in two ways: they have to be enough alike from the learner's point of view, from occasion to occasion, to afford him a basis of similarity to generalize upon, and they have to be enough alike from simultaneous distinct points of view to enable the teacher and learner to share the appropriate occasions. (*WO* 7)

Quine does not use the word 'stimulation' in this passage, but, his idea is that instances resemble one another in virtue of similar stimulations. But the two ways which Quine distinguishes are quite different. Recall that a stimulation is the activation of some subset of the subjects sensory receptors. Granted that these receptors stay fixed, or relatively fixed, over time, there is no problem in saying that someone has a similar stimulation as before; if much the same receptors are triggered on two occasions, we can say that the subject undergoes the same stimulation on these occasions. But when it comes to intersubjective likeness of stimulation, when we want to say that the teacher and the student *share* an occasion, the account cannot be so simple. What seems to be needed is the assumption that the sensory receptors are similar from one individual to another.

While it may not be an unreasonable hypothesis that sensory receptors are similar from one individual to another, a theory of translation should hardly depend on it. In some parts of *Word and Object*, Quine did explicitly reject the relevance of this hypothesis. That is the moral of the example of the trimmed bushes.

> Different persons growing up in the same language are like different bushes trimmed and trained to take the shape of identical elephants. The anatomical details of twigs and branches will fulfill the

> elephantine form differently from bush to bush, but the overall outward results are alike. (*WO* 8)

Now, while this tension is not made explicit in *Word and Object*, it became so only five years later (1965) in the paper "Propositional Objects". There Quine wrote:

> If we construe stimulation patterns my way, we cannot equate them without supposing homology of receptors; and this is absurd, not only because full homology is implausible, but because it surely ought not to matter. (*OR* 157)

The point is not that the assumption is wrong or could not be made sense of, but that this kind of assumption should not be needed.[7] For one thing, the linguist carries out his work in ignorance of neurophysiology, both on his part and the native's. But, from Quine's point of view, this assumption surely seemed to be needed, for how else could language be public, how else could the teacher and the student have something relevant in common – how could they share the occasion. In "Propositional Objects" Quine left the issue with much uneasiness:

> I leave you, therefore, with a problem of theoretical formulation that carries no evident practical problem with it. It is the problem of saying in general what it means for two subjects to get the same stimulation, or, failing that, what it means for two subjects to get more nearly the same stimulation than two others. (*OR* 159-160)

Later, Quine took a rather radical step in dealing with this problem. Instead of rejecting the conclusion, he gave up the idea of intersubjective likeness of stimulation. We could say that he gave a sceptical solution to the problem of intersubjective likeness of stimulation.

> The view that I have come to, regarding intersubjective likeness of stimulation, is rather that we can simply do without it. The

---

[7] The problem with this assumption is not the sceptical claim that we cannot know whether other people are like us, whether there are other minds etc., nor is it the sceptical claim that our senses are not reliable. It is rather that a theory of translation should not depend on neurophysiology.

> observation sentence 'Rabbit' has its stimulus meaning for the linguist and 'Gavagai' has its for the native, but the affinity of the two sentences is to be sought in the external of communication. (*PT* 42)

From the time of *Word and Object* to the present, the role of stimulations in Quine's philosophy has changed in a fundamental way. Stimulations are no longer a common coin, and neither are stimulus meanings.

Stimulus meaning

The first sentences to be translated by the linguist are, as mentioned earlier, those which appear to be directly caused by isolated events. The linguist is walking in the woods with his informant, a rabbit scurries by and the informant says 'Gavagai'. And the linguist experimentally translates that sentence as 'Rabbit'. But what exactly is the *justification* for that translation? At the time of *Word and Object*, Quine wrote that meaning is what a sentence shares with its translation, and since 'Gavagai' and 'Rabbit' share nothing but certain stimulations, we could arrive at the following crude notion of meaning:

> . . . *affirmative stimulus meaning* of a sentence such as 'Gavagai', for a given speaker, [is] the class of all the stimulations (hence evolving ocular irradiation patterns between properly timed blindfoldings) that would prompt his assent. (*WO* 32)

The negative stimulus meaning of a sentence for a given speaker was defined in a similar way as the class of all the stimulations that would prompt his dissent from that sentence at that time. And, assuming that the natives will not both assent to and dissent from the same sentence at the same time, affirmative and negative stimulus meanings of a sentence are mutually exclusive, i.e. a particular stimulation cannot prompt both assent and dissent to a particular sentence at a given time. But the affirmative and negative stimulus meanings of a sentence do not determine each other since there may be various stimulations that belong to neither. (*WO* 33)

According to the definition given in *Word and Object*, stimulus meaning was obviously relative to a given speaker as well as to time and modulus of stimulation. The reason for relativity to a given speaker was this: A speaker

may assign a stimulation to some expression that some other individual does not assign to that same expression, even if their assignments agree for the most part. But despite this relativity, stimulus meaning played a central role in translation in *Word and Object.*

> The imagined equating of 'Gavagai' and 'Rabbit' can now be stated thus: they have the same stimulus meaning. (*WO* 33)

Or, in other words, the justification for translating 'Gavagai' as 'Rabbit' was the similarity of the stimulus meaning of 'Gavagai' for the native and the stimulus meaning of 'Rabbit' for the linguist.

On Quine's later view, this picture has changed dramatically. As we have already seen, the matching of the native's and the linguist's stimulations cannot be, or at least, should not be assumed. And since no two speakers share any stimulations, they can not assign the same stimulation to any expression. Instead of having a difference in stimulus meaning from speaker to speaker that is a matter of a degree, we have a total difference. Despite the rejection of intersubjective likeness of stimulation, Quine did not give up the notions of affirmative and negative stimulus meanings, and, in fact, this change did not call for a rephrasing of the definition of stimulus meaning. What changed was its role in the translation procedure.

In his latest view the situation is not that of comparing stimulations of different individuals, but assigning the stimulus meaning of a sentence in the home language to some sentence in the target language.

> The observation sentence 'Rabbit' has its stimulus meaning for the linguist and the observation sentence 'Gavagai' has its stimulus meaning for the informant. The linguist observes natives assenting to 'Gavagai' when he, in their position, would have assented to 'Rabbit'. So he tries assigning his stimulus meaning of 'Rabbit' to 'Gavagai' and bandying 'Gavagai' on subsequent occasions for his informant's approval.[8]

---

8 "Three Indeterminacies", *Perspectives on Quine,* R. Barrett and R. Gibson eds., Cambridge Ma. 1988, p. 3.

But even if this does not in the end change his thesis of indeterminacy, it changes the balance between what is private and what is public. In *Word and Object* Quine claimed that stimulations were socially checked to some degree (*WO* 31), but once he moves stimulations into one's private affairs, translation can not be justified in terms of stimulations. That would hardly be consistent with his behavioristic orientation. But what could take over this justificatory role? In *Pursuit of Truth* Quine writes:

> The observation sentence 'Rabbit' has its stimulus meaning for the linguist and 'Gavagai' has its for the native, but the affinity of the two sentences is to be sought in the externals of communication. (*PT* 42)

What has replaced intersubjective likeness of stimulus meaning as the justification for particular translations is fluency of communication.

> [the] distinctive factuality [of rival manuals of translation] is blurred now by the disavowal of shared stimulus meaning. What is utterly factual is just the fluency of conversation and the effectiveness of negotiation that one or another manual of translation serves to induce. (*PT* 43)

This latest move, as radical as it may look, seems to be in a better coherence with Quine's holism, which, just as his behaviorism, has been one of the prominent themes in his philosophy from "Two Dogmas of Empiricism" until the present. (See e.g. *PT* 14) If intersubjective likeness of stimulus meaning made sense, it seems that translation of individual sentences could be checked; i.e. for some sentences rival translation manuals could be compared sentence for sentence just by comparing the nonverbal stimulus meanings assigned to these sentences by natives and linguists. According to Quine's latest view, this will not be possible since no such comparison will be possible. What can be compared, what is factual, is simply the overall use of each manual. And when a problem arises, it cannot be traced to any one sentence, not any more than dilemmas in sciences can be traced to individual hypotheses.

Types of sentences

The starting points of radical translation are sentences where the correlation between certain stimulations and verbal behavior is as direct as it gets. These are *occasion sentences* as opposed to *standing sentences.*

> Occasion sentences, as against standing sentences, are sentences such as 'Gavagai', 'Red', 'It hurts', 'His face is dirty', which command assent or dissent only if queried after an appropriate prompting stimulation. (*WO* 35-36)

A sentence like 'The winter has arrived' may be prompted by a particular cold morning and, in that respect, it would be similar to an occasion sentence. But the difference between this sentence and a sentence like 'His face is dirty' is that the former can be repeated throughout the day as well as the following days, even though the weather gets warmer, whereas the latter would need repeated prompting stimulations. An assent to the sentence 'The *Herald* has come' will be prompted daily by the arrival of the newspaper, but only by the current day's paper, i.e. the prompting stimulation must be repeated daily. This sentence appears to be somewhere in between 'The winter has arrived' and 'His face is dirty' as far as the need of reprompting stimulations goes. Standing sentences grade off toward occasion sentences as the interval between possible repromptings diminishes. An occasion sentence is the extreme case where that interval is less than the modulus. (*WO* 36) The difference between occasion sentences and standing sentences is then not an essential difference but one of a degree, and a standing sentence can be changed into an occasion sentence by lengthening the modulus, and vice versa. If the modulus of stimulation is one day, the sentence 'The *Herald* has come' will be an occasion sentence, but if it is only a few minutes or an hour, it will be counted as a standing sentence.

For a paradigm standing sentence, the stimulus meaning does not give many clues to how that sentence should be translated, but even for occasion sentences, the stimulus meaning has its shortcomings. What enters into the stimulus meaning of an occasion sentence can be all sorts of collateral information that would not count as part of the meaning of the sentence.

> There may be a local rabbit-fly, unknown to the linguist, and recognizable some way off by its long wings and erratic movements; and seeing such a fly in the neighborhood of an ill-glimpsed animal could help a native to recognize the latter as a rabbit.
>
> And, to be less fanciful, there are all those stimulations that incorporate verbal hints from native kibitzers. (*WO* 37)

The role of collateral information in stimulus meaning need not, however, refute the hypothesis that 'Gavagai' should be translated as 'Rabbit'. In *Word and Object* Quine wrote that translation is carried out, not by identity of stimulus meanings, but by significant approximation of stimulus meanings. (*WO* 40) As we saw earlier, Quine has given up the idea of intersubjective likeness of stimulus meanings and, instead of the translation being carried out by significant approximation of stimulus meanings, it will be carried out by a significant degree of success in assigning the stimulus meaning of 'Rabbit' to 'Gavagai'. If the linguist would assent to 'Rabbit' in significant majority of cases when the natives assent to 'Gavagai', the linguist might try to assign his stimulus meaning of 'Rabbit' to 'Gavagai'. Or, in other words, if a translation works in the majority of cases, the linguist is not put off by a few cases where it does not yield the expected result.

In some cases it may, however, be difficult to distinguish between collateral information and meaning of a sentence. The problem is that the linguist does not learn the language of the natives and their theories about the world separately. He is like a layman among scientists, who, unless becoming a scientist himself, may not be able to say when the scientists' assertions are prompted by something more or less directly observable, or when they are backed up by one or another theory.

Going back to the story about the rabbit fly, we can suppose that seeing the flies is the most common way of identifying rabbits, and that this is something the linguist has not realized. We might add to the story that the natives are not hunting rabbits but just watching them; there will be rabbit-watchers among them like there are bird-watchers among Canadians. Then, in the majority of cases where the natives will assent to 'Gavagai', the linguist will actually dissent from 'Rabbit' since he does not see any. All he sees are those weird

looking flies, and, perhaps, some ill-glimpsed movements in the grass. Granting this, the linguist will not find himself successful in assigning his stimulus meaning of 'Rabbit' to 'Gavagai'. Now, 'Rabbit' might still be a proper translation of 'Gavagai' for all we know.

There are, however, some sentences whose assent or dissent is rarely susceptible to collateral information, namely *observation sentences.*

> There is on this score a significant contrast between 'Red' and 'Rabbit' even when 'Red' is taken on par with 'Rabbit' as announcing not a passing sense datum but an enduring objective trait of the physical object. . . . there is less scope for collateral information in deciding whether a glimpsed thing is red than in deciding whether it is a rabbit. In the case of 'Red', therefore, sameness of stimulus meaning comes unusually close to what one intuitively expects of synonymy. (*WO* 41)

> Occasion sentence whose stimulus meanings vary none under the influence of collateral information may naturally be called observation sentences, and their stimulus meanings may without fear of contradiction be said to do full justice to their meanings. (*WO* 42)

This was at the time of *Word and Object,* but although Quine has not changed his mind about the role of collateral information in prompting assent or dissent, the last sentence depends on his discarded notion of intersubjective likeness of stimulation. However, before I consider Quine's latest views, let's consider the line of thought developed in *Word and Object.*

What seems to be the idea behind it is that there is nothing more to seeing colors than meets the eye; we do not have to make a projection from what meets the eye to the nature of the things that cause the visual stimulation. When we observe rabbits the situation is quite different. The nature of the thing observed is not something determined by the visual impact. And, if what meets the eye, as well as the eyes, is the same from person to person, there is no problem in saying that the stimulations are the same. Thus, on the hypothesis that sensory receptors are similar from person to person, we could say that the stimulus meaning of 'Red' is the same from person to person. But as we saw when discussing stimulations, it is one thing to say that the same individual has the same stimulation from time to time, and saying that two different

individuals have the same stimulation, even if it is at the same time in the same situation.

After his revision of the role of stimulus meanings, Quine can hardly say that the stimulus meaning of an observation sentence does full justice to its meaning; stimulus meanings are unavoidably bound to individuals but meanings of sentences are not supposed to be private in that way.

In later writings, Quine simply rejected his old definition of observation sentences from *Word and Object*, and replaced it with a twofold definition. He first defined what it is for a sentence to be observational for a given speaker.

> An observation sentence is an occasion sentence that the speaker will consistently assent to when his sensory receptors are stimulated in certain ways, and consistently dissent from when they are stimulated in certain other ways. If querying the sentence elicits assent from the given speaker on one occasion, it will elicit assent likewise on any other occasion when the same total set of receptors is triggered; and similarly for dissent.[9]

Then, Quine counts a sentence observational for the whole community if it is observational for each of its members. (*PT* 40) In this approach, two things must be noted. Firstly, whether a sentence is observational for a whole community, may depend on which community we have in mind. Thus, a sentence can be observational for a community of scientists, but not for some laymen, who, nevertheless are member of the same speech community. (*PT* 6) The second thing to notice is that the definition does not exclude the possibility for a sentence to be observational for the whole community, without it having the same meaning for each member of that community; it may even be such that someone assents to it, while another dissents from it on exactly the same occasion. But this need not be a defect of the definition. Most of us have doubtless witnessed or been involved in a dispute about whether certain thing is of this or that color. The dispute will turn on whether a certain sentence, say "This book is red" is true. The disagreement is not due to some collateral information since it is not disputed that the sentence in question is a genuine observation sentence. The dispute is, in a way, quite simple, and at the same

---

9 *Theories and Things*, Cambridge Ma. 1981, p. 25. I will refer to this book as TT.

time, unresolvable. For how could such a dispute be resolved? A scientific definition of the color red in terms of wavelengths of light would be of no help. Instead of a disagreement about the truth-value of the sentence "This book is red", we would have a disagreement about the appropriateness of the definition. The disagreement is, in fact, not about the physical properties of the book, but about the meaning of an expression. If the criterion for the application of that expression is only its stimulus meaning, then each member of the dispute will have his private criterion which, in principle, cannot be directly compared to others' criteria.

Logical connectives

In *Word and Object*, Quine proposed a rather straightforward method of translating the logical connectives 'not', 'and' and 'or'. The method was parallel to that of identifying the truth tables for the connectives with 'truth' and 'falsity' replaced by 'assent' and 'dissent'. The semantic criterion for identifying negation, Quine said, is that it turns any sentence that one would assent to into a sentence that one would dissent from, for conjunction the criterion is that it produces compounds that one assents to just in case he assents to each component, and, finally, the criterion for disjunction is that it produces compounds that one assents to just in case he assents to at least one if its components. For the purpose of these translations, the sentences put to assent or dissent could be either standing sentences or occasion sentences. (*WO* 57)

> When we find that a native construction fulfills one or another of these three semantic criteria, we can ask no more toward an understanding of it. Incidentally we can then translate the idiom into English as 'not', 'and', or 'or' as the case may be, but only subject to sundry humdrum provisos; for it is well known that these three English words do not represent negation, conjunction, and alternation exactly and unambiguously. (*WO* 58)

Quine did not see any serious obstacles in translating the logical connectives, even though he was well aware of the fact that the English words 'not', 'and' and 'or' do not represent the logical connectives. He may have had in mind sentences like "He undressed *and* took a shower", where the order of the

conjuncts cannot be changed, or replies like "Yes and no" which are not taken to be a conflation of the law of noncontradiction.

But these semantic criteria will, Quine claimed, not only suffice to identify these logical connectives, but it will also exhaust what understanding these connectives amounts to, or, in his words: "when we find that a native construction fulfills one or another of these three semantic criteria, *we can ask no more toward an understanding of it*" (my emphasis). Why are these semantic criteria all there is to the understanding of the logical connectives? Before we look at this questions, let us first consider the very task of identifying logical connectives in the native language.

Why should the linguist expect to find anything expressing the logical connectives? To this question there is a rather straightforward answer. We are simply not interested in languages that do not have these, or at least similar devices. An anthropologist or a linguist might wonder whether some previously unknown language was rich enough for sentential logic to be expressible in it, but Quine is neither concerned with anthropology nor linguistics in this sense. Quine's concern are natural languages comparable to English, i.e. a language that enables its speakers to communicate, not only about rabbits or bricks, but also about sciences. And, if we consider sentential logic to be necessary for this purpose, then, we should only consider languages where sentential logic can be expressed.

Granting that Jungle is rich enough for sentential logic to be expressible in it, we can ask further questions, namely about the identification of the connectives and the logical laws which the natives ascribe to: "Why should the linguist adopt the above semantic criteria for identification of the logical connectives?" and "Why should he take the natives to ascribe to the same logical laws as are commonly recognized in introductory courses in philosophy?". Now, can we answer these questions separately? In *Word and Object*, Quine's view was that these questions were inseparable.

> This approach ill accords with the doctrine of "prelogical mentality". To take the extreme case, let us suppose that certain natives are said to accept as true certain sentences translatable in the form '*p* and not *p*'. Now this claim is absurd under our semantic criteria. (*WO* 58)

Quine's point here is that given his semantic criteria, if the natives assent to a sentence that the linguist wants to translate into English in the form '*p* and not *p*', then the linguist should reconsider his translation of some native expressions as 'and' or 'not'. But what if certain native expressions fit the semantic criteria in majority of cases, but, on rare occasions, yield translations in the form '*p* and not *p*'? This, Quine insists, is not to be taken as a sign of the natives flouting the law of noncontradiction, and, after all, this is sometimes the case in English.

> Thus when to our querying of an English sentence an English speaker answers 'Yes and no', we assume that the queried sentence is meant differently in the affirmation and negation; this rather than that he would be so silly as to affirm and deny the same thing. (*WO* 59)

The point is, in a way, quite simple. If a sentence which the natives assent to is translated in the form '*p* and not *p*', there are two ways of accounting for this surprising result; they might ascribe to a different logic, or the translation could be wrong. And Quine rules out the former possibility as absurd. In *Philosophy of Logic* he writes: "It would seem that such an idea of deviation in logic is absurd on the face of it. If sheer logic is not conclusive, what is?"[10] and a little later, "The canon 'Save the obvious' bans any manual of translation that would represent the foreigners as contradicting our logic". (*PL* 83)

Now, any reader of Quine's "Two Dogmas of Empiricism" will, doubtless, recall his claim that no statement is immune to revision.

> Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery [i.e. observation sentence] can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws. Conversely, by the same token, no statement is immune to revision.[11]

---

10 *Philosophy of Logic*, second edition, Cambridge Ma. 1986, p. 81. I will refer to this book as *PL*.

11 *From a Logical Point of View*, second edition, Cambridge Ma. 1961, p. 43. I will refer to this book as *FLPV*.

But did Quine not say, both in *Word and Object* and *Philosophy of Logic*, that logical truths could not be changed, that they were indeed immune to revision? Here we must be careful, since it is not always clear when we are changing our theory, and when we are changing the meanings of our expressions. This problem is not a trivial one, neither in philosophy nor in the sciences. Quine's claim is that it does not make any sense to say that someone means genuine conjunctions by 'and' and genuine negation by 'not' and still holds that a sentence of the form '*p* and not *p*' can be true. When the truths of logic are changed, it is not a change in theory, but a change of subject. (*PL* 81) Quine sums this up in his reactions to the San Marino conference.

> Tension has been sensed in my attitude toward changes in logic: whether they are a change of heart or a change of subject; whether we are changing our theory or, as the analyticity suggests, just changing the meaning of our signs. I disavow the tension. At this level a change of theory is itself a change of meaning, though not always conversely. If we just write 'and' for 'or' and vice versa, we change meanings but not theory. If we abandon the law of excluded middle, we change meanings *and* theory; the law does not survive in any rendering.[12]

The point is, in a way, trivial. If to be analytic is to be true in virtue of meaning, and logical truths are analytic, then they cannot be revised without changing the meanings of the terms by which they are expressed.

Let us put the question of whether deviant logic is possible aside for a moment and look at other problems concerning translation of logical connectives. It is not so clear whether the behavioral evidence available to the radical translator, is sufficient to identify the logical connectives. As early as 1969 Quine noticed that the semantic criteria for translating the logical connectives he provided in *Word and Object* was insufficient.

> Since writing *Word and Object* I have observed . . . that the determinacy of translation even of the truth functions is less than complete. In the case of conjunction the gap is due to the fact that a speaker may dissent from a conjunction without dissenting from either component. Alternation has a similar gap, dually situated.[13]

---

12 *On Quine*, pp. 351-352.

13 *Word and Objections*, p. 314.

The reason for this problem is that some sentences, even if they are declarative, command neither assent nor dissent. This fits, actually, quite well into Quine's remarks in *Word and Object* that stimulus meaning is not totally defined:

> Yet the affirmative and negative stimulus meanings do not determine each other; for many stimulations may be expected to belong to neither. (*WO* 33)

In *Roots of Reference,* Quine takes up the issue and introduces a new kind of logical functions, *verdict functions.*

> Verdict logic is three-valued, the three verdicts being assent, dissent, and abstention. A compound sentence is a verdict function of its components if a verdict to the compound is determined for each assignment of verdicts to the components.[14]

It turns out that negation is at once a verdict function and truth function. But conjunction and alternation are not. The verdict table for conjunction is as follows:

p \ *q*	assent	abstain	dissent
assent	assent	abstain	dissent
abstain	abstain	?	dissent
dissent	dissent	dissent	dissent

The verdict table for alternation will also have a blind quarter. It is as follows:

p \ *q*	assent	abstain	dissent
assent	assent	assent	assent
abstain	assent	?	abstain
dissent	assent	abstain	dissent

14 *The Roots of Reference: The Paul Carus Lectures,* La Salle 1974, p. 77. I will refer to this book as *RR.*

If we cannot fill in the table for the question mark, we will not have verdict functions in Quine's sense; verdict to a compound consisting of components from which one abstains, is not determined. In order to obtain verdict functions, one might stipulate abstention at the center of the tables. But that seems not to be a viable way. A conjunction of the sentences "It is a mouse" and "It is a chipmunk" will be denied, even if both components are neither affirmed nor denied, but if the components are "It is a mouse" and "It is in the kitchen" and they are neither affirmed nor denied, the compound will perhaps neither be affirmed nor denied. The situation is similar for disjunction. Someone may affirm neither "It is a mouse" nor "It is not a mouse", but still affirm the disjunction of these two sentences. But even sentences of the form '*p* or not *p*' may be assigned abstention. Consider for example the observation sentences "This is a heap" and "This is not a heap". If they are being queried while pointing at a pile of 10 stones, or while pointing at one single stone, they will receive a definite answer. But, somewhere in between, perhaps if the stones are 3 or 4, neither will be affirmed nor denied, and that may also be true of their disjunction.[15]

It turns out that the logical connectives are not as easy to deal with as Quine originally supposed. In *Roots of Reference,* Quine concludes that:

> [Verdict functions] are more primitive than the genuine truth-functional conjunction and alternation, in that they can be learned by induction from observation of verdictive behavior. They are independent of our parochial two-valued logic, and independent of other truth-value logics. Truth values represent a more advanced, more theory-laden level of linguistic development; and it is in terms of theory, different theories for different subject matters, that we eventually learn (if at all) what verdict to give to the center of the verdict tables. (*RR* 78)

Having taken this step, it seems that Quine has left some space for alteration of logical principles. What must be built into every translation manual will not

---

15 This would, however, not mean that some truth of classical logic is held false, or that some classical contradiction is held true, but that a truth of classical logic or a logical contradiction is assigned abstention.

be truth-functional logic corresponding to the semantic criteria for truth-functions, but verdict logic corresponding to the verdict tables. And they are incomplete. After identifying expressions in the native language that yield translations corresponding to the determined parts of the verdict tables, the linguist can still fill in the gaps in more than one way.

We might say that the incomplete verdict tables provide a basis for the linguist in trying to identify expressions in the native language corresponding to the English words (not the truth-functions) 'and' and 'or'. These tables do not provide exhaustive criteria for the translation, but the criteria may be good enough. The filling in of the tables to obtain verdict functions will be a theoretical development and, as for other theories, whether physical or logical, this will be done in indirect ways.

Quantification

The question whether the existential and universal quantifiers will be recognizable in the native language depends on the availability of a semantic criteria for their identification. For the moment we are assuming that the native language actually has these kinds of quantifiers.

As we have already seen, Quine considers observation sentences to be translatable; there will be, in principle, no difficulty in translating some native sentence as "There is a rabbit in the yard". But, since this is a typical existential sentence, will the quantificational devices not be as readily available as the translation? The answer is "No". The reason why a translation of some native sentence as a quantificational sentence falls short of providing evidence for the natives' quantificational devices, are parallel to the reasons why the translation of the native sentence 'Gavagai' as 'Rabbit' falls short of equating the terms 'gavagai' and 'rabbit'.

> The difficulty is fundamental. The categorical depend for their truth on the objects, however external and however inferential, of which the component terms are true; and what those objects are is not uniquely determined by stimulus meanings. (*WO* 61)

The problem is the following: The linguist may successfully translate 'Gavagai' as 'Rabbit', but this only means that the natives assent to (or dissent from) the sentence 'Gavagai' in the same, or significantly similar situations as the linguist would assent to (or dissent from) the sentence 'Rabbit'. What is not established is that 'gavagai' and 'rabbit' taken as terms, rather than one word sentences, *denote* the same objects, or even the same kind of objects.

If the natives had a term synonymous to 'animal', the linguist might be able to identify sentences like 'All rabbits are animals'. But, in order to identify a native expression synonymous to 'animal', he would have to overcome all the same difficulties as when translating the term 'gavagai' as 'rabbit'. In short, he would have to identify the ontology of the native language, and that is exactly the stumbling block of translation.

There is, however, an alternative to the above picture. If we take the quantifiers to be *substitutional* rather than *objectual*, they will not carry with them the same kind of ontological burden. An existential substitutional quantification is true, if and only if there is an expression which, when substituted for the variable in the open sentence after the quantifier yields a true sentence. A universal substitutional quantification is counted as true if and only if no such substitution yields a false sentence. Consider for example the sentence "Someone was the teacher of Alexander the Great". If we take the existential quantifier to be objectual, we count the sentence true just in case there is some individual who taught Alexander the Great. If the quantifier is considered substitutional, we count the sentence true just in case there is an expression, say a name or a singular description, such that if 'someone' is replaced with that expression, the resulting sentence comes out true. On the latter approach, no ontological questions have to be settled in order to settle the truth value of the quantificational sentence.[16] All we need to do is to identify the appropriate substitution class, and check whether it contains an expression which turns the corresponding open sentence into a true sentence.

The difference between objectual and substitutional quantification does not only lie in the quantifiers themselves, the variables also play a different role. In

---

16 This is not quite correct, since substitutional quantification requires a non-empty substitution class, and our ontology must include the entities in that class. In this case, the substitution class would, most naturally, be a class of names, and our ontology would have to include names.

objectual quantification, a variable is a part of a denoting phrase, but in substitutional quantification variables are strictly placeholders, they play no role whatsoever in denotation.

> Behavioral conditions for interpreting a native construction as existential substitutional quantification, then, are readily formulated. We fix on parts of the construction as candidates for the roles of quantifier and variable; then a condition of their fitness is that the natives be disposed to dissent from a whole quantified sentence when and only when disposed to dissent from each of the sentences obtainable by dropping the quantifier and substituting for the variable. A second condition is that the natives be disposed to assent to the whole whenever disposed to assent to one of the sentences obtainable by dropping the quantifier and substituting for the variable. (*OR* 104-105)

As Quine recognizes, these conditions do not wholly settle assent, but, he maintains, they go far. There are various practical problems which the linguist would have to overcome. Suppose, for example, that the linguist is having a dinner with the natives, they are eating a rabbit stew. The natives may all assent to a sentence that could be translated as "Someone caught the rabbit" without being willing to assent to any substitution instance of this sentence since they may not know which one of them caught the rabbit that ended up in the stew. This problem is similar to the one we encountered in the case of the logical connectives. The semantic criterion for universal quantification has a parallel slack; someone may not be willing to assent to a universally quantified sentence but still be willing to assent to each substitution instance.

Now that we have distinguished between these two kinds of quantifiers, we might want to ask which one is rather found in natural languages. I do not see how we could answer this question for all natural languages, the best we can do is to look at English and try to determine how to understand expressions of the form "there is . . . such that . . ." or "for all . . . such that . . ." etc. While it may seem compelling to interpret the quantifier and the variable in the sentence "There is a thing x such that it is a rabbit" as objectual, it is not as obvious how to construe it in the following sentences: "There is a distance x such that it is three feet" or "There are three feet in a yard". We might end up saying that in

English, quantifiers and variables may sometimes be objectual and sometimes substitutional. But that won't take us far.

In *Roots of Reference*, Quine tries to answer the question whether English quantifiers are objectual or substitutional in terms of simplicity.

> The variable of the 'such that' construction, which is in effect the relative pronoun, is a substitutional variable at its inception. The words 'is a thing x such that' are learned by an equivalence transformation that is explicitly substitutional in character. And this variable, surely, is the variable at its most primitive. It is a regimentation of the relative pronoun. Variables begin as substitutional. (*RR* 99)

The point is that when we learn to use quantifiers and variables, we do this by means of substitution; we do this by taking the variable as a placeholder for another expression, usually a name. In this process, the focus is on the variable and not the quantifier. But Quine maintains that this is just how we learn the device of quantifiers, once we have mastered the technique, we begin to treat it as objectual.

> Once the relative clause or 'such that' construction has done its important work of siring quantification, a vital change takes place in the character of its pronoun or variable: it goes objectual. (*RR* 99)

This change, Quine claims, is "an irreducible leap in language learning". (*RR* 99) The reason for this is that the paradigm examples for quantificational expressions are "An apple is a fruit", "A rabbit is an animal" and the like, and in these cases it seems unnatural, if not absurd, to imagine names or singular descriptions for all apples and rabbits. Since most objects over which we are used to quantify are *in fact* both nameless and not uniquely identifiable by some linguistic devices such as singular descriptions, it is unnatural to construe the truth-conditions for ordinary quantificational sentences as depending on there being names or singular descriptions for these objects. This is not to say that quantification in English is never substitutional, only that it is objectual in the paradigm cases.

But, Quine claims, the objectual quantification carries with it the burden of ontological commitment. This, however, is not trivial. Since the same

evidence will verify the sentences "There is a rabbit in the garden", "There is a rabbit stage in the garden" and "There is an undetached rabbit part in the garden", it seems that we could assess the truth-value of a corresponding Jungle sentence without determining the ontology of the natives. Even when we quantify over abstract objects, we may not be involved in any ontological presuppositions. If the sentence is "There is a prime number between 10 and 20" the evidence will be computation and we do not need to settle on some ontology of numbers. But it will be much harder to say what evidence could support or refute sentences like "There are objects" or "There are numbers". (*OR* 97) This is not to say that these matters are arbitrary, but that whether we do or do not include these things in our ontology depends on reasons that are not so straight forward. Far from being arbitrary, matters are far removed from the most concrete facts.

> Our theory of nature grades off from the most concrete fact to speculations about the curvature of space-time, or the continuous creation of hydrogen atoms in an expanding universe; and our evidence grades off correspondingly, from specific observation to broadly systematic considerations. Existential quantifications of the philosophical sort belong to the same inclusive theory and are situated way out at the end, farthest from observable fact. (*OR* 98)

When discussing the truth-functions we saw that the verdict tables for conjunction and alternation were incomplete, and, consequently, the translation of these connectives was not behaviorally determined. When it came to quantification, similar things took place. It is possible to give rough semantic criteria for substitutional quantification, even if they depend on the logical connectives and inherit their indeterminacy. But when it comes to objectual quantification, no such criteria are possible; it belongs rather to theory of nature than to semantics. And as for any other theory of nature, it is perfectly reasonable to ask whether the natives do at all have anything corresponding to it. It is possible that we would be unable to arrive at any satisfactory translation of 'there is', and hence of existential quantification.

> Some languages are perhaps so unlike ours that any translation of 'there is' or '$\exists x$', however cunningly contextual, would be too farfetched and Procrustean to rest with. To entertain the notion of an

> ontology at all, known or unknown, for the speakers of such a language would be an unwarranted projection on our part of a parochial category appropriate only to our own linguistic circle. (*PT* 28)

In *Word and Object* Quine saw the problem with objectual quantification as that of determining the ontology of the natives. Now he goes so far as to allow for the possibility of no ontology at all.

Synonymy and analyticity

There are two distinct problems related to synonymy, synonymy of terms and synonymy of sentences. It seems to be a rather obvious consequence of what has already been said about denotation of terms, that synonymous terms cannot be recognized in radical translation. As for synonymy of sentences it can be recognized for some sentences, not for others. But what kind of definition or explication are we after? Obviously we do not just want to say that sentences are synonymous if they have the same meaning, for we do not yet know what that would be. What is needed is some extra-linguistic definition of a synonymy relation, such that for any arbitrary sentences S_1 and S_2, we can say whether they bear this relation to one another.

For occasion sentences, Quine says that intrasubjective stimulus synonymy captures the envisaged notion of synonymy pretty well. (*WO* 46 and 62) This may sound strange in the view of Quine's earlier claims that stimulus meaning does not do justice to the meaning of non-observational occasion sentences. But, if there is something strange here, it is only the sound of it. The reason why stimulus meanings did not do justice to the meanings of non-observational occasion sentences was that collateral information, which by no means can be counted as part of the meaning of such a sentence, may be part of its stimulus meaning. But, for any synonymous sentences, any collateral information that prompts assent to (or dissent from) the one would do the same for the other as long as we stick to one speaker.

Two expressions are then called synonymous for a whole community if they are synonymous for each member of that community. This notion of synonymy can even be extended to two-language cases by means of bilingual speakers. A bilingual speaker will represent two linguistic communities, but

whether he is a good sample will be checked by observing the fluency of his communication in those communities and by comparison to other bilinguals. (*WO* 47)[17]

The notion of stimulus synonymy meets the requirement of being extra-linguistic; it is defined in terms of stimulus meaning which is not linguistic. To broaden the notion of synonymy, we might say that two sentences are synonymous if they "command assent concomitantly and dissent concomitantly, and this concomitance is due strictly to word usage rather than to how things happen in the world". (*WO* 62) But how far can we push this notion of synonymy? By lengthening the modulus of stimulation the relation of stimulus meaning will become tighter, but there are practical limitations to how long the modulus can be. Moreover, some sentences are such that they do not stand in any direct correlation to experience. These would be paradigm standing sentences.

> The significant trait of other [i.e. standing] sentences is that experience is relevant to them largely in indirect ways, through the mediation of associated sentences. (*WO* 64)

For standing sentences straightforward stimulus synonymy will not suffice; on that account all affirmed standing sentences would come out synonymous. To try to extend the notion of synonymy to standing sentences one might define synonymy such that sentences are synonymous if they are stimulus synonymous with respect to any background theory, or associated sentences, as Quine puts it. Calling the background theory S, the proposed definition of synonymy becomes:

> S_1 and S_2 are synonymous if for every S the conditional compound of S and S_1 and that of S and S_2 are stimulus-synonymous. (*WO* 64)

In other words, S_1 and S_2 are synonymous for a given speaker if the sentences '$S \rightarrow S_1$' and '$S \rightarrow S_2$' are stimulus synonymous for any S for that speaker. (*WO* 65) But does this notion of synonymy take us any further than the notion of

---

[17] In his "Reactions" from the San Marino conference, Quine drops the expression 'stimulus synonymy' and uses instead 'perceptual equivalence' but this does not indicate any change in his theory. See *On Quine*, pp. 349-350.

stimulus synonymy did? One disadvantage of this definition is that it is not just given for any arbitrary S_1 and S_2, but for some set of associated sentences S as well. But what sentences are in S? If S itself contains standing sentences the resulting notion of synonymy will be relativized to a particular language, it will be intralinguistic. But if S contains only observation sentences, it seems that the assessment of '$S \rightarrow S_1$' and '$S \rightarrow S_2$' will be arbitrary, at least in some cases. Let S_1 and S_2 be some arbitrary standing sentences that are compatible with all possible observation. Then both conditionals will be consistent and the stimulus synonymy of the conditionals will boil down to the question whether S_1 and S_2 are assigned the same truth value. But we did not just want to say whether these sentences were both true, but whether they were synonymous.

The notion of stimulus synonymy, even as socialized, has its shortcomings even in the case of occasion sentences. It does not guarantee that the sentences in question have the same meaning for each member of the community. Suppose that sentences A and B are such that each member of the community will assent to A just in case he would assent to B. This does not mean that each member of the community would assent to A and B under the same circumstances. The meaning assigned to these sentences by one member of the community might, for that matter, be contrary to that of some other member. What the notion of stimulus synonymy is capable of doing is only to group sentences together, not to identify their meaning.

The notion of analyticity meets all the same obstacles as that of synonymy since they are interdefinable with one another.

> Sentences are synonymous if and only if their biconditional (formed by joining them with 'if and only if') is analytic, and a sentence is analytic if and only if synonymous with self-conditionals ('If *p* then *p*'). (*WO* 65)

And as synonymy is related to analyticity, so is stimulus synonymy related to stimulus analyticity; i.e. the notion of analyticity fails to provide a firmer relation between sentences than that of stimulus synonymy.

But can we not define analyticity in some different way? Quine considers the following proposal for the definition of analyticity that is not given in terms of the notion of synonymy: Analytical sentences are those we are prepared to affirm come what may. But before this proposal can be taken seriously, the

phrase 'come what may' must be made clear. If it means, come what stimulation may, what we get is, at best, stimulus analyticity, on whose account the sentence 'there have been black dogs' will turn out analytic. (*WO* 66) And that is surely not what we were after.

How far does the linguist get?

Now that we have seen Quine's setting of radical translation and the central notions involved, it is time to bring things together; see how far the linguist can get and what sentences he can translate and by what certainty. Let's start with Quine's own summary in *Word and Object*.

> (1) Observation sentences can be translated. There is uncertainty, but the situation is the normal inductive one. (2) Truth functions can be translated. (3) Stimulus analytic sentences can be recognized. So can sentences of the opposite type, the "stimulus-contradictory" sentences, which command irreversible dissent. (4) Questions of intrasubjective stimulus synonymy of native occasion sentences even of non-observational kind can be settled if raised, but the sentences cannot be translated. (*WO* 68)

Standing sentences are still outside the scope of possible translation and so are most occasion sentences, namely those that are not observation sentences. But more importantly, the linguist cannot determine what words are terms, much less what terms are coextensive. In fact, the linguist has merely been able to give an account of a fragment of the target language, and in later years, Quine has seen this fragment become smaller as he does not consider logical connectives to be translatable with the same certainty as before.

Analytical hypotheses

To pass these bounds, the linguist forms what Quine calls *analytical hypotheses*. As for establishment of hypotheses in the sciences in general, there are two steps to the linguist's establishment of his analytical hypotheses. The first step is to identify regularities in the diverse data.

> [The linguist] segments heard utterances into conveniently short recurrent parts, and thus compiles a list of native "words". Various of these he hypothetically equates to English words and phrases, in such a way as to conform to (1)-(4). (*WO* 68)

On the bases of observed regularities, the linguist proposes some hypotheses which he then tests against what has already been established under (1)-(4) above.

> The sentence translations derivable from the analytical hypotheses are to include those already established under (1); they are to fit the prior translation of truth functions, as of (2); they are to carry sentences that are stimulus-analytic or stimulus-contradictory, according to (3), into English sentences that are likewise stimulus-analytic or stimulus-contradictory; and they are to carry sentence pairs that are stimulus-synonymous, according to (4), into English sentences that are likewise stimulus-synonymous. (*WO* 68)

At least three things should be noted about analytical hypotheses. (i) They "exceed anything implicit in any native's dispositions to speech behavior" and, thus, "extend the working limits of translation beyond where independent evidence can exist"; (*WO* 70) (ii) observable behavior gives them only negative support, i.e. it won't determine whether they are right, but it may prove them wrong; and (iii) when analytical hypotheses are tested against evidence established under (1)-(4), they are not tested one by one, but as a whole. All this has its parallel in other sciences where theories exceed the available data, experiments provide refutation rather than confirmation, and theories are not tested one by one but as a part of bigger wholes.

The translation methods are essentially the same even if the linguist becomes bilingual, and, consequently, the correlations of sentences of the two languages made by a bilingual, are no more uniquely right than are the relations made by a monolingual linguist, since:

> . . . another bilingual could have a semantic correlation incompatible with the first bilingual's without deviation from the first bilingual in his speech dispositions within either language, except in his disposition to translate. (*WO* 74)

II. Indeterminacy of Translation

We are now able to distinguish Quine's thesis of indeterminacy from other indeterminacy theses. In Quine's writings, three theses of indeterminacy have figured conspicuously: indeterminacy of translation, inscrutability of reference and underdetermination of scientific theory. Indeterminacy of translation is related to both the other indeterminacy theses, but it is not the same as either of them.

Underdetermination of scientific theory

Underdetermination of scientific theory has many forms. For present purposes I will distinguish two, (i) a weaker form, according to which scientific theory is underdetermined by all past evidence, or even by all past and future evidence; and (ii) a stronger form which claims that scientific theory is underdetermined in principle. According to the weaker form, there will always be incompatible theories that are equally in accordance with all possible evidence. The stronger form claims that there will always be such incompatible theories, but, moreover, it claims that all possible evidence along with an ideal organon of scientific method will not be able to prove one rather than the other right, nor even better.[18] The stronger form, I should emphasize, is not tantamount to the claim that there is no fact of the matter as to which theory formulation is correct. That claim would only follow given some sort of instrumentalism, i.e.

---

[18] See Michael Friedman, "Physicalism and the Indeterminacy of Translation", *Noûs*, Vol. 9 No. 4 1975, pp. 356-357.

it would only follow given the additional premise that there are no facts that are unknowable in principle, a premise that Quine explicitly rejects.

It is not always clear which form Quine has in mind when discussing underdetermination of scientific theory. Sometimes, however, he is clearly talking about the stronger form, for example in *Word and Object* when he writes:

> The incompleteness of determination of molecular behavior by the behavior of ordinary things is hence only incidental to this more basic indeterminacy: *both* sorts of events are less than determined by our surface irritations. This remains true even if we include all past, present, and future irritations of all the far-flung surfaces of mankind, and probably even if we throw in an in fact unachieved ideal organon of scientific method besides. (*WO* 22)

But in what sense can rival theories be incompatible? Can they be contradictory, or will they just be such that neither can be reduced to or interpreted in terms of the other? To answer these question we have to look at Quine's account of the relation between theory formulation and observation. A key notion in Quine's discussion of these matters is that of *observation categoricals*. These are sentences like 'Where there is smoke there is fire' or 'When night falls the lamps are lit', and their role is to link observation sentences, say 'There is smoke' and 'There is fire' to formulation of theories which will consist of one or more standing sentences, say 'Smoke is only caused by fire' or 'The twilight turns the lamps on'. The relation between theory formulation and observation categoricals is not that some observation categoricals imply some theory formulation, but the other way around. Some theory formulation will imply some observation categoricals, and, moreover, different theory formulations may imply the same observation categoricals. This then gives a way to say what it is for two theories to be empirically equivalent.

> A theory formulation merely implies its observation categoricals, and is not implied by them, unless it is trivial. Two theory formulations

> may thus imply all the same observation conditionals[19] without implying each other. They can be empirically equivalent without being logically equivalent. (*TT* 28)

Given this understanding of empirical equivalence of theories we can be more specific about the degree of incompatibility that rival theories can enjoy. Suppose that T is a theory formulation and _ is the set of observation categoricals implied by T. There may be some sentence S such that neither it nor its negation implies any observation categoricals, and both are consistent with the rest of T. We may then take the conjunction of T and S to be a theory formulation, as well as the conjunction of T and not-S. These theory formulations will be empirically equivalent but logically incompatible, i.e. their conjunction will be unsatisfiable. To this one might object that the sentence S is totally redundant and the parts of the theory formulations which matter are not incompatible.[20]

Another way of construing empirically equivalent but logically incompatible theories is by taking a pair of terms that appear in T but do not appear in any sentence in _ and switching them. Suppose for example that the terms 'molecule' and 'electron' appear in T but not in _. Then we can make a new theory formulation T* by switching these terms. The two theories will be logically incompatible but empirically equivalent. To this proposal one might, however, object that the two theory formulations are not two theories but only one, just in slightly different terms; in any case, the one can easily be translated into the other just by switching the terms again. (See *TT* 28-29)

---

[19] This, I think, is a mistake and should be "categoricals". Quine's notion of 'observation conditionals' is an older relative of 'observation categoricals'. An observation conditional is a sentence of the form 'If there is smoke then there will be fire'. The disadvantage of these sentences is that they do not have the same generality as the observation categoricals, since the initial conditions expressed in the antecedent refer to times and places at some removed from what is referred to in the consequent. The problem is this: "At the place-time where the predicated observation is due, how does the experimenter know that the supposed initial conditions were fulfilled a while back and some way off?". See *Theories and Things*, p. 27.

[20] Quine does not mention this way of construing empirically equivalent but logically incompatible theory formulations. This way differs from the one below, which Quine mentions, in a significant way since it does not mention terms and meanings of terms and might, therefore, be recognized in a radical translation, while the other might not be recognizable.

But suppose now that we have two empirically equivalent theory formulations that we see no way of reconciling by reinterpreting some words or by deleting some redundant sentences. This would be the case in Poincaré's story where one theory formulation states that the universe is infinite but the other that it is finite. But unlike Poincaré's story, it may well be unknown that the theory formulations are empirically equivalent. But even this inconsistency may be resolved.

> However, the specter is easily laid, by a move just as trivial as our recent switch of 'molecule' and 'electron'. Being incompatible, the two theory formulations that we are imagining must evaluate some sentence oppositely. Since they are nevertheless empirically equivalent, that sentence must contain terms that are short on observation criteria. But then we can just as well pick out one of those terms and treat it as if it were two independent words, one in the one theory formulations and another in the other. (*TT* 29-30)

It should be clear from Quine's general motivation that he holds that scientific theories are ontologically underdetermined. However, this strong underdetermination does not give room for theory formulations whose conjunction is unavoidably logically unsatisfiable. For the empirical slack gives room for reinterpretation of some of the theoretical terms in those theory formulations, and instead of inconsistent formulations we end up with theory formulations that contain terms irreducibly alien to one another, and, consequently, the formulations themselves can not be reduced to one another nor to a third theory. But, as Quine notes:

> Our language can embrace the full vocabulary of both theories, and our truth predicate can then apply to each on its separate merits.[21]

It may seem tempting to say that the empirically equivalent but incompatible theories arise because of a mixing of translation manuals. However, that is not right. As already noted, these theory formulations can occur in the same language and, as Quine says, "our truth predicate can apply to each on its own merits". But translations of one and the same theory formulation from one

---

21 *Perspectives on Quine*, p. 15.

language to another guided by different translation manuals, may be empirically equivalent and logically incompatible. But that is true of any standing sentences.

Inscrutability of reference

Inscrutability of reference arises also independently of indeterminacy of translation. Successful translation of observation sentences, does not fix the reference of the terms of which those sentences are composed, nor will such translation determine what words are terms. This is what Quine's hypothetical example of 'gavagai' was originally meant to show, (*PT* 51), and the Japanese classifiers would be a real example of this. (*OR* 35-38)

Further difference between inscrutability of reference and indeterminacy of translation is highlighted by Quine's notion of ontological relativity. Questions about ontology are not raised at the level of translation of observation sentences. Whether 'gavagai', as a term, refers to rabbits, rabbit stages, undetached rabbit parts etc., does not matter since the stimulus meaning of the corresponding observation sentences is identical. Here we are confronted with the same problem as when considering the possibility of semantic criteria for objectual translation. Given that we have identified the quantificational expression, we are still unable to infer what kind of objects the variables range over.

Now we have seen that underdetermination of scientific theory and inscrutability of reference arise independently of indeterminacy of translation, but is indeterminacy of translation independent of the other two?

Kinds and scope of indeterminacy of translation

Before we go any further, two matters should be settled: (i) What kind of indeterminacy does translation admit of, i.e. is translation only epistemically underdetermined or is it ontologically underdetermined, and (ii) what is the scope of this indeterminacy, i.e. to what expressions of the language does this indeterminacy extend.

As we have already seen, Quine holds that scientific theories are underdetermined in the strong sense, i.e. underdetermined in principle, and this underdetermination is carried over to translation, even at the level of observation sentences. The translation of 'Gavagai' as 'Rabbit' is not based on *reduction* to behavioral facts (or even to physical facts); it is justified by reference to behavioral facts, but, alternatives like 'Rabbit stage' or 'Undetached rabbit part' are equally justifiable. A justification, or rejection, of 'Rabbit' as a translation of 'Gavagai' will be based on the fluency of communication that a manual, where 'Gavagai' is given that translation, serves to provide. Alternative manuals that work just as well will be just as good. This, however, is not a defect of the translation of 'Gavagai' as 'Rabbit'.

To be more specific about how underdetermination of scientific theory enters into indeterminacy of translation, we should remember how translation of observation sentences is carried out. The translation is carried out by significant success of assigning the stimulus meaning of an expression in the home language to an expression in the target language. This assignment will be a kind of theory formulation and will imply some observation categoricals, say 'Whenever 'Gavagai' is true, 'Rabbit' is true'[22]. But alternative translations of 'Gavagai' such as 'Rabbit stage' will yield all the same observation categoricals. It is only when we proceed from sentences to terms that the difference between 'Rabbit' and 'Rabbit stage' becomes visible. But that is to say that the terms 'rabbit' and 'rabbit stage' are short of observation criteria, which, as we just saw, gives room for empirically equivalent but logically incompatible theory formulations.

Even at a higher level of observationality, certain underdetermination will still occur. Instead of identifying a universal form of an event as 'Green' or 'Green is present', we might as well identify it as 'Grue is present'. Choosing 'Green' rather than 'Grue' as a translation of a certain native utterance could not be justified on the bases of the stimulation belonging to the stimulus meaning of green rather than grue. In this respect 'green' and 'grue' are equal. And even if we, contrary to Goodman, take 'grue' to be positional but 'green' not, the problem will not be solved. On that account, the sentence 'This is

---

[22] Strictly speaking, this observation categorical should be 'Whenever the natives would assent to 'Gavagai', I would assent to 'Rabbit''.

green' will have a higher degree of observationality than the sentence 'This is grue'. But a degree of observationality is not in any way given in the circumstances, it is neither part of the natives utterance nor part of the event. The question of whether the native's utterance should be translated as 'This is green' or 'This is grue' may, however, not survive all possible future observation, and the hypothesis that 'This is green' should be preferred to 'This is grue' is, therefore, not epistemically underdetermined, at least not in the strong sense. We might say that it is presently underdetermined. And since stimulus meaning is relative to time, preference of 'green' over 'grue' will not be justified by reference to stimulus meaning.

All the underdetermination occurring at the level of observation sentences will be carried over to other sentences of the language through the analytical hypotheses, since their acceptance depends, first and foremost, on their conformity to prior translation of occasion sentences. Thus, since scientific theory is underdetermined in the strong sense, so is translation. But, at any time, translation is also weakly underdetermined, or, as I say, presently underdetermined. This has been established without any mention of inscrutability of reference. Now we shall see how the inscrutability of reference affects indeterminacy of translation.

Indeterminacy of translation seems to follow directly from inscrutability of reference; if there is no way of telling whether the term 'gavagai' refers rather to rabbits than rabbit stages, any sentence containing 'gavagai' as a constituent will be translatable into sentence containing either 'rabbit' or 'rabbit stage' as a constituent. It seems that we have arrived at the same conclusion as before, just from a different starting point. But, have we really given an independent argument for the indeterminacy of translation?

The analytical hypotheses, whose role is both to distinguish certain words as terms and to fix the reference of those terms, are ultimately based on observation sentences. Now, what would justify our translation of the *term* 'gavagai' as 'rabbit' rather than 'rabbit stage'? If the analytical hypotheses giving 'rabbit' as a translation of 'gavagai' would conform to the already established translations of occasion sentences, but the alternative translation as 'rabbit stage' could not yield similar conformity *nor* conform to any other equally justifiable translation of occasion sentences, then 'gavagai' should certainly be

translated as 'rabbit' rather than as 'rabbit stage'. But this is impossible as we have already seen; the indeterminacy of translation is already present in the translation of observation sentences.

In the case of 'gavagai' inscrutability of reference seems then to do little more than reflect what has already been established by underdetermination of scientific theory. However, this is not the general situation. Given a particular translation of observation sentences, it is still possible to construct the analytical hypotheses in more than one way assigning radically different references to the same terms.

In "Ontological Relativity" Quine takes an example of a Japanese term that can either be taken as a individuative term for bovines or as a mass term for live cattle by appropriate adoption of the classifier that goes with it. (*OR* 35-37) What this shows, according to Quine, is that:

> Between the two accounts of Japanese classifiers there is no question of right and wrong. The one account makes for more efficient translation into idiomatic English; the other makes for more of a feeling for the Japanese idiom. Both fit all verbal behavior equally well. (*OR* 38)

The inscrutability of reference in the case of these Japanese classifiers arises given any particular translation of observation sentences and any particular scientific theory.

This point can also be illustrated in a more theoretical way by means of natural numbers.

> . . . the sentence '3 ∈ 5' goes into a true sentence of set theory under von Neumann's way of construing natural numbers, but goes into a false one under Zermelo's way.[23] (*ORIT* 183)

In other words, if we were to translate sentence about natural numbers, then, given a particular translation of observation sentences as well as set-theoretical sentences, we could construe the analytical hypotheses in such a different way that some sentences would come out true under one set of analytical hypotheses and false under some other.

---

[23] W.V. Quine, "On the Reasons for Indeterminacy of Translation", *The Journal of Philosophy*, Vol. 67 1970.

> . . . rival systems of analytical hypotheses can conform to all speech dispositions within each of the languages concerned and yet dictate, in countless cases, utterly disparate translations; not mere mutual paraphrases, but translations each of which would be excluded by the other system of translation. Two such translations might even be patently contrary in truth value, provided there is no stimulation that would encourage assent to either. (*WO* 73-74)

If all this is true, then given a certain translation of observation sentences there will still be room for incompatible translations of various standing sentences. And this incompatibility is not due to some empirical slack in our scientific theories, since it arises given a particular scientific theory. Thus, inscrutability of reference is not just a consequence of the indeterminacy of translation and the underdetermination of scientific theory. It is additional.

On top of all this there is another problem, namely when to say of two predicates which are alike in extension that they have the same meaning. In "Two Dogmas of Empiricism" Quine argues that even within a language, there is no way of giving a general account of the notions of synonymy and analyticity. And if these notions do not make sense as applied to terms of a single language, their application across languages will be all the more problematic. But inscrutability of reference shows that even before we can so much as formulate the problem of meaning or intension, we face a severe problem, namely that of identifying the extension of a term.

The empirical and the ontological aspect

Let's now return to the two forms of the indeterminacy thesis mentioned at the beginning of the chapter, the epistemological form and the ontological form. Translation is epistemically indeterminate if there can be incompatible translation manuals such that there is no evidence as to which is correct. It is ontologically indeterminate if there is no fact as to which manual is correct.

The question we want to answer is whether these two forms should be seen as two different theses or merely as two forms of the same thesis. The answer will depend on what facts are relevant to translation. If relevant facts are exhausted by behavioral facts, it does not make any difference whether we

allow for some additional facts. The additional facts will be irrelevant, and so will the difference between the two forms of the thesis. But if there are some additional facts that are, or could be relevant, then there will be a substantial difference between the two forms of the thesis.

By allowing irreducible semantic facts, the matter could be settled quite easily; it could be an irreducible semantic fact that 'gavagai' and 'rabbit' are coextensive terms. But this kind of an account of synonymy or coextensivity leaves us completely in the dark about the nature of semantic facts. Since Quine holds that all genuine facts are physical facts or supervene on physical facts and the question is whether he saw these two theses as fundamentally distinct or not, I shall only consider facts that can be accommodated within the limits of his physicalism.

When discussing the relation between indeterminacy of translation and underdetermination of scientific theory I said that the underdetermination of scientific theories would carry over to translation, and since the former is strongly underdetermined, so is the latter. The reason for this was that alternative translation could imply all the same observation categoricals. The question I am now going to explore is whether there could be any facts other than behavioral facts that constrain translation. This task has two parts; firstly, instead of restricting the observation sentences that form the observation categoricals implied by the translation manual to sentences about behavior, I allow some observation categoricals that mention physiological facts, and secondly, I shall consider the possibility of facts that are relevant to meaning, but that could not by mentioned in any observation categorical, because they are in principle not observable. The former questions the relevance of Quine's behaviorism within his general physicalistic framework, i.e. whether indeterminacy of translation follows from his restriction of possible evidence to behavioral evidence and could be resolved by bringing in some additional scientific evidence. The latter point is not thus concerned with the relevance of different kinds of evidence but with the question whiter there is a fact, accessible or not, as to which of alternative translation manuals are better.

What must be shown, in order to establish relevant difference between the epistemological and the ontological forms, is that there are non-observational physical facts in virtue of which rival translation manuals are right or wrong.

These facts could apply to any level of the translation and to any kind of expressions, i.e. not only to sentences but also to terms, connectives, etc.

If there are physical but non-observational facts that are constraining for translation, the correlations of sentences by a truly bilingual speaker should be within the limits of those constraints. The bilingual's translations should correspond to some physical condition in him which corresponded further to those constraints. But Quine explicitly denies this.

> . . . it is only to say that the bilingual has his own private semantic correlation – in effect his private implicit system of analytical hypotheses – and that it is somehow in his nerves. My point remains; for my point is then that another bilingual could have a semantic correlation incompatible with the first bilingual's without deviating from the first bilingual in his speech dispositions within either language, except in his dispositions to translate. (*WO* 74)

On Quine's view it seems to boil down to the same thing whether we restrict evidence for translation to behavioral facts alone, or whether we allow for other physical facts as well. The additional physical facts will not make any difference. That is why Quine says that in linguistics, one has no choice but to be a behaviorist.

> In psychology one may or may not be a behaviorist, but in linguistics one has no choice. Each of us learns his language by observing other people's verbal behavior and having his own faltering verbal behavior observed and reinforced or corrected by others . . . There is nothing in linguistic meaning, then, beyond what is to be gleaned from overt behavior in overt circumstances.[24]

But could we not say that the native refers actually to something in particular by the term 'gavagai', just as an English speaker refers to rabbits when he says 'rabbit'? Is it not a fact about the English speaker that he refers to some rabbit, and not to some undetached part of a rabbit? And if it is a fact about the English speaker, then there will also be a fact of the matter that the native refers to, say, a rabbit and not to some rabbit stage or some undetached rabbit part, even if that fact is inaccessible in principle to anyone but the native himself.

---

24 "Indeterminacy of Translation Again", *The Journal of Philosophy*, Vol. 84 1987, p. 5.

But another native may refer to rabbit stages, and not to rabbits by the same word. And if this difference is inaccessible in principle, and hence unresolvable, it is very doubtful whether we could say that the terms 'gavagai' and 'rabbit' are coextensive rather than 'gavagai' and 'rabbit stage'. And this last indeterminacy is not due to some hidden fact, like the fact as to whether someone intends to refer to a rabbit or a rabbit stage; this last indeterminacy is an ontological indeterminacy – not due to our limitations but to the fact that there are no facts of the matter.

Mental states and physical states

Is Quine's behaviorism a fundamental premiss on which everything else depends? I don't think it is, even though Quine has insisted on its importance rather than argued for it. To bring out the relevance of other facts, let's amend the story of the linguist a bit. Let's suppose that the linguist has access to all physical facts about the natives so that, for example, when a rabbit scurries by and a native says 'Gavagai', the linguist can observe what sensory receptors are triggered, and, in short, all the physical changes that take place. Will the linguist be able to determine whether 'Gavagai' should be translated as 'Rabbit' rather than 'Rabbit stage' given this additional information? I do not see how, since both translations are compatible with all the physical facts about the native. But what if we add all the physical facts about the native when he learned the sentence 'Gavagai'? It seems that all these facts will be equally compatible with the native having learned a sentence that should translate 'Rabbit' as one that should translate 'Rabbit stage'.

What would be needed for these physical facts to be constraining for translation? It is not enough that certain physical state be associated with certain expression, say 'Gavagai', since that will not give the linguist any clues to what English expression to associate with that state. What if some physical state of some native could be associated with a unique English expression as well? This might be possible if the linguist was bilingual. Perhaps some native's physical state would be associated with both 'Gavagai' and 'Rabbit' and not the rival alternatives. Would this determine translation? No, Quine would

insist, since some other bilingual native could have 'Rabbit stage' and not 'Rabbit' associated with the same physical state as 'Gavagai'.

Another proposal for physicalistic constraints on translation is to say that if the native's physical state when he would assent to 'Gavagai' is identical to the physical state of the linguist when he would assent to 'Rabbit', then 'Gavagai' should be translated as 'Rabbit'. But it does not make much sense to say of two individuals that they are in an identical physical state. It is hard to say of two cups when they are in the same physical state, let alone a cup and a saucer. And I do not even know what it could mean to say that the native and the linguist were in an identical physical state.

But what then about mental states? Could we not say that if the content of some mental state of the native is correctly expressed by 'Gavagai', and that state is identical to some mental state of the linguist that is correctly expressed by 'Rabbit', then 'Gavagai' should be translated as 'Rabbit'? Let's suppose that it is unproblematic to say that two individuals are in the same mental state. For two native English speakers we might say that they are in the same mental state if these states are both expressed by the same or synonymous expressions. But how can we extend this criterion to two language cases? We might try to say that if a native expresses his mental state with expression A in Jungle and the linguist expresses his mental state with an expression B in English, and A and B are synonymous, then the mental states of the native and the linguist are identical. But then the criterion of identity of mental states depends on the notion of synonymy, and we are back where we started. We wanted mental states to provide some constraints on translation, and now we end up with a criterion of identity of mental states that depends on translation.[25]

---

[25] See Dagfinn Føllesdal's paper "Indeterminacy and Mental States" in *Perspectives on Quine*, pp. 98-109.

Part II

Rules and Private Language

In his book *Wittgenstein on Rules and Private Language*[1] Saul Kripke lays down a paradox to the effect that it is impossible to follow rules or, for that matter, mean anything by anything. But he also offers a solution to this paradox, a sceptical solution, i.e. a solution which does not point out a fault in the reasoning that led to the paradox but which reviews the notion of following a rule. Kripke takes this to be the central theme in Wittgenstein's *Philosophical Investigations*, and his main argument against the possibility of a private language. In this chapter I will give an exposition of Kripke's work and not try to determine whether Kripke got Wittgenstein right.

In part I, where I spell out the sceptical paradox, and in part II, where I discuss Kripke's rejection of several proposals for a solution to the paradox, I will follow Kripke's text rather closely. In part III I will summarize the previous discussion and, following José Zalabardo,[2] consider two interpretations of Kripke's text. Finally, in part IV I will discuss Kripke's solution to the sceptical paradox.

I. The Rule Following Paradox

Kripke develops the paradox with respect to a mathematical problem, but the relevance of the paradox extends to all meaningful discourse. Suppose that I

---

1 Saul A. Kripke, *Wittgenstein on Rules and Private Language: An Elementary Exposition*, Oxford 1982. I will refer to this book with page numbers in brackets.

2 José L. Zalabardo, "Kripke's Normativity Argument", *Canadian Journal of Philosophy*, forthcoming.

am given the following problem to solve: '68 + 57'. This is meant to be just an ordinary mathematical problem, but one thing is crucial, namely that all my previous computations involved only numbers lower than 57. So, it is a new problem, and its arguments exceed all my previous computations. Since the natural numbers are infinite there will always be such a number, and for the present purposes I assume that 57 will do.

I perform the computation and obtain the answer '125'. But not only do I get this answer, I am *confident* that this is the right answer, both in the arithmetical sense that the sum of 68 and 57 is 125, and in the metalinguistic sense that '+' or 'plus', as I use the word, stands for a mathematical function which, when applied to the numbers which I call '68' and '57' gives the result that I call '125'. So, I am confident that the computation is correct, and that I use the expressions correctly.

But as uncontroversial as this may seem, it can be questioned. Suppose a bizarre sceptic confronts me and questions my certainty about the metalinguistic sense. The sceptic questions my use of the term 'plus' and maintains that, as I used it in the past, the answer to '68 + 57' should have been '5'. The sceptic points out that I have never seen this actual problem before, that all my previous computations involved numbers smaller than 57, and that in the past I gave myself only a finite number of examples instantiating the function denoted by 'plus'. So, perhaps in the past I used 'plus' and '+' to denote a different function, let's call it quus and symbolize it by '⊕'. It can be defined in the following way:

$$x \oplus y = x + y, \text{ if } x, y < 57$$
$$= 5 \text{ otherwise.}$$

Since the claim that I meant this function is consistent with my previous usage of 'plus' and '+', who is to say that this is not the function I meant by those expressions? Kripke admits that this is a crazy hypothesis, that it is most certainly false, but that it is not *a priori* impossible. But if it is false, shouldn't there be some fact about my past usage that can be cited to refute it?

The sceptic's challenge is twofold. He questions whether

(a) there is any fact which determines that I mean plus and not quus, and

(b) whether I have any reason to be so confident that now I should answer '125' rather than '5'.

A successful answer to the sceptic must, accordingly, satisfy two conditions.

(c) It must give an account of what fact it is that determines that I mean plus and not quus, and

(d) that fact must, in some sense, show how I am justified in giving the answer '125' to '68+57'.

In answering the sceptic, I need not restrict myself to one or another methodological principle, there are no behavioristic limitations to the facts that I can cite, and there are no epistemic barriers that can prevent me from finding the relevant fact. So, even if the initial formulation of the problem may suggest that it is merely an epistemological one, it is ontological; what the sceptic maintains is that there is no fact about me, be it accessible or not, which constitutes my meaning one thing rather than another. The sceptic does not hold that we are prevented from finding the meaning determining fact, but that there is nothing to find. (14 and 21)

It is important to notice that the sceptic is not questioning the correctness of my answer – he does not maintain that '5' is the right answer – but my confidence. If I arrived at my answer by a flipping of a coin, it might have been right or wrong, and if it is right then it is just as right as if it was the result of a careful computation. But, if I would merely have flipped a coin, I could not be *confident* that I had given the right answer, unlike when I *follow* a rule. And it is primarily this confidence that the sceptic questions. But if the sceptic is right, it seems as if my careful computations are no better than a flipping of a coin; it seems that my application of a rule is merely an unjustified stab in the dark – that I have no reason to be confident.[3]

---

[3] The arbitrariness which is a result of a flipping of a coin is not a result of a metalinguistic problem as is the case with 'plus' and 'quus'. When we flip a coin we may arrive at a particular answer, but the problem is that we are not guided by the rule and, hence, the relation between the arguments and the answer has nothing to do with the function in question. The sceptic, however,

If the sceptic is right and there is no fact as to whether I meant addition or quaddition by 'plus' in the past, then there can be no fact in the present either. But then there can be no fact that I mean one thing rather than another by anything I say – I will not even be able to formulate the sceptical paradox – and that is absurd.

II. Proposed Solutions to the Sceptical Paradox

Kripke discusses seven proposals for a straight solution to the paradox, all of which he rejects. Discussion of dispositional proposals takes most space but not because they are what he thinks Wittgenstein was primarily aiming at nor because they are more plausible than the other proposals. It may be argued that Wittgenstein was primarily concerned with, what might be called the classical empiricist picture, according to which expressions derive their meaning from mental images with which they are associated. But Wittgenstein's criticism of the classical empiricist picture is relatively well known, whereas his criticism of something along the dispositional proposal is not so clear or well received. But, be this as it may, Kripke's ideas and not Wittgenstein's are my main concern here.

Rule for a rule

A natural response to the sceptic is this: "I did not extrapolate the whole addition table from some finite number of examples, but I internalized instructions – I learned a rule – which determines how addition is to be continued for any arbitrary numbers." An explication of the rule or the instructions can be something along the following lines. Adding is really a way of counting, or can be spelled out in terms of counting. So, if you want to add x and y you can count out x marbles in one heap, then y marbles in another, you put the two heaps together and count the number of marbles in the union thus

---

does not question whether the answer I give is guided by a rule, but whether it is guided by the right rule. But if there is no way of making sense of 'the right rule', it will also be impossible to make sense of the notion of following a rule at all.

formed. In particular, if I have 68 marbles in one heap and 57 in another and I put them together, I will end up with one heap of 125 marbles. So, what I do when I add, whether it is a new addition problem or not, is that I proceed according to an *algorithm* for addition. This algorithm does not involve marbles or apples or any physical objects, but in principle there is no difference. What is important is that I count, and I can count on my fingers or in my mind.

Promising as this response may sound, the sceptic has an easy time undermining it. Sceptical doubts concerning my use of 'plus' will also apply to 'count'. If 'count', as I used the term in the past referred to the act of counting, then 'plus' must have stood for addition. But I applied 'count' only to finitely many cases, perhaps I never counted up to 57. And by 'count' I actually meant qount, where to qount a heap is to count it in the ordinary sense unless the heap was formed as the union of two heaps, one of which contained 57 or more items, in which case the answer is automatically '5'. At this point it is, of course, of no use to refer to some algorithm for counting, for the game can be played over again; a non-standard interpretation of one term calls for a non-standard interpretation of other terms.

The problem is not limited to 'plus' and 'count', it is perfectly general and well known from Wittgenstein's text. An interpretation cannot determine the meaning of a word or the application of rule. There are no bedrock rules such that once we arrive at those rules, the sceptical challenge can not be repeated.

What about the feeling I get when I have been struggling to learn the rule and finally 'get it' as it were. I get the feeling that 'now I can go on in the same way'. But this will not be of much help, for what 'going on in the same way' means, is determined by what we call the 'right' response. I learn the rule from finitely many cases and there will be infinitely many ways of proceeding. As before, if 'count' refers to the act of counting, then going on in the same way when adding means responding with '125' to '68+57'. But if by 'count' I meant qount, then, if I go on the same way as before, I should respond with '5' when queried about '68 + 57'.

Simple dispositional analysis

The dispositional proposal accepts the sceptic's claim as far as actual thoughts and actions go, but it maintains that there are dispositional facts that can differentiate between someone meaning plus or quus. (23) But what are these dispositional facts? Kripke considers two dispositional analyses, a simple analysis and a more complex one. Kripke finds them both failing, but not quite for the same reasons. I shall start considering the simple analysis and then go on to consider the more advanced one.

According to the simple analysis, to mean addition by 'plus' is to be disposed, when asked for the outcome of '$x + y$' to give the sum of x and y, so in particular, to respond with '125' when asked about '68 + 57'. So, even if I did not do this particular computation in the past, I *would* have responded with '125' had I been asked about '68 + 57'.

This crude analysis is unable to distinguish between what is a right and what is a wrong answer. Someone who is disposed to err systematically in his calculations, say, he answers '12' when queried about '5 + 8', will not be making a computational error on this account. He will just mean something different with '+' than most of us do. And if his concern is to be consistent with his past use, he should indeed not respond with '13' when queried about '5 + 8', since then he would not be consistent with his use of '+'. And this is absurd.

Another problem for this simple dispositional analysis is my finiteness and the infinity of the addition table.

> The dispositional theory attempts to avoid the problem of the finiteness of my actual past performance by appealing to a disposition. But in doing so, it ignores an obvious fact: not only my actual performance, but also the totality of my dispositions, is finite. It is not true, for example, that if queried about the sum of any two numbers, no matter how large, I will reply with their actual sum, for some pairs of numbers are simply too large for my mind – or my brain – to grasp. (26-27)

The infinity of the addition table seems fatal to the simple dispositional analysis. Within the framework of this dispositional analysis we cannot

distinguish between what might be called proper dispositions on the one hand and capacities on the other; it is limited to my actual capacities and we cannot ask what I would do if I would, say, live to be 200 years old.

Apart from the above problems, this proposal seems to be misplaced. It may well explain why I do now respond with '125', but it does not provide a *justification* for my response. And this holds true whatever my response is. What this crude dispositional analysis can provide is, at most, a response to the first part of the sceptics worries, i.e. it may satisfy condition (c) above in citing a candidate for a fact that constitutes my meaning something by an expression. But it does not provide anything in the way of a justification.

The simple dispositional analysis faces thus three fatal problems: (i) it cannot make a distinction between a right and a wrong answer, (ii) it cannot explain how a finite agent can master an infinite rule, and (iii) it does not provide anything in the way of a justification for my response.

Advanced dispositional analysis

The problem with the simple dispositional analysis is, one might say, that it is too simple. The right response should be something like this: "That I am disposed to respond with the sum when queried about '68 + 57', means that I would give the sum if queried under ideal conditions. And, in general, if I were queried about '$m + n$', for any m and n, then I would respond with the sum and not the quum, given ideal conditions." (28)

Before we go any further we should try to make the notion of ideal conditions clearer. We might call conditions ideal if there is no 'noise', if there are no disturbing factors. It may be true that I might not respond with '125' when queried about '65 + 57' – I might be upset or sick, but I would still be disposed to respond with the sum in the sense that in the absence of disturbing factors I would respond with '125'. The second kind of ideal conditions concern capacity. Suppose, for example, that I am given an addition problem involving numbers that are so big that I will run out of paper and pencils before I can complete the computations. In this case I will not respond with the sum, since I cannot finish my computations; not because of some 'noise' or some disturbing factors, but because I run out of resources. Ideal conditions would involve

enough paper and pencils as well as time and food etc. to complete the computations.[4]

An advantage of this analysis over the crude one, is that it can make sense of a right and a wrong answer as well as the infinity of one's dispositions.

According to the advanced dispositional analysis, an answer is wrong if it does not accord with what would be the answer if conditions were ideal. Unlike the crude analysis, it does not simply equate correctness and performance, but correctness and performance under ideal conditions. How far does this take us? This may, indeed, make sense of some distinction between right and wrong answers, but the plausibility of this approach depends heavily on how we can spell out the ideal conditions.

Kripke detects two problems for the spelling out of the ideal conditions. One concerns the distinction between right and wrong in instances where the numbers involved are small enough for us to deal with. The second problem concerns the infinity of the addition table. In addition to these problems, he considers the question whether dispositional facts can at all be justificatory of my responses.

Let's start with the first problem. I cannot simply say that ideal conditions are those under which I give the right answer, since that is flatly circular. What is needed is a specification of these conditions in non-semantic terms. It may be useful to consider ideal conditions for the application of color predicates (even if Kripke does not do so). We might say that ideal conditions for ascribing color predicates are that the object be in daylight, not far away, in the middle of the visual field and nothing relevant to color perception is affecting the subject's perception of the object. Calling these conditions *O*, we can give the following formulation of the dispositional analysis:

> For any subject *S*, object *x* and color-predicate *R*: if *O*, then *S* will judge *R* to apply to *x* only if *R* applies to *x*.[5]

---

[4] See C.B. Martin and John Heil, "Rules and Powers", *Philosophical Perspectives*, forthcoming. They provide a model for dispositions that can accommodate their infinitude as well as their normative character. This is accomplished by approaching dispositions from an ontological point of view and not from an epistemological perspective as is central to Kripke's Wittgenstein's objections.

[5] See for example Paul A. Boghossian, "The Rule-Following Considerations", *Mind*, Vol. 98, No. 392 1989, p. 538.

In other words, there must be conditions under which subjects are immune from error about judgements involving color predicates. But, it does not take long reflection to see that this will not provide us with the meaning of color predicates, since people do disagree about ascription of color predicates even when conditions are ideal. But even if this dispositional analysis is not able to fix the meaning of color predicates across the community, will it specify for each individual facts in virtue of which they are justified in ascribing color predicates as they do?

Suppose that I come across a color I have never seen before, say some shade between green and blue, and that ideal conditions hold. Suppose also that I am wondering whether to apply predicate A or B. Will my response be justified in terms of my dispositions to apply color predicates. Well, whatever I do will, by definition, be correct. But then, it makes no sense to talk about right or wrong – whatever strikes me as right will be right. And that is not what we wanted.

A further problem would be suggested by Nelson Goodman's riddle of induction. Who is to say that by 'green' I don't mean grue? No reference to ideal conditions will resolve the problem, since given ideal conditions, whatever they are, I can suddenly start to call the sky green and the grass blue. I need not have changed the meaning of the term; I may be going on in the same way, my 'same way' just not being the common one. (See e.g. p. 20 and note 46)

Let's leave color predicates and go back to Kripke's discussion. What the dispositionalist is suggesting is that the function someone means – the meaning of 'plus' – is to be read off from his dispositions. But if he makes systematic errors, it is hard to justify the attribution of the addition function to him. A certain unique function, call it skaddition, will correspond to the subjects dispositions, including his disposition to make errors. If the function is to be read off from his disposition then, Kripke maintains, the right function will be skaddition, not addition. (30)

There is a further complication here. Could we not object on the ground that we might make the person recognize his errors *as* errors, so that it is not his disposition to skadd which is the determining disposition, but some more basic disposition, say his disposition to count? In other words, can we not combine the first proposal with the dispositional proposal to get the desired result? I

think not. It is not any harder to imagine someone who has a disposition to count incorrectly. The problem boils down to this: It will be impossible to make sense of the right disposition without presupposing a notion of which function is meant.

How about the infinitude of the addition table and our own finiteness? The more advanced dispositional analysis can make sense of a distinction between dispositions and capacities. The ideal conditions would, perhaps, involve infinite capacities, so, for example, if I need paper and pencils to carry out the calculations, I would have an infinitude of both, as well as infinite time and other resources necessary to carry out my dispositions. This way of construing the ideal conditions allows the dispositional analysis to accommodate my finiteness and the infinitude of the addition table.

But how are we to determine what the disposition is, even if it is infinite? The idea behind the dispositional analysis is that the function is to be read off the disposition. The problem is that our only evidence is the manifestation of the disposition, and since the plus-hypothesis and the quus-hypothesis are equally compatible with this evidence, we cannot favor one over another without circularity. The problem is that there is no way of reading off the disposition to which function it corresponds.

Now let us put the above problems aside for a moment and suppose, for the sake of the argument, that there are dispositional facts that determine the satisfaction conditions of our predicates. Could these facts justify my application of a predicate? Kripke rejects the idea that dispositional facts construed along these lines can be justificatory of my response on the grounds that (i) I cannot know what the dispositions are, and (ii) I do not, as a matter of fact, apply predicates in light of my knowledge of the past physiology of my brain.

> Am I supposed to justify my present belief that I meant addition, not quaddition, and hence should answer '125', in terms of a *hypothesis* about my *past* dispositions? (Do I record and investigate the past physiology of my brain?) Why am I so sure that one particular hypothesis of this kind is correct, when all my past thoughts can be construed either so that I meant plus or so that I meant quus? (23)

The problem is that dispositional facts do not meet condition (d) above. They may show why I do respond in one way rather than some other but they fail to justify my responses.

The advantage of the advanced dispositional analysis over the simple one is then twofold: It can make sense of some distinction between right and wrong by means of ideal conditions, and it can make sense of infinite dispositions by means of a distinction between dispositions and capacities. But, nevertheless, it runs short of providing a satisfactory answer to the sceptic for two reasons. Firstly, as Kripke puts it, "the idealized dispositions are determinate only because it is already settled which function I meant" (28), and secondly, whatever the dispositional facts are they do not provide me with a justification for my response.

Machines

After rejecting the dispositional proposals, Kripke goes on to consider the possibility of building the function into a machine. The machine response goes something like this: "Can we not define the right function by building it into a machine?" This, Kripke claims, is in a way just a variant of the dispositional proposal since we can be viewed as machines. Just as the function was to be read off from our dispositions, so will it be read of from the operation of the machine, and ideal conditions will be conditions under which the machine works properly.

One thing must be made clear at the outset. The term 'machine' is ambiguous. Sometimes it refers to programs, as is usually the case when we talk about Turing-machines, and sometimes it refers to physical machines, like a clock or a computer. If we are talking about programs, then we are back to the first proposal, namely interpreting rules with rules. The program consists of instructions that I am free to interpret any way I like. What we need are physical machines. But where does that take us? We will need (i) a criterion to determine when the machine works properly and when not, (ii) a way of determining the dispositions of the machine for cases that exceed its actual capacity, (iii) some means of interpreting the output from the machine.

The last point brings us, in a way, back to the first proposal. A machine which I invent to interpret some symbol cannot provide me with anything but another symbol. But, waiving this problem, we are left with (i) and (ii), which are parallel to the problems which the dispositional proposals proved unable to answer. These problems are, I think, clearer in the case of a machine than in the case of an individual. The philosophical relevance is the same.

Whether a machine malfunctions, say, whether an actual Turing-machine works properly, is not a property of the machine itself as a physical object, but is well defined only in terms of its program as intended by its designer (35). But if the program is the criterion for correctness, then *it* and not the machine, would be definitive for the function.

How then about the finitude of the machine? Obviously, the machine is a finite object, but could it have infinite dispositions? Consider a Turing-machine which is composed of devices to scan, write, move left and right and to change states. This I shall refer to as the basic part of the machine. In addition to the basic part, it has a string on which it writes '1' or '0'. As a matter of fact, the string will be finitely long. We could say that the basic part of the machine has dispositions to do more than the string allows it to do. This we say simply because we can say that if the string was longer, it could do computations that involved higher numbers than it can actually do. We could even say that for any natural numbers, we could imagine the string being long enough for the machine to be able to compute them. In this sense we can say that the basic part of the machine has infinite dispositions. But even if we can say that the machine has infinite dispositions in virtue of its basic part having infinite dispositions, we are still left with the problem of determining what those dispositions are. And since the actual machine is finite, we are left with finitely many computations and, hence, we cannot distinguish between infinitely many possible hypothesis in virtue of its actual performance. The only way to determine which disposition is built into the machine is to look at the program – and that renders the actual machine irrelevant.

Simplicity

"Is the hypothesis that I meant plus rather than quus not to be preferred on the grounds that it is simpler? Simplicity considerations play an important role both in sciences and every day life. The present problem is this: How are we to determine which of two empirically equivalent hypothesis we should adopt. Could we say: Let's adopt the simpler one, namely the plus-hypothesis?"

It is true enough that simplicity considerations are important both in sciences and every day life but, this response is misguided from the start. What simplicity considerations can do is to help us to decide between competing hypotheses, but they cannot tell us what these hypotheses are. The problem in deciding whether to adopt the plus-hypothesis or the quus-hypothesis is that we do not know what could make them right – we are at a loss about what they are – and no simplicity consideration can help us determine that. The problem is not that we have found a certain fact and are wondering whether it supports the plus-hypothesis or the quus-hypothesis, but that we have the hypotheses but the fact is missing.

When scientists are faced with two or more empirically equivalent but incompatible hypotheses, say, in how they depict the structure of the world, simplicity considerations may help. Suppose, for example, that we have two incompatible theories about electrons. All available evidence is the behavior of gross objects, and our access to electrons is only indirect. In this kind of circumstances simplicity might help, but it would only help us finite beings. If God were to determine which was the right hypothesis, he would not have any use for simplicity considerations, he would simply see all relevant facts about electrons directly, and, hence, see which hypothesis was right if either was.

If hypotheses do not depict anything – if they are not about factual matters – then one cannot be favored over another as the *right* hypothesis on grounds of simplicity. It may still be favored, say, because it makes computations simpler, but then it is not taken to be right or wrong, it is just taken to be an instrument and should, perhaps, not be counted as hypothesis at all but some methodological instrument.

The sceptic holds that there is no fact as to what I mean, and, hence, there is nothing to be right or wrong about. He maintains that even if God could look into my mind and observe all available facts, he could not determine whether I meant plus or quus. And, hence, no simplicity considerations can help either.

The classical empiricist picture

"Since I know immediately and with fair certainty that I mean plus, it must be that I know this in the same way as I know that I have a headache or a blue afterimage etc. Just as it does not make sense to say: "I thought for a moment that I was in pain, but I wasn't", neither does it make sense to say: "I thought for a moment that I meant addition by 'plus', but I didn't". So, perhaps, meaning addition has its own irreducible quality, and the fact that I mean addition by 'plus' is to be identified with my possession of an experience of this quality." (41)

But this approach seems to fall short of any satisfactory answer to the sceptic. Suppose, for example, that I do indeed have a very special feeling when I think of the '+' sign or hear the word 'plus'. How could this help me figure out whether '125' or '5' is the right answer to '68 + 57'? Whatever the feeling is, it will be compatible with my either meaning addition or quaddition by 'plus'.

The classical empiricist picture was one of Wittgenstein's main targets. But it does not do justice to this view to discuss it only in relation with mathematical functions. On that front, it may not seem so promising. What the classical empiricist picture claims is that words get their meaning, or their meaning is determined, by way of association with mental images; the word 'red' would then get its meaning, or its meaning for me would be determined, by my association of a red mental image with the word, similarly for the word 'cube'. And isn't this promising?

Kripke cites two kind of problems with this view. Firstly, many of us use words such as 'cube' without any such image coming to mind. But, secondly, even for someone who does associate a mental image of a cube with the word 'cube', that image does not determine when or how that picture fits or fails to fit the word 'cube'. Perhaps it fails to fit somewhat in the following way: "If that picture comes to me and I point to a triangular prism for example and say it's a

cube, then this use of the word doesn't fit the picture". (42) But we can easily imagine a method of projection which makes the mental picture fit the prism. What seems to be wanted in addition to the mental picture is a determinate way of *using* the picture. But the picture can be used in different ways; it does not suggest a single determinate use.[6] In order to overcome this problem, one might say that it is not only the picture that comes to mind but some method of projection as well. If the method of projection is in the form of a mental picture similar to my mental picture of the cube, then we will still face the same problem. That picture can be interpreted in non-standard ways. If it is in the form of a set of rules, the mental picture of the cube is perhaps irrelevant. In either case, we are back to the first proposal – we are trying to determine the application of a rule by referring to other rules, and as we have already seen, that does not solve the problem.

On top of this, it seems very dubious to associate a special qualitative character to the experience of meaning addition by 'plus' (whatever that experience is). (44) What could this qualitative character be? Perhaps some feeling that I had when I first mastered addition, when I suddenly felt that I had grasped the rule for addition. This seems dubious enough, but, suppose for the sake of the argument that there was such a moment, that I had such a feeling at that moment, and that feeling came back whenever I thought of the '+' symbol or the word 'plus'. The problem still arises, for how am I to determine that it was actually the addition rule that I got and not the quaddition rule? Here, there will be no use in referring to some feeling of being able to go on the same way, for what counts as going on the same way depends on what counts as the right answer. (19n)

Kripke is not denying that there may be a certain 'feel' to a meaningful use of words, but that there is a particular qualitative experience that we have when and only when we use a word with a certain meaning. Whatever role feelings play, they do not determine the meaning of words. (See note 29 p. 46)

So, the classical empiricist picture fails to provide an answer to the sceptic since any qualitative experience or mental image which may be associated with

---

6 I do not even know if it makes sense to talk about 'a single determinate use' in this context. Is the application of the picture in relation to new objects the same use or a different use? It is difficult to define 'the same use' in a non-circular way.

the use of a word, can be compatible with the word being used in a non-standard way, so, in particular, whatever feeling or mental image I associate with the word 'plus', it will be compatible with my either meaning plus or quus. If we try to supplement the feeling with, say, how we first learned the rule in question or our mental images with directions for use, we will just run into the first objection; these will only be rules for interpreting rules.

Platonism

The Platonist holds that mathematical entities have independent existence, and hence, there is no problem how the addition function can contain within it all its instances. It is simply in the nature of this mathematical object which is infinite. To grasp the addition rule is simply to have an idea, or some mental representation, of this mathematical object.

Kripke's discussion of this proposal is rather short, since its shortcomings should be obvious. The problem will simply be this: What determines that the idea, or the representation, is a representation of the addition rule and not the quaddition rule. Even if the mathematical object is infinite we are only familiar with a finite part of it and since this finite part with which we are familiar could belong to an indefinite number of mathematical objects, we can not favor one over another without circularity. The problem is quite general; even if the mathematical objects need no interpretation there will always be a need for a mental entity about which the sceptical problem can be raised.

III. Summary and Interpretation

Now let's sum up what we have established so far. The starting point was the criterion for any fact to be what we might call a meaning determining fact. A fact is a meaning determining fact if and only if it satisfies two conditions: (i) it determines that I mean whatever I mean, say meaning addition by 'plus', and (ii) it must, in some sense, show how I am justified in giving my actual responses, for example replying '125' when queried about '68 + 57'.

But all attempts at finding such a fact have failed and we seem forced to conclude that there is no fact about me whatsoever that determines that I mean

one thing rather than another. Now, the sceptic claims that this is not due to some empirical slack – it is not because we cannot get to these facts that we haven't been able to find them – but that there is no such fact at all. The problem is an ontological one.[7]

The proposed solutions to the sceptical problem which Kripke considers fail mainly for two reasons: they fail to fulfill the justificatory role or they fail to ground the original rule. The dispositional proposals fail for the former reason. They do point out a fact, but that fact cannot be justificatory in the required sense. The classical empiricist picture fails to ground the performance. It points out a fact – a mental image – but that fact can only serve as an interpretation of the original rule and, hence, runs up against Wittgenstein's well known thesis that an interpretation cannot determine the application of a rule. Platonism faces the same problem as the classical empiricist picture. And it should be clear that it will be of no help to point to some theory about the nature of mathematics since the sceptical problem has nothing to do with mathematics as such, it is not scepticism about mathematics. One can be a Platonist or structuralist about mathematics and Kripke's sceptical paradox remains the same. The simplicity objection is totally misdirected since it does not point out any fact whatsoever.

Normativity

So what is the kernel of Kripke's argument for this paradoxical conclusion? What Kripke is often credited with is an argument for the normativity of meaning – normativity which cannot be captured by any descriptive account. The first premiss in that argument would be something like this:[8]

> (A) The fact that determines the satisfaction conditions of a predicate licenses evaluative claims about ascription of the predicate.

---

7 Kripke sets the problem up as if it is an empirical problem, but, as he makes clear, the problem is not just empirical but ontological. It is, however, not clear whether the distinction between what would be an ontological problem and an empirical problem is at all relevant, since any meaning determining fact must be accessible if it is to fulfill the justificatory role.

8 In the following discussion I follow José L. Zalabardo.

This seems to be a trivial claim: If 'red' applies to all and only red things, then I will use the predicate 'is red' *correctly* if I ascribe it to red things, and I will use it *incorrectly* if I ascribe it to things which are not red. That the predicate has a definite meaning gives rise to normative claims about my use of the predicate; we get a criterion for a right and a wrong use.

The argument against the dispositional accounts would then proceed something like this:

(B) Dispositional facts are descriptive.

(C) Descriptive facts cannot license evaluative claims.

From these three premisses it follows directly that no dispositional fact can determine the satisfaction conditions of a predicate.

There are, however, several reasons not to read Kripke along these lines. Firstly, Kripke does not try to argue for premiss (C) which is not trivial.[9] By evaluative claims I mean claims such as "You ought to answer '125' and not '5' when asked about '68 + 57'". It was perhaps one of Hume's lessons that an ought-statement could not be based on is-statements, and that it would be doomed to fail to look for a descriptive fact that determined that I ought to do one thing rather than another. But even if Hume had an argument in support of this thesis, Kripke does not offer any, nor does he refer to Hume on this point. And moreover, it is, in a way, quite easy to derive an ought-statement from is-statements. In a game, say soccer, we can say that if the players want to play soccer, then they must do so and so, i.e. they must follow the rules of the game since that is part of what it means to play soccer. They cannot, for example, play the ball with their hands. The ought-statement is thus the consequent of a conditional: If you want to play soccer, then you ought to do so and so on a given occasion. Similarly for addition: If you mean addition by 'plus' then you ought to reply '125' to '68 + 57'. An advocate of the standard normativity argument would, perhaps, object to this conditional on the grounds that it does not describe a fact – that it is not a proper is-statement. But

---

9 See José L. Zalabardo pp. 5-11.

Kripke's sceptic does not question this conditional, he only questions the certainty with which its antecedent is asserted.

That the above reading is faulty should also become clear if we consider the difference between someone who obtains an answer through careful computations and someone who just calls out numbers at random. The former may be confident that he gave the right answer while the latter will not. But their answer may be equally right or wrong; a fact which determines the application of the rule will license all the same evaluative claims about their answers, but the sceptic's challenge does not affect the one who just calls numbers out at random.

The sceptic's question was twofold: (i) "Is there a fact which determines that you mean plus and not quus?", and (ii) "How can you be so certain that '125' is the right answer?" It may well be a fact that '125' is the right answer without that fact supporting my certainty. The advanced dispositional proposal did, for example, point out a fact in virtue of which we could make sense of a distinction between a 'right' and a 'wrong' answer. What it could not provide was a justification for giving one answer rather than another – it could not provide a fact in virtue of which I could be confident that my answer was right.

What it comes down to is this: We can imagine a situation where we can make evaluative claims about predicate application, without the sceptic's challenge having any force. It is not just that we apply predicates (or follow rules) correctly or incorrectly, but that we are, at least sometimes, *confident* that we are applying predicates correctly. So, it appears that what the sceptic is after is something more than just satisfaction facts.

Justification

What is Kripke's argument if it is not the normative argument laid down above? What we need is a fact that can support my confidence and such a fact must, according to Kripke, tell me somehow what I ought to do in each new instance. Without such a fact we cannot distinguish between someone who gives a justified response and someone who just calls out numbers at random. This is what Kripke is getting at towards the end of his discussion of the simple dispositional analysis.

> So it does seem that a dispositional account misconceives the sceptic's problem – to find a past fact that *justifies* my present response. As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition on such a candidate . . . that it should *tell* me what I ought to do in each new instance. (24)

Before we go any further, let us distinguish between *satisfaction facts* and *meaning determining facts*. Let us call any fact that determines the extension of a predicate a satisfaction fact. A satisfaction fact would determine how I ought to apply a predicate provided that I want to apply it correctly.[10] Let us, on the other hand, call a meaning determining fact a fact that determines what I mean by a given predicate. A basic condition on a meaning determining fact is, according to Kripke, that it justifies my present response and it does that by telling me what I ought to do in each new instance.

Now we might ask whether satisfaction facts are identical with meaning determining facts? Does the fact that the predicate 'red' applies to all and only red things justify my application of the predicate by telling me how I ought to apply it? Obviously not since despite the fact that 'red' applies to all and only red things, I can use it totally arbitrarily. But can the fact that tells me how I should apply the predicate in each new instance be independent of the satisfaction fact? It seems to be necessary that the meaning determining facts be somehow related to the satisfaction facts. The question remains what exactly this relation should be.

To begin with let us note that we want to say that some procedures for answering questions like 'What is the right answer to 68 + 57?' are justified while others are not. A proper procedure would presumably involve certain computations while calling out numbers at random would be an unjustified procedure. But we still need an account of *why* certain procedures are justified; we need an *explanation* of why certain procedures are justified not merely a list of justified procedures.

---

[10] Someone might object to this formulation on the grounds that the extension of various predicates, even all predicates, is not well defined. But for the sake of the argument, let us suppose that good sense can be made of this notion of satisfaction facts.

Kripke does not give any conclusive account of this relation. When discussing the proposal that I should answer '125' because that is what I was disposed to do in the past and I want to be consistent with my past usage, he replies:

> How does any of this indicate that – now *or* in the past – '125' was an answer *justified* in terms of instructions I gave myself, rather than a mere jack-in-the-box unjustified and arbitrary response? Am I supposed to justify my present belief that I meant addition, not quaddition, and hence should answer '125', in terms of a *hypothesis* about my *past* dispositions? (Do I record and investigate the past physiology of my brain?) (23)

What Kripke is maintaining is that if the answers I should give were determined by a certain dispositional fact, then a justified procedure for arriving at an answer would have to draw on this fact. So, according to Kripke, it will be a minimal requirement for any justified procedure for predicate application that it involves a conscious engagement with the satisfaction facts.

Now we are in a position to summarize Kripke's arguments beginning with these two premisses. I will refer to this argument as the justification argument.

(A) If my application of a predicate is justified then the procedure that I use for deciding whether to apply the predicate must be justified.

(B) If my decision procedure is justified then it involves conscious engagement with the satisfaction fact.

The justification argument differs from the standard normativity argument in an important respect: The fact in question must not only license evaluative claims about my use of the predicate but it must *justify* my decision to use the predicate in one way rather than some other. This fact must not only determine which objects fall under the predicate but it must *tell* me to which objects *I should* apply the predicate.

Against the advanced dispositional proposal Kripke argues that (i) dispositional facts are not able to justify my decision procedure for predicate application whether or not they can license evaluative claims about my application of predicates, and (ii) that they fail because they do not guide me –

do not give me a reason to act in one way rather than another – whether or not they can provide facts which determine the extension of the predicates of the language.

When arguing against the advanced dispositional proposal, we need only the following additional premiss:

(C) I decide whether to apply predicates to objects in ignorance of whether I would do so under ideal conditions.

Premise (C) differs from the previous two in that it is purely empirical. As a matter of fact, we might say, we do not contemplate how we would apply predicates under ideal conditions, nor do we even consider what would be ideal conditions for each predicate before we use it.

From premiss (B) and (C) we can conclude the following:

(D) If facts about how I would apply predicates under ideal conditions were identical to satisfaction facts, then my decision procedures would be unjustified.

And from (D) and (A) we can conclude the anti-dispositional conclusion:

(E) If facts about how I would apply predicates under ideal conditions were identical to satisfaction facts then my predicate application would be unjustified.

The classical empiricist picture maintained that we determined how to apply a predicate by bringing to mind a mental image. So, in particular, that we determined how to apply a predicate like 'red' by bringing to mind a red mental image. But, Kripke argues, the mental image does not as such determine how it should be applied.

Suppose that I do indeed always think of a certain image whenever I use a particular word, say, I have a red mental image whenever I use the word 'red'. Will this fact determine to which objects I should apply the predicate? Well, the mental image is just another object, and it seems that I would already need a fact that determined whether the word 'red' applied to my mental image. So, I need something else to justify my application of 'red' to my mental image, perhaps another mental image, and another . . . The details of how the mental

image is supposed to be of help in determining whether an object falls under a predicate do not matter. What is important is that it purports to justify my predicate application with another predicate application, but that only leads to an infinite regress.

So, when arguing against the classical empiricist picture we only need the following additional premiss:

(F) A mental image does not determine to which objects I ought to apply a predicate.

From this premiss and premisses (A) and (B) we can conclude that if a mental image would constitute the satisfaction facts for a predicate, then my predicate application would be unjustified.

But how exactly does this lead to a paradox? The paradox arises in connection with sentences like "I mean addition by 'plus' ". The justification argument states that if I can justifiably assert this then I must have a conscious engagement with a fact that determines that I mean addition by plus. But we are at a loss to find such a fact – there seems to be no such fact about me which determines that I mean addition by 'plus' – and hence we cannot justifiably assert that I mean addition by 'plus'. But not only can we not assert this justifiably, but we cannot justifiably assert any meaning attributing statement. Now from this it follows that I cannot be certain that I am using any of my words correctly, since if I could be certain, I should be able to assert it justifiably that I am using it correctly. And if no certainty is possible, it seems that our language use is nothing but a guessing game.

IV. The Solution to the Paradox

Before we go any further we should note that the paradox is not a contradiction. But we only have to add what seems to be an uncontroversial premiss in order to derive a contradiction. The sceptical paradox is the following:

P1 There is no fact about me that justifies the attribution of one meaning rather than some other to whatever I say.

The picture to be replaced maintains that a declarative sentence gets its meaning by virtue of its *truth conditions*. (72) This is a correspondence theory of meaning and, perhaps, the most natural way of explaining meaning. But as natural as this approach may seem, it is perhaps the central target of Kripke's Wittgenstein. He wants to show that truth conditions do not capture the essence of meaningful speech. Instead of truth conditions we should look for *assertability conditions*. (73-74) Instead of asking what would make a sentence true we should ask the two following questions: "Under what conditions may this form of words be appropriately asserted (or denied)?" and "What is the role, and the utility, in our lives of our practice of asserting (or denying) the form of words under these conditions?"[11]

Before we try to spell out the assertability conditions, we should make clear what is involved in giving up premiss P2.

One way of giving up premiss P2 is to adopt non-factualism about meaning. A non-factualism about meaning is analogous to emotivism in ethics where sentences of the form "X is good" are not taken to state a fact – namely the fact that X is good – but as having a logical form similar to "Hurrah for X!". On this reading, we could adopt the following semantic principle for statements such as the one above:

(1) All sentences of the form 'X is good' are not truth-conditional.

If non-factualism about meaning is the right interpretation, Kripke's Wittgenstein should perhaps adopt the following semantic principle for meaning attributing sentences:

(2) For all S and p: $\ulcorner S$ means that $p\urcorner$ is not truth-conditional.[12]

---

[11] Kripke's Wittgenstein proposes a more radical view than this passage suggests since the expression 'assertability condition' indicates a primacy to declarative sentences. Such a primacy is, in fact, what Wittgenstein was up against, as Kripke notes. But for the present purpose we can, as Kripke does, focus on a range of cases where we attribute certain intention to language users.

[12] Here 'S' ranges over names of sentences and 'p' ranges over sentences.

But this is highly counter-intuitive. If I mean addition by 'plus', then it is true that I mean addition by plus, and the reason why it is true is that the possible fact that I mean addition by 'plus' obtains.

But, moreover, if the aim of Kripke's Wittgenstein was just to show that meaning-ascribing sentences should be analyzed along the lines of, say, performatives or imperatives, then he would not, no matter how successful in his task, have shown a fundamental defect in the truth-conditional picture. His conclusion would have no direct consequences for sensation language and, hence, have little relevance for the private language argument. But Kripke's Wittgenstein was not just trying to limit the scope of the truth-conditional picture but to show that the picture is wrong in a fundamental way.

But what alternatives are there apart from non-factualism on the one hand, and premiss P2 on the other? And how are we to understand Kripke's claim that there is no fact about me that I mean one thing rather than another?

Before we try to answer these questions we should get a bit clearer about what a truth-conditional theory of meaning consists in. In the *Tractatus* Wittgenstein put forward such a theory of meaning.

> The simplest, most basic idea of the *Tractatus* can hardly be dismissed: a declarative sentence gets its meaning by virtue of its truth conditions, by virtue of its correspondence to facts that must obtain if it is true. For example, "the cat is on the mat" is understood by those speakers who realize that it is true if and only if a certain cat is on a certain mat; it is false otherwise. The presence of the cat on the mat is a fact or condition-in-the-world that would make the sentence true (express truth) if it obtained. (72-73)

This formulation of the truth conditional picture of the *Tractatus* may seem hard to resist, but, as Kripke notes:

> Nonetheless, as Dummet says, "the *Investigations* contains implicitly a rejection of the classical (realist) Frege-*Tractatus* view that the general form of explanation of meaning is a statement of the truth conditions". In the place of this view, Wittgenstein proposes an alternative rough general picture. (72-73)

The 'alternative rough general picture' of the *Investigations* did not, according to Kripke, imply that declarative sentences do not have truth conditions, but

that the truth conditions are not *explanatory* of meaning. The difference between a sentence having truth conditions and the truth conditions being explanatory of the meaning of the sentence is intimately related to the difference between a fact that licenses evaluative claims about predicate application and a fact that justifies its application. A fact that *explains* the meaning of a predicate would not only determine its extension but tell me somehow to which objects it applied, and, hence, satisfy the basic conditions of a meaning determining fact specified above. But we can think of a situation where a fact determines the extension of a predicate and hence determines whether sentences in which it occurs are true or false without it explaining the meaning of the predicate.

The alternative picture of the *Investigations* did not, then, aim at undermining the two following semantic principles:

(a) A predicate is meaningful iff it stands for a property.

(b) A predicate applies to an object iff the object has the property for which the predicate stands.

These principles are general principles about predicate meaningfulness and application but are not explanatory of *why* predicates have the meaning they have. Principle (a) specifies minimal condition which all predicates must satisfy in order to be meaningful,[13] principle (b) is an *interpretive* principle for predicate application, it specifies minimal conditions for correct predicate application. An instance of principle (b) could be something along the following lines:

(b') The predicate 'red' applies to an object iff it is red.

Principle (b') does license evaluative claims about the predicate 'red', and if we could specify the property for which the predicate 'red' stands in a noncircular way, we could define the extension of the predicate.

---

[13] There is a complication here since some predicates such as "does not apply to itself" are paradoxical. These predicates are not meaningless, but it is a question whether they stand for a property, or if so, what kind of a property they stand for. In the present context I will simply ignore this complication.

What must be rejected according to the alternative picture of the *Investigations* are the corresponding explanatory principles.

(c) A speaker understands a predicate *because* he has knowledge of the property for which the predicate stands.

(d) A predicate is meaningful *because* speakers have a conceptually prior grasp of a certain property and have formed the intention to associate the predicate with that property.

An explanation of the meaningfulness of a particular predicate along the lines of principles (c) and (d) will be a reductive explanation; the meaningfulness is reduced to (i) the property for which the predicate stands, and (ii) the intention of the speakers to use the predicate to stand for that property.

What the sceptic questioned in the original formulation of the paradox was the role of the knowledge of properties in principle (c) and the intention in principle (d). How do I know that in the past I intended to use 'plus' to refer to addition rather than quaddition? If we cannot answer this question, then, according to the truth-conditional picture, our explanation of the meaningfulness of the term 'plus' is incomplete. But how could I form the relevant intention? Did I have some direct, nonlinguistic nonrepresentational grasp of the addition function in virtue of which I could form the intention to use a certain expression to refer to it? Kripke's Wittgenstein rejects this idea, and suggests in its place a picture according to which my grasp of the function coincides with, rather than precedes, my understanding of a term for it.[14]

What is given up by giving up the truth-conditional picture is then the idea that corresponding facts explain the meaningfulness of sentences. But once we give up this explanatory role of corresponding facts – once we stop looking for the corresponding fact that makes a sentence meaningful – we give up the demand for a reductive explanation of meaningful speech.

> All that is needed to legitimize assertions that someone means something is that there be roughly specifiable circumstances under

---

[14] See Scott Soames, "The Sceptical Solution Without the Sceptical Paradox", unpublished, pp. 15-16.

which they are legitima
them under such cond
that 'facts correspond' t

Under the influence of t
that would justify an asserti
'plus'", to make the sentence
that could justify the assert
was to conclude that meanin
that is to go too far. What t
fact about Jones's mental or
means addition by 'plus', no
do the job.

But what is our evidence t
our justification for asserting
by 'plus'"? Instead of lookin
be it a disposition or a men
asserting or denying the sent
judge whether Jones meets
since these communal standa
accept the sentence "Jones
conscious engagement with
shift from the truth-conditio
thus a shift from the internal

What the truth-condition
would have to be preceded b
in order to be a proper use
associated with the right int
object would be so because it
property for which the pred
justification argument state.

According to the truth-c
following:

(4) 'P' applies to all and

where 'P' stands for a certain property, would underlie our linguistic competence. So, in particular, we would have conscious engagement with redness, we would understand that the predicate 'red' applies to all and only red things, and this would *guide* us in our application of the predicate in new cases.

In the new picture this is turned upside-down. It is our linguistic competence that explains our understanding of principles like (4). It is not that in the case of predicate application we have conscious engagement with properties, form tentative hypotheses similar to (4) about the relation between the predicates of the language and these properties and then decide which way we are going to apply the predicates. Rather, linguistic competence is a matter of satisfying conventional non-semantic standards regarding language use.

The new picture

According to the new picture each declarative sentence has associated with it both truth conditions and assertability conditions. Truth conditions are interpretive of these sentences by specifying the conditions under which they are true. Assertability conditions are, on the other hand, conditions under which these sentences can justifiably be asserted and do not determine their meaning.

This has a parallel in the case of predicate application. Instead of truth conditions we have extensions which determine to which objects the predicate applies and will, thus, license evaluative claims about its application such as "'Red' applies correctly to all and only red things". But we also have application conditions which are conditions under which predicates can justifiably be applied to objects. And just as the assertability conditions do not determine the meaning of declarative sentences, application conditions do not determine the extensions of predicates.

But what are assertability conditions? Kripke emphasizes that we should not argue *a priori* what conditions should license particular assertions and what role such assertions should have but *look and see* what circumstances *actually* license particular assertions and what role they *actually* play. Sticking to the

original formulation of the paradox, we should look and see what circumstances actually license claims such as "Jones means addition by 'plus'". Now, the situation would be something like this: Jones is asked about '68 + 57' and he replies '125' and he does so unhesitatingly and without any thought of the possibility that a quus-like rule might have been appropriate. He replies *with confidence* but *without justification*. If asked why he gave that particular reply, he would probably explain that he added 7 and 8, put down '5' and carried '1', etc. But this kind of an explanation can only suffice to resolve some mathematical doubts that we might have and is *not* an answer to the sceptic's metalinguistic doubts, as it will only bring us to new rules which he followed with the same confidence and ultimately he gets to a point where no justification is possible. This, Kripke concedes, is the central point of the sceptical argument.

> The entire point of the sceptical argument is that ultimately we reach a level where we act without any reason in terms of which we can justify our action. We act unhesitatingly but *blindly*. (87)

What can we say about an individual in isolation with respect to this situation? All we have to go on, Kripke claims, is the individual's psychological state and his external behavior. At each application of the rule he acts with confidence but without justification, just as he should. But there is no room to question his performance and there is no way of distinguishing between a right and a wrong answer. As Kripke writes: "There are no circumstances under which we can say that, even if he inclines to say '125' he *should* have said '5', or *vice versa*." (88) We cannot say whether he is consistent with his past intentions or following the right rule since any reply will be just as good as any other. So, considering a person in isolation, all we can say is that he can apply the rule the way it strikes him – he can react to the signs just as he is disposed to.

This is obviously not what we mean by saying that someone is *following* a rule. We seem to have fallen into the ditch of the simple dispositional proposal which we found incapable of justifying assertions such as the one in question.

If we, on the other hand, consider an individual in a community, we are able to make sense of the necessary normativity; in a community 'going on the

anything else". (91) The demand for such a justification is mistaken; all it can provide is a rule to interpret a rule.

Jones's inclinations are, of course, not enough to justify claims such as "Jones means additions by 'plus'". What is needed is a third person perspective. Others need not accept Jones's authority on these matters; they need neither agree with his general inclination that he can 'go on' nor do they have to accept his particular responses in particular cases. But if Jones does reply with the sum and not the quum to most addition problems and his behavior seems sensible, then he can, justifiably, be judged to mean addition by 'plus' and be accepted into the community of adders.

But how is one accepted into a speech community, say the community of adders? The conditional assertion that if Jones is following the addition rule then he should reply with the sum, plays a central role. So, in particular, if he means addition by plus he should reply '125' to '68 + 57'. But, as should be obvious, Jones can satisfy any finite conjunction of such particular conditionals without actually following the addition rule but some other quus-like rule. That is why the priorities must be reversed. In practice, Jones would be judged by the community to mean addition by 'plus' if he does not give it a reason to the contrary. So, instead of saying "If Jones means addition by 'plus' then he should reply '125' to '68 + 57'", the emphasis is on the contrapositive form, i.e. "If Jones does not reply '125' to '68 + 57' then he does not mean addition by 'plus'". This conditional would, of course, have to be more complex in order to make room for mathematical errors.

What justifies the assertion that Jones means addition by plus is then not any mental fact about Jones – it is not his intention to add nor is it any disposition or a mental image – but the fact that up to now he has replied to addition problems in a way that satisfies the communal standards for an adder. The facts that justify the assertion are thus indirect in the way that they are not the facts that make the assertion true, i.e. the sentence "Jones means addition by 'plus'" is not synonymous with the assertion that in the past Jones replied so and so to particular questions in particular circumstances. Moreover, the facts which justify the claim that Jones means addition by 'plus' do not exclude the possibility that he means something else. For all we know, he might reply

differently, even bizarrely, to the next mathematical problem and still maintain that he is just following the same rule as before.

But why should the community have these standards? Why should it bother accepting Jones as one of its members as opposed to, say, not pay attention to him? To answer these questions we must describe the role in our lives of the practise of accepting people into the speech community, i.e. we must describe the role of assertions such as "Jones, like many of us, means addition by 'plus'".

Kripke brings out this utility by an example of a man buying something at the grocer's. The customer will expect the grocer to add, count etc. like he does and not according to some bizarre quus-like rule; i.e. the customer will expect that in many cases the grocer will come up with the same answer as he would have given himself. (92-93)

> Our entire lives depend on countless such interactions, and on the 'game' of attributing to others the mastery of certain concepts or rules, thereby showing that we expect them to behave as we do. (93)

Our interactions with other people are filled with expectations that are not infallibly fulfilled. While any behavior can be compatible with any rule, it is not so with respect to our expectations since among them are expectations about particular responses in particular circumstances. The expectations of the customer put substantive restrictions on the behavior of the grocer, and this kind of restrictions cannot be made sense of if we consider an individual in isolation.

But can the sceptical challenge not be reinvoked at this point? Could not the sceptic ask why the grocer should attribute one expectation to the customer rather than some other or, to stick to the original formulation of the paradox, why we are so confident that others will give the same answer as we are inclined to give? But even if this challenge looks much like the original sceptical challenge, it is not about meaning but induction. (See 58)

This needs clarification. The new challenge is not about my knowledge of my past intentions, but why I should predict that one thing rather than some other will happen; it is not "How can I know that in the past I meant addition by plus and, hence, should reply '125' and not '5'?" nor even "How do I know that the grocer means addition by 'plus'?", but "Why should I predict that the

grocer will reply '125' rather than '5'?". I know why I do predict it – this is the answer I would give and usually people give the same replies to simple addition problems as I do – but does this provide a justification for my prediction? A demand for justification at this level is, according to Kripke's Wittgenstein, misguided. The fact that we generally agree must be taken as primitive.

This brings us to Wittgenstein's notions of a form of life. Agreement and form of life go hand in hand. Kripke writes: "The set of responses in which we agree, and the way they interweave with our activities, is our *form of life.*" (96) But how are we to identify or define a form of life, and how are we to distinguish one form of life from another? It is one of the lessons to be learned from the sceptical paradox that we can neither define nor explain a form of life; such a definition or explanation would constitute a straight solution to the sceptical paradox. This is not to say that other forms of life are impossible, only that they would be incomprehensible to us.

So, what is basic in this new picture is not the understanding of the individual, his conscious engagement with properties nor his grasp of certain rules, but his shared form of life. This means that we cannot explain our shared form of life in terms of the individuals' understanding – we cannot say that we all respond with '125' to '68 + 57' because we have the same understanding of the addition rule. We have the shared understanding because we have a shared form of life.

> . . . our license 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. (97)

If the notion of assertability conditions sounds puzzling, it may be helpful to compare it to standards for theory acceptance in the sciences. Theories are not checked by observing directly the facts that would make them true, but by indirect evidence. The theories predict certain things allowing us to formulate conditionals of the form "If theory A is right, then, given certain circumstances, so and so will happen". But no finite conjunction of such conditionals will determine that theory A, rather than some other, is right. That is why the emphasis is on the contrapositive form, i.e. theories are checked to see whether

they should be rejected. If no experiment provides a counterexample to the predictions of the theory, it can be accepted by the scientific community. Such an acceptance does, however, not mean that it could not be refuted at some later time. So, in short, the evidence which justifies acceptance of theories is indirect and only provides negative support.

And just as we raised the question why a community should have standards for acceptance of new members, we may ask why the scientific community should have standards for theory acceptance. An answer to this question along the lines of Kripke's Wittgenstein would describe the role of this practice for the scientific community.

The analogy with science should, however, not be brought too far since scientific discourse is just one of a variety of discourses and dependent for its existence on more primitive discourses such as that of attributing meaning to others' expressions.

Individual and community

What then is my response to the sceptic's challenge? The sceptic questioned the confidence with which I reply '125' to '68 + 57'. I would now justify my reply by reference to the community and its acceptance both of me as an adder and of my particular response. But suppose that no one except the sceptic was present when I gave my reply. Could he then not challenge me further by asking me why I am so certain that the speech community into which I have been accepted is a community of adders and not one of quadders? That is, if I answer the sceptic with a reference to a community, can the sceptical challenge not be repeated at that level? Are we not back to the 'rule for a rule' reply, only this time we have a community instead of a rule? So, in order to answer this challenge I would need to point out a fact that determined that the community is a community of adders and, say, not quadders, but, moreover, that fact would have to justify my assertion that it is a community of adders and not quadders.

Of course it won't suffice to point to the communal standards and say that they determine that the community is one of adders, since any formulation of these standards will be compatible with the community following an indefinite number of quus-like rules.

There are two issues here; firstly, is there a fact that determines that the community is one of adders and not of quadders and, secondly, is there a fact that justifies the assertion that it is a community of adders and not of quadders? Let's consider the latter question first. For any tentative hypothesis as to the nature of the community, I can accept or reject it. And since I am a part of this community, I am justified in replying just as I am inclined to so long as the community does not question it. So, in particular, I am justified in asserting that the community is one of adders and not quadders *because* I reply '125' to '68 + 57'. I follow the rule as it strikes me, and I predict that other members of the community will give the same reply.

But even if we can cite certain facts in support of the claim that the community is one of adders and not, say, quadders, we have not identified any facts that would make the statement true. Kripke would, perhaps, say that it is a primitive fact about the community – its form of life – that it is a community of adders. If that is right, we cannot explain that in terms of other facts about the community. The best we can do is, perhaps, to describe the role of addition in the community.

Summary of the sceptical solution

How then would I answer the sceptic when he questions the confidence with which I reply to the mathematical problem? I may be able to give a partial justification for my answer by explaining what computations I did. That would, however, only suffice to answer some mathematical doubts but not the sceptic's metalinguistic doubts and will only provide rules for interpreting rules, and ultimately I reach a level where I can give no justification. If I consider only myself – my mental and physical state – I can not distinguish between *following* these rock bottom rules or just *acting at random.*

In order to support the claim that I am actually following some rules and not just doing whatever strikes me as the right thing, I must, according to the sceptical solution, cite facts about the community to which I belong. The community has a criterion for an adder. Those who satisfy this criterion will be judged by the community to mean addition by 'plus' and not something else such as quaddition. So, in support of my confidence I cite the communal

approval of my use of the term 'plus'. These communal standards are, however, not interpretive of my utterances in the way truth conditions are.

The communal approval is based on conditionals like this: "If I do not reply '125' to '68 + 57' then I do not mean addition by 'plus'". So, in general, people are accepted as adders if they do not give the community a reason to the contrary.

At this point two further questions can be raised. Why should the community accept me as an adder, as opposed to, say, not pay attention to me? and Why does a conjunction of conditionals like the one above provide a criterion for following the addition rule as opposed to some quus-like rule since there will be indefinitely many rules that are compatible with any finite conjunction of such conditionals? The first question would be answered by describing the role of accepting the categorical assertion that I mean addition by 'plus'. That is Kripke's point in describing the interactions between the customer and the grocer. My only answer to the second question could be to give examples and see if the community agrees with me. My inclination to reply one way rather than some other must be taken as primitive, and, it is a primitive fact about the community that it agrees, more or less, in its responses to addition problems. This agreement cannot be explained but must be taken as primitive.

What then is the fact that I mean addition by 'plus'? That fact is, perhaps, a primitive fact and not reducible to anything else. But Kripke's Wittgenstein has little to say about it, his interest is in the fact that justifies the assertion that I mean addition by 'plus', which is a different one; it is the fact that I satisfy the community's criterion for an adder. This is not a superlative fact about me; it is not a fact about my physical state and, to borrow a jargon not used by Kripke, it does not even supervene on my physical state.[15] As meaning attributing assertions are justified by reference to the role and utility of such assertions, so must we justify other assertions. The assertion that it is a fact that Jones means addition by 'plus' will be justified by reference to the communal criterion for such assertions, not by those very facts.

---

[15] See Alex Byrne, "On Misinterpreting Kripke's Wittgenstein", *Philosophy and Phenomenological Research*, Vol. 56, No. 2 1996.

Kripke distinguishes two views regarding the need for an outward criterion: what we can call the 'official' view and the 'liberal' one. On the official view there must be, for each particular rule, a conditional of the form: "If Jones follows the rule, he should do so-and-so", and if this kind of a conditional is to have any content, it must be contraposed, we judge that Jones is not following the rule if he does not do so-and-so. But if there must be such a conditional for *each rule*, it follows that for *each rule* there must be an external check – there must be the 'so-and-so'. On this view, then, even if the speaker has been accepted into the community of adders and his sincere utterance has become the criterion for correctness, it must still be possible to check his performance with respect to each rule, so, in particular, it must still be possible to check his multiplication. The liberal version, on the other hand, allows that once a speaker has been accepted into a community, there can be rules where his mastery of those rules is presumed simply on the basis of his membership in the community.

When someone is taught multiplication by means of repeated addition, does not his grasp of the addition rule *explain* his understanding of multiplication? The liberal version of the requirement for an outward criteria explained above and endorsed by Kripke would suggest a positive answer to this question. If someone is accepted to the community of multipliers on the grounds of his mastery of the addition rule and his sincere statement that he knows how to multiply, then his grasp of the addition rule and his statement are taken to *justify* the attribution to him of the mastery of multiplication. But does this mean that we could save the truth-conditional picture for some higher level rules? Would there be *a fact about me*, in the sense of the justification argument, that justified the assertion that I meant multiplication by 'x'? I don't think so. The attribution of multiplication to someone on the grounds that he satisfies the criterion for an adder, has a reference to the community and is not based on conscious engagement with the truth conditions of such an assertion. So, in particular, the acceptance of the sentence "Jones means multiplication by 'x'" is not based on conscious engagement with its truth conditions but on

certain facts that are not identical with the truth conditions.[16] Even if it is a fact that I mean addition by plus, it is not 'a fact about me' in the above sense, and, consequently, neither will it be 'a fact about me' that I mean multiplication by '×'.

---

16 Scott Soames claims that the truth-conditional picture can be saved for higher order rules. See his "The Sceptical Solution Without the Sceptical Paradox", pp. 18-20.

PART III

The New Picture

That there are important connections between Quine's and Wittgenstein's philosophy of language is generally accepted and recognised both by Quine and Kripke. In *Word and Object* Quine says in a footnote:

> Perhaps the doctrine of indeterminacy of translation will have little air of paradox for readers familiar with Wittgenstein's latter-day remarks on meaning. (*WO* 77)

And at the outset of his discussion of the sceptical paradox, Kripke remarks that:

> Wittgenstein's sceptical problem is related to some work of two other recent writers who show little direct influence from Wittgenstein. . . . The first is W. V. Quine, whose well-known theses of the indeterminacy of translation and the inscrutability of reference also question whether there are any objective facts as to what we mean. (K 55)

But what is the relation between these two approaches? Even if both Quine's theses of the indeterminacy of translation and the inscrutability of reference and Kripke's interpretation of Wittgenstein have been discussed at length in recent years, exact comparisons of these approaches are rare. One attempt in that direction is Dirk Koppelberg's paper from the San Marino conference on Quine's philosophy.[1] But his account, as well as Kripke's own remarks, are both

---

[1] Dirk Koppelberg, "Skepticism about Semantic Facts", *On Quine: New Essays*, P. Leonardi and M. Santambrogio eds., New York 1995.

short and focus mainly on the differences between these approaches rather than what they have in common.

I will begin by noting some general similarities between Quine's approach and Kripke's interpretation of Wittgenstein, then I discuss Kripke's remarks about the differences that he sees between these approaches, but at last I will go on to give my own account.

I. Contacts

Both Quine and Wittgenstein rejected what was, and perhaps still is, a mainstream in the philosophy of language: what I have called the truth-conditional picture. But their affinities are not restricted to their criticism of other theories, their constructive accounts do share important features.

Both Quine and Kripke's Wittgenstein reject the notion of meaning as an explanatory notion. Kripke's solution to the sceptical paradox involved a rejection of the claim that truth conditions and extensions explain the meaningfulness of sentences and predicates as well as peoples' understanding of such expressions. Quine's rejection is, however, more dramatic. In *Pursuit of Truth* he writes:

> But I would not seek a scientific rehabilitation of something like the old notion of separate and distinct meanings; that notion is better seen as a stumbling block cleared away. (*PT* 56)

Quine's thesis of inscrutability of reference goes even further than rejecting the explanatory role of truth conditions as it explicitly rejects the idea that truth conditions are interpretive in the sense that they give the meaning of sentences. This follows from the fact that the truth conditions of sentences such as "There is a rabbit in the yard" and "There is an undetached rabbit part in the yard" are the same even if their meanings are different.

Further affinities between these two approaches are evident in the emphasis on communal agreement. In Kripke's picture the justification for accepting sentences such as "Jones, like many of us, means addition by 'plus'" would be the fact that his application of the term 'plus' is in agreement with the rest of the community. What is factual is the agreement, and that must, moreover, be

taken as primitive. Quine says strikingly similar things, for example in *Pursuit of Truth* where he is discussing the factuality of translation manuals:

> What is utterly factual is just the fluency of conversation and the effectiveness of negotiation that one or another manual of translation serves to induce. (*PT* 43)

Central to the approaches of both Quine and Kripke's Wittgenstein is the presupposition that semantic facts cannot be primitive. By a primitive fact I mean a fact that is not knowable a priori from a set of non-intentional facts. Quine, as is well known, claims that all genuine facts must be physical facts or dependent on physical facts. Kripke's Wittgenstein does not make any such claim but his requirement that any candidate for a meaning determining fact must show how I am justified in attributing meaning to utterances of mine or others' and his subsequent discussion of various candidates for a meaning determining fact, makes it clear that he does not allow for primitive semantic facts.

Determining facts

Up to now I have assumed that the notion of determination is unproblematic. But it is not. We can say that mean kinetic energy determines temperature, and also that atomic structure determines chemical properties, that the truth of premisses in a valid argument determines that the conclusion will be true, and these are different ways of determination. The first is an identity statement, the second is a theoretical reduction, and the third is a logical consequence. I want, however, to make a more general distinction between what I will call *a priori* determination and ontological determination. A set of facts F determines *a priori* that, say, Jones means addition by 'plus', if it can be demonstrated that Jones means addition by 'plus' by using F and basic logical principles. In order for F to determine ontologically that Jones means addition by 'plus', it is enough that it be impossible for F to be the case and, at the same time, that Jones does not mean addition by 'plus'.

In what sense does Quine claim that translation is not determined by any possible facts? Before we try to answer this question it is useful to recall Quine's

thesis about underdetermination of scientific theory. In *Word and Object* Quine wrote:

> [Molecular behavior is not determined by the behavior of ordinary things] even if we include all past, present, and future irritations of all the far-flung surfaces of mankind, and probably even if we throw in an in fact unachieved ideal organon of scientific method besides. (*WO* 22)

Quine's point here is that scientific theory is not determined by all possible evidence together with an ideal organon of scientific method. This is what we might call epistemic underdetermination: our theories about molecular behavior are not derivable from all possible evidence along with any ideal methodological principle. Is the relation between surface irritations and molecular behavior then contingent? Quine would deny this. He would maintain that this relation is in accordance with the principles of nature, and thus, ontologically determined. So, there are facts of the matter about molecular behavior and surface irritations of people, the problem is just that these facts are, in principle, hidden.

Quine's thesis about indeterminacy of translation goes further than underdetermination of scientific theory. Not only are translation hypotheses epistemically underdetermined, they are also ontologically underdetermined. It is not that the facts are hidden – there are no facts. Quine's reason for this strong thesis is that any fact that is relevant to translation must be observable – facts that are in principle hidden cannot determine translation – and since no observational facts determine translation, no facts at all determine translation.

Quine has sometimes been criticized on the grounds that his behaviorism, the principle that facts that determine translation must be observable, excludes certain facts that can be determining for translation. In this direction are Michael Friedman's remarks in his paper "Physicalism and the Indeterminacy of Translation".

> [the causal theory of reference] contrasts with the Quinean skeptical approach according to which the only semantically relevant physical relations between words and non-linguistic entities relate our uses of words to sensory stimulation's, stimulus meanings. Since different referents can yield the same stimulus meanings, we end up with the

doctrine of inscrutability of reference.[2]

What is the alternative to Quine's behaviorism? Friedman suggests that there could be non-behavioral physical facts that determine reference. These facts would be hidden in the sense that they would not tell us, under ordinary circumstances at least, that we were referring to one thing rather than some other. They would be like the physical facts that make us thirsty; we know that we want something to drink, but we do not know the physical facts that cause this desire. The reference determining facts would determine that we refer to, say, rabbits and not rabbit stages, but they would not determine this by 'telling' us what the reference of our expressions are.

These hidden facts would violate Kripke's Wittgenstein's basic requirement that any meaning determining fact must somehow 'tell' me how to apply expressions in new circumstances, since hidden facts do not tell me anything. Kripke's Wittgenstein would argue that a fact that determined my responses without 'telling' me how to respond, would not support my confidence – it would not justify my response. Such a hidden fact is, in that respect, no better than, say, my dispositions.

It seems then that both Quine and Kripke's Wittgenstein are talking about a similar kind of determination. We might perhaps, without too much simplification, attribute the following principle about meaning determination for sentences to Quine.

(D_1) A set of facts F determines that a sentence *S* means that *p* iff the sentence ⌜*S* means that *p*⌝ is derivable *a priori* from F,

where '*S*' is a name of a sentence and '*p*' is a sentence. Principles about predicate application would be similar.

In Quine's case the problem is that no fact determines to what kind of an object a certain expression refers; in Kripke's Wittgenstein's case the problem is that no fact determines how a certain individual should respond so the corresponding principle would be something like this:

---

[2] Michael Friedman, "Physicalism and Indeterminacy of Translation", *Noûs*, Vol. 9, No. 4 1975.

(D_2) A set of facts F determines that an individual *I* means *p* by *S* iff the sentence ⌜*I* means *p* by *S*⌝ is derivable *a priori* from F.

The crucial thing is that meaning determination – be it determination of translation or use – would consist in *a priori* derivability.

But is this notion of derivability consistent with Quine's physicalism? Quine holds that chemical facts are determined by physical facts even if we cannot derive chemical facts *a priori* from physical facts. The determination relation between physical facts and chemical facts seems then to be rather that the latter supervene on the former; they are ontologically determined even if they are not *a priori* determined. It appears then that Quine uses two notions of determination, *a priori* determination when he says that no facts determine translation, ontological determination when he argues for his physicalism. But does this threaten the validity of Quine's reasoning? As such, there is nothing wrong with using various notions of determination. The issue boils down to the plausibility of Quine's behaviorism, as Quine admits.

> Critics have said that the thesis [of indeterminacy of translation] is a consequence of my behaviorism. Some have said that it is a *reductio ad absurdum* of my behaviorism. I disagree with this second point, but I agree with the first. I hold further that the behaviorist approach is mandatory. In psychology one may or may not be a behaviorist, but in linguistics one has no choice. (*PT* 37-38)

In Kripke's Wittgenstein's approach, similar questions about ambiguity of the idea of determination can be raised. Why could it not be a primitive fact that I mean one thing rather than some other? Such a primitive fact could determine my use of certain words in the sense of ontological determination. Kripke's Wittgenstein would not have to reject that the notion of ontological determination is intelligible, for as Quine, he can hold that different notions of determination apply to different subject matters. But he would argue – or insist – that a set of facts does not determine that someone means one thing rather than some other by a certain expression unless it is an *a priori* consequence of those facts; i.e. he would argue – or insist – that facts determine meaning by way of telling us how to apply expressions in new circumstances.

II. Alleged differences

Despite these important similarities, these two approaches are in many ways different. It is, however, not clear what these differences amount to. There are obvious methodological differences, but are there also substantial differences between these two approaches? Are they perhaps incompatible?

Kripke identifies three differences between his understanding of Wittgenstein and Quine's theses. The first involves Quine's behaviorism, the second is about Quine's formulation of problems about meaning as problems of dispositions to behavior, and the third is about Quine's concern about the degree to which even infallible and unlimited dispositions determine interpretation. (*K* 56-57) Of these three differences Kripke seems to think that the second one is of most importance and that it does, in fact, show a defect in Quine's approach. The problem is that in Quine's approach there is no room for the normativity of meaning which is a central feature of the notion of meaning in Kripke's interpretation of Wittgenstein.

Quine's emphasis on behaviorism and Wittgenstein's extensive introspective experiments should point out an obvious difference in their approaches. But, far from denying this, I think that this difference is neither as radical nor of as much an importance as sometimes is thought. Even in *Word and Object* Quine would consider non-behavioristic facts and argue that they will not make any difference:

> . . . one can protest still that the sentence and its translations all correspond to some identical even though unknown neural condition in the bilingual. Now let us grant that; it is only to say that the bilingual has his own private semantic correlation – in effect his private implicit system of analytical hypotheses – and that it is somehow in his nerves. My point remains; for my point is that another bilingual could have a semantic correlation incompatible with the first bilingual's without deviating from the first bilingual in his speech dispositions within either language, except in his dispositions to translate. (*WO* 74)

This quote is not meant to be a justification for Quine's behaviorism. He is just answering the stubborn objection that bilinguals will certainly make the *right* correlations in virtue of their physical properties. But what kind of behaviorist

is Quine? Quine has rarely made an effort to justify his behaviorism, but it seems to me that he holds an underlying principle not so different from Wittgenstein's 'look, not think' principle together with the basic condition on a meaning determining fact. Quine's emphasis is on *describing* what words we learn first, how we learn words, etc. or how a radical translator could proceed. From this he draws the general conclusion that:

> There is nothing in linguistic meaning, then, beyond what is to be gleaned from overt behavior in overt circumstances.[3]

This does not mean that he would endorse the verificationstic claim that the meaning of a declarative sentence is reduced to the mode of verifying it. That is a reductionism about meaning which Quine would never accept. Quine's point is rather that the evidence for correlating an expression in the target language with an expression in the home language, say correlating 'gavagai' with 'rabbit', cannot be anything over and above what is observable. This does not mean that Quine denies the existence of what is not observable, only that facts – physical facts – that are not observable in 'overt circumstances' are irrelevant as far as linguistic meaning goes. That was Quine's problem with the homology of nerve ends; whether or not such homology is actually true, it should not matter.

How does all this fit into Kripke's interpretation of Wittgenstein? There are three points in Kripke's interpretation that are relevant here: (i) the demand for outward criteria, (ii) the 'look, don't think' principle, and (iii) the basic condition on a meaning determining fact that it must *tell* me how to apply the expressions of the language. Let's start with the outward criteria. As Kripke notes, Wittgenstein does not assume at the beginning of his discussion that an inner process stands in need of outward criteria, he deduces it towards the end, and the role of this principle in Kripke's Wittgenstein's reasoning is, hence, significantly different from that of Quine's behaviorism. But the 'look, don't think' principle and the basic requirement are premisses. The question is whether they amount so some sort of behaviorism in Quine's sense. The 'look, don't think' principle does not determine what facts will be relevant and, in

---

3 "Indeterminacy of Translation Again", *The Journal of Philosophy*, Vol. 84 1987, p. 5.

resemble one another . . ." (*WO* 27) According to this understanding, a description of a language would be exhausted by a description of the verbal dispositions of speakers of that language, and there seems to be no room for normative judgments. There is no standard in virtue of which we could say that even if the speakers have a disposition to do one thing rather than another, say to reply '125' to '68 + 57', they should do something else. But normativity is at the heart of Kripke's understanding of Wittgenstein; a basic condition on a meaning determining fact is that it should tell me how I ought to reply in the future, whether or not I am disposed to reply that way.

Dirk Koppelberg raises the further question of how Quine would or could react to this complaint. He suggests that Quine could respond in either of two ways: firstly, he could try to modify his dispositional account or, secondly, he might doubt that the normativity requirement is well defined by questioning whether there is "a clear-cut demarcation between the correct and the incorrect use of a word or between understanding or misunderstanding a linguistic expression" (342) Koppelberg then quotes a passage from Quine's *Pursuit of Truth* which he takes to be indicative of the latter option.

Koppelberg concludes that Kripke's understanding of Wittgenstein and Quine's approach are incompatible because there is no room for the normative character of the notion of meaning within Quine's framework. But while it is obvious that Quine does not give a normative account of the notion of meaning – in his extensive writings on meaning he rarely, if at all, raises questions about normativity – I am sceptical that this indicates a chasm between Quine and Kripke's Wittgenstein.

To see how this can be, we must make clear what the main aims of Quine and Kripke's Wittgenstein are and the role of the normativity requirement in the latter's approach.

Quine's thought experiment of radical translation is meant to show what we, in our own language, can say about the meanings of expressions in a target language. Quine does not raise questions about the meaning of expressions in the home language at the initial stage and neither does he question the appropriateness of independent utterances of the speakers of the target language. The question that Quine wants to answer is the following:

Q_1 What justifies the attribution of one meaning rather than some other to an expression of an alien language?

Kripke's Wittgenstein is, on the other hand, not concerned with semantic relations between languages, but with the attribution of meaning to utterances of speakers in a community. So, Kripke's Wittgenstein's central question becomes:

Q_2 What justifies the attribution of one meaning rather than some other to an individual's utterance?

These are very different questions. One questions the uniqueness of an interpretation of an alien language, the other questions the attribution of a certain intention to an individual.

The normativity requirement was that if I wanted to be consistent with my previous usage of an expression, say 'plus', then I ought to do one thing rather than some other when applying that expression, say replying '125' rather than '5' to the question '68 + 57?'. But as unquestionable as this seems to be, there is no room in Kripke's interpretation of Wittgenstein for applying such a requirement to the speech community itself. That would allow us to make normative judgments about forms of life but forms of life are indefinable and any judgments, normative or not, about forms of life are impossible.

According to Kripke's Wittgenstein, the normativity requirement is necessary to distinguish between someone following a rule and someone acting at random. The point is that without a fact that would tell me how I ought to apply an expression in new cases I could not justify my novel applications of it. The thrust of the sceptical paradox was that as long as we consider an individual in isolation we cannot meet this requirement; nothing about my mental or physical state justifies my attribution of one meaning rather than some other to my utterances, and, hence, nothing about my mental or physical state justifies the claim that if I want to be consistent with my previous language usage, then I ought to do one thing rather than some other. The sceptical solution accepted this conclusion, but met the normativity requirement by considering the individual as a member of a certain community of speakers. But the nature of the community is left indeterminate.

We might, perhaps, justify an attribution of a particular meaning to certain expressions by integrating the community into yet another community. This could be possible if the first community was, say, a community of scientists, and the second one a community of speakers of some language. But ultimately we reach a level where no such justification is possible, and we cannot say that the community ought to assign one meaning rather than some other to any expression. If this is right, no normativity requirement can be applied to the whole community and the best we can do is, perhaps, to describe how expressions are actually used in the community; i.e. describe the individuals' dispositions to verbal behavior and the standards for accepting individuals into the community. The former is what the radical translator tries to do, the latter is the task of lexicography.

Understanding and misunderstanding

Kripke's Wittgenstein maintains that my understanding of a term determines how I ought to use it in the future. So, in particular, if I understand that 'plus' stands for addition, this would uniquely determine how I ought to reply to any problem of the form 'x + y'. In this sense there are strict boundaries between understanding and misunderstanding. This is perhaps what Koppelberg has in mind when he interprets the following excerpt from Quine's *Pursuit of Truth* as questioning whether my understanding of a term determines how I ought to use it.[5]

> Lexicography has no need for synonymy, we saw, and it has no need of sharp distinction between understanding and misunderstanding either. The lexicographer's job is to improve his reader's understanding of expressions, but he can get on with that without drawing a boundary. He does what he can, within a limited compass, to adjust the reader's verbal behavior to that of the community as a whole, or of some preferred quarter of it. The adjustment is a matter of degree, and a vague one: a matter of fluency and effectiveness of dialogue. (*PT* 59)

I think that Koppelberg misunderstands the whole point of this passage and,

---

5 See his "Scepticism about Semantic Facts" p. 342.

far from being a rejection of Kripke's Wittgenstein's claim, fits nicely into the sceptical solution.

What would it mean if the lexicographer could draw boundaries, if he could draw sharp distinctions between understanding and misunderstanding? Let's consider the original sceptical challenge. We want to say that someone who replies '5' to '68 + 57' has misunderstood the '+' symbol, whereas someone who replies '125' has probably understood it, and in general, there is a uniquely determined answer for any two arguments. In this sense, there are clear cut boundaries between understanding and misunderstanding. But any finite formulation of a standard for correct use will leave the meaning of the '+' symbol undetermined. The problem arises because of the infinity of the addition table and the fact that the criterion for a correct use must be finite. It should be obvious from the discussion of the sceptical paradox, especially Wittgenstein's well known claim that an interpretation cannot determine meaning, that any such attempt will be incomplete. The lexicographer cannot hope to do any better than describe the standards of the community of speakers for accepting someone as a member, and as we saw when discussing the solution to the sceptical paradox, these standards do not determine the meaning of the expressions of the language.

According to the solution to the sceptical paradox, the nature of the community is indeterminate; we are justified in asserting that, say, Jones means addition by plus, by considering him as a member of a community of adders. That the community is actually a community of adders and not of people following some quus-like rule must be presupposed. If we could determine the nature of the community with respect to addition as well as other rules, we could define its form of life. But as Kripke emphasises, the form of life must be taken as primitive. Now, if the lexicographer could draw exact boundaries, if he could define a correct use of the '+' symbol for any argument, he could thereby determine whether the community was a community of adders, and, hence, define its form of life.

From this it should be clear that Quine's understanding of the lexicographer's job, so far from being a rejection of it, is in perfect coherence with Kripke's interpretation of Wittgenstein.

The third difference which Kripke mentions between his interpretation of Wittgenstein and Quine's proposal, is that Quine is "concerned to show that even if dispositions were ideally seen as infallible and covering all cases, there are still questions of interpretation that are left undetermined". (*K* 57) By this he means that Wittgenstein has nothing directly comparable to Quine's theses of the indeterminacy of translation and the inscrutability of reference.

From what I have already said about the normativity requirement, it should not be surprising that Wittgenstein does not offer anything comparable to Quine's theses. Since Quine's theses apply primarily to a community of speakers and not to speakers in a community, they are simply outside the scope of his considerations. But even if the sceptical solution is not primarily about a community of speakers but speakers in a community, it is suggestive about the nature of such a community. As I have already mentioned, it follows from the sceptical solution that there is no fact which determines the nature of the community, but that is the same conclusion as Quine arrives at. Kripke's Wittgenstein claims that there is no fact which determines whether a community is a community of adders since any specification of the nature of the community will be compatible with the community following some quus-like rules. This conclusion may seem radical enough, but even so Quine arrives at a still more radical conclusion. He claims that there are specific hypotheses the correctness of which is not determined by any possible facts, but as for any specific quus-like alternative there are possible facts which will determine whether it can be true.

III. Rationality and linguistic competence

If the facts of the matter of meaning attribution are just facts about fluency of conversation or membership in a community and not anything about mental states, what does it mean to say that I know what I mean by a particular predicate? I think that both Quine and Kripke's Wittgenstein would give a similar answer: "I know what I mean by a predicate if I know how to use the predicate." And the criterion for correct use is just fluency of conversation, as

PART IV

Resisting the Arguments

Should we try to resist these arguments? and if so, How could we do that? We could doubt the general claim that the only facts relevant to meaning attribution are facts about fluency of communication. Both Quine and Kripke's Wittgenstein seem to agree on this point. Or we may raise more specific doubts about each approach.

According to Quine's theses, discussions about right or wrong ontology, and metaphysics in general, are pure speculations and admit of no extra-linguistic standards by which they can be judged. Quine does not reject the meaningfulness of such discourse, as did the logical positivists, but he claims that ontological questions cannot be settled in principle.

Quine argues that our frame of reference is not just the objects to which we refer, but also the individuative tools of our language, the most important of which is the notion of identity, and that there is no semantic criterion for identifying it in the target language. But *given* a term in the target language for identity, ontological relativity would be resolved and we could determine which of the alternatives such as 'rabbit', 'rabbit stage' and 'undetached rabbit part' is correct. The problem is, however, that identity belongs to the theoretical discourse of science and is, therefore, among those parts of the language that admit of most indeterminacy. That is the point of the following passage from *Ontological Relativity*.

> Our theory of nature grades off from the most concrete fact to speculations about the curvature of space-time, or the continuous creation of hydrogen atoms in an expanding universe, and our evidence grades off correspondingly, from specific observation to

> broadly systematic considerations. Existential quantifications of the philosophical sort belong to the same inclusive theory and are situated way out at the end, farthest from observable fact. (*OR* 98)

Similar problems arise in Kripke's interpretation of Wittgenstein. We could determine that, say, someone means addition by 'plus' *given* that the speech community to which he belongs is one of adders but we have to accept the conclusion that the nature of the community is totally indeterminate. Not even God could know whether the community to which I belong is a community of adders and not one of people following some quus-like rule. It becomes difficult to say that the community follows any such infinite rule. But if the nature of the community is ontologically indeterminate, then it becomes questionable whether it is, in the end, a fact that I mean one thing rather than some other.

But is it such a bad situation that identity cannot be translated with fair certainty or that the nature of the community is indeterminate? Is it not unquestionable that in order to demonstrate anything, be it the meaning of an expression or a mathematical theorem, we need basic principles that must be taken as given? While that is true the problem here is a bit different. Even when we have to take theorems as given, we still want them to be factual. A basic principle, say the principle that the shortest distance between two points is a straight line, is supposed to be factual; even if it cannot be proved or demonstrated it is supposed to state a fact. But Quine and Kripke's Wittgenstein have questioned the very factuality, not only the demonstratability, of meaning attribution. This is what the theses of the indeterminacy of translation and the inscrutability of reference state explicitly, and, as for Kripke's Wittgenstein, there is no fact that determines the nature of the community to which I belong.

I. Ontological Relativity

Let me start with my concerns about Quine's thesis of inscrutability of reference. What is needed in order to reject it? Quine gives two kinds of arguments for his thesis: The actual example of the Japanese classifiers, and a theoretical argument to the effect that there is no semantic criterion available to determine the ontology of the language, so in particular that there is no

semantic criterion available to determine what kind of a thing a term like 'gavagai' refers to. I will not try to reject the example of the Japanese classifiers, and I don't think I need to, since these classifiers do not raise questions about ontology. My aim is not to resolve all inscrutability of reference, only to limit the scope of the thesis to non-ontological discourse. Now there are three ways to proceed: we could argue (i) that we can determine whether one ontology should be preferred to some other within each language, (ii) that ontological differences will show up in linguistic behavior and affect the fluency of conversation, or (iii) we can try to identify semantic criterion that will allow us to determine which of alternative translations of, say 'gavagai', is correct. The first two ways are, roughly, the classical ways of doing metaphysics, the third, which I do not take, is more along the lines of Quine's own considerations.

Let me start with the classical metaphysical approach. Quine maintains that there can be ontologically incompatible translation hypotheses all of which are equally compatible with the totality of verbal dispositions. In other words, a term such as 'gavagai' can be translated either as 'rabbit', 'rabbit stage' or 'undetached rabbit part' and no behavioral evidence can determine which alternative is right. My doubts about this conclusion are twofold. Firstly, it is not clear in what sense these alternatives indicate different ontologies, and secondly, I think that in as much as they are different they will prompt different scientific questions and that this difference will show up in the behavior of the scientists.

Let's begin with the second point. What is the relation of ontology and science? Is it really possible that there could be two bilinguals, Albert and Niels, who, as Quine would put it, do not deviate in their speech dispositions within either language except in their dispositions to translate? (See *WO* 74) Suppose that Albert translates 'gavagai' as 'rabbit' while Niels translates it as 'rabbit stage'. As long as these two men confine themselves to the woods pointing out what they see, their differences will not surface within either language. But we have to allow them more complex activities such as the practice of science. It appears to me that Albert and Niels will be interested in quite different questions. In particular, Niels might want to know how long a rabbit stage lasts, why one rabbit stage produces another, etc. But I am very sceptical that Albert would take interest in those questions. The differences prompted by these

views will be even more compelling if Albert and Niels gained an interest in psychology, for Niels would have to explain how he, who is nothing but a collection of objects whose only relation is that of cause and effect, has memories, self-identity, etc.

But in what sense do 'rabbit', 'rabbit stage' and 'undetached rabbit part' indicate different ontologies? Whichever of these alternatives we adopt, we need something in virtue of which, say, one rabbit stage is of the same kind as another rabbit stage and different from, say, an elephant stage; i.e. we need some classification principle. What can that be? We might identify stages in virtue of shape, purposes, powers etc. But whichever alternative we take, I do not see how we could avoid introducing some such principle as forms of things. And, unless we become idealists, we would maintain that these things are made of matter. The forms alone do not constitute anything. But what then is in the way of identifying these temporal stages as objects in an Aristotelian sense – a unity of form and matter? I cannot see how temporal stages will be anything but short-lived objects. If that is right, the difference in ontologies between a translation manual which translates 'gavagai' as 'rabbit', and another manual that translates it as 'rabbit stage', is not one of objects as opposed to something else. The difference will be about the nature of these objects, i.e. the nature of gavagai. Someone who translates 'gavagai' as 'rabbit' will look at them as indivisible objects in the sense that they are not composed of similar things, while another who translates it as 'rabbit stage' would treat them as a collection of similar objects.

The difference between the translation of 'gavagai' as 'rabbit' and as 'rabbit stage' is, therefore, better seen as an attribution of different scientific theories to the speakers of Jungle. These scientific theories will be empirically equivalent but, nevertheless, incompatible. The situation will be similar to that in Poincaré's example of the two theories one of which postulates infinite space of rigid objects, the other which postulates a finite space with shrinking objects. But if that is right, Quine's ontological relativity is perhaps nothing over and above underdetermination of scientific theory.

All this would need much longer discussion. My points are (i) that the alternatives that Quine mentions are not as much a difference in ontology as, say, physics or biology, and (ii) that our scientific interests are guided by our

At the outset of the sceptical solution Kripke writes:

> We have to see under what circumstances attributions of meaning [Jones means addition by 'plus'] are made and what role these attributions play in our lives. Following Wittgenstein's exhortation not to think but to look, we will not reason *a priori* about the role such statements *ought* to play; rather we will find out what circumstances *actually* license such assertions and what role this license *actually* plays. (*K* 86-87)

Given this methodological principle, is it surprising that Kripke's Wittgenstein reaches the conclusion that an individual considered in isolation can not be said to mean one thing rather than some other by whatever utterances he makes? I think not. Suppose that we have not had the opportunity to consider an individual in isolation, and that the circumstances that actually license meaning attributions to individuals in a community have reference to the community. It seems to me that from these suppositions and the methodological principle above, we can conclude that any meaning attribution must involve a reference to a community. But, without further reasoning, this is a conflation of actuality and possibility.

Now, the question is whether Kripke's Wittgenstein is guilty of this modal conflation. I think he is. A central feature of Kripke's sceptical solution is the attribution of expectations that are not infallibly fulfilled. After describing interactions between a customer and a grocer, Kripke writes:

> Our entire 'game' of attributing to others the mastery of certain concepts or rules, thereby showing that we expect them to behave as we do.
>
> This expectation is *not* infallibly fulfilled. It places a substantive restriction on the behavior of each individual, and is *not* compatible with just any behavior he may choose. (*K* 93)

It may well be true that this 'game' of attributing expectation does only take place in a community and consists in attributing expectations to others. But this does not exclude (i) the possibility of someone attributing expectations to himself, (ii) that these 'private' expectations are not infallibly fulfilled, and (iii) that this be noticed by the person in question. Suppose, for example, that Robinson Crusoe is stranded on a deserted island. He might make marks on

trees to work his way around the island. In following these marks, he will expect that he gets to a certain destination, say back to his bungalow. But what if he gets lost in the jungle? He will most certainly recognize that his expectation was not fulfilled. And how will he explain this? He may think that he didn't follow the marks correctly, he may even swear to himself to follow them more carefully next time. And here we have an example of someone recognizing that his 'private' expectations are not fulfilled by his own actions.

It must be recognized that rule following does not only take place in language but is a substantial part of our ways of going about in our every day life and those ways can be our own, as C. B. Martin notes:

> There are natural procedural ways of behaving directed to a particular outcome (that may or may not be successful) that need not be learned from others. These procedures are reinforced by success of outcome and not necessarily by reward from others. . . . The physical environment itself has a role to play in our learning by rewarding with success and punishing by failure our endeavours.[7]

The indispensability of external reality is that without it we could not distinguish between a right and wrong way, a success and failure.

III. External Reality

I have not objected to the general picture that Quine and Kripke's Wittgenstein propose; the idea that our linguistic competence is based on practical knowledge rather than theoretical knowledge, that our understanding of expressions does not consist in a conscious engagement with truth- or application conditions, and that meanings of expressions cannot be reduced to some mental states. But I think they go too far in their emphasize on a community of speakers for all we need is external reality, whatever that is.

Kripke's Wittgenstein argues that the only way to make sense of the normativity of meaning is to have a community of speakers. But this is plainly wrong. An individual considered in isolation from other people can realize

---

7 C.B. Martin, "Proto Language", *Australasian Journal of Philosophy*, Vol 65, No. 3 1987, p. 284.

that he went wrong even if he carries out his actions with great confidence. That was the point of my example of Robinson Crusoe. If, on the other hand, someone were to do everything in his mind, I cannot see how this normativity requirement could be met. The key notion in the establishment of the normativity requirement is that of expectation; the expectations of the grocer and the customer about certain computations or Robinson Crusoe's expectation that he will get back home. These expectations are not infallibly fulfilled. If I am, on the other hand, confined to my mind, I am not even sure how to make sense of the idea that my expectations were not fulfilled. The key thing about external reality is that it is an authority independent of our will and wishes. I may expect a dry sunny day, but certain circumstances will prove that my expectations did not come true.[8] If we were, however, confined to do things in our mind, there would be no authority independent of our will and wishes. All we could say is that whatever strikes us as right is right, which means only that talk about 'right' and 'wrong' does not make any sense.

I reject then Quine's thesis about ontological relativity. That, however, does not resolve the inscrutability of reference or the indeterminacy of translation. It will only limit the scope of these theses to non-ontological discourse. I also reject Kripke's Wittgenstein's claim that the only way of making sense of the normativity of meaning is by reference to a community of speakers. My claim is that instead of a community we can do with external reality. My objections to the claims of Kripke's Wittgenstein have further consequences for his thesis about the impossibility of private language since a private language would be impossible only in the sense that one needs the company of reality independent of one's will.

---

[8] One might object that we could always insist that it was indeed dry and sunny, even in the worst of rainstorms, just by claiming hallucinations. But against such extreme measures, there is no defense.

Bibliography

Paul A. Boghossian, "The Rule-Following Considerations", *Mind*, Vol. 98, No. 392 1989, 507-549.

Alex Byrne, "On Misinterpreting Kripke's Wittgenstein", *Philosophy and Phenomenological Research*, Vol. 56, No. 2 1996.

Michael Friedman, "Physicalism and the Indeterminacy of Translation", *Noûs*, Vol. 9 No. 4 1975, 353-373.

Dirk Koppelberg, "Skepticism about Semantic Facts", *On Quine: New Essays*, P. Leonardi and M. Santambrogio eds., New York 1995.

Saul A. Kripke, [*K*] *Wittgenstein on Rules and Private Language: An Elementary Exposition*, Oxford 1982.

C.B. Martin, "Proto Language", *Australasian Journal of Philosophy*, Vol. 65, No. 3, 1987, 277-289.

C.B. Martin and J. Heil, "Rules and Powers", *Philosophical Perspectives*, forthcoming.

On Quine: New Essays, P. Leonardi and M. Santambrogio eds., New York 1995.

W.V. Quine, [*WO*] *Word and Object*, Cambridge Ma. 1960.

— [*FLPV*] *From a Logical Point of View*, second edition, Cambridge Ma. 1961.

— [*OR*] *Ontological Relativity and Other Essays*, New York 1969.

— [*ORIT*] "On the Reasons for Indeterminacy of Translation", *The Journal of Philosophy*, Vol. 67 1970, 178-183.

— [*RR*] *The Roots of Reference: The Paul Carus Lectures*, La Salle 1974.

— [*TT*] *Theories and Things*, Cambridge Ma. 1981.

— [*PL*] *Philosophy of Logic*, second edition, Cambridge Ma. 1986.

— "Indeterminacy of Translation Again", *The Journal of Philosophy*, Vol. 84 1987, 5-10.

— [*PT*] *Pursuit of Truth*, revised edition, Cambridge Ma. 1992.

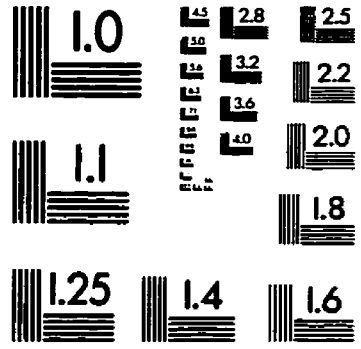
Scott Soames, "The Sceptical Solution Without the Sceptical Paradox" unpublished.

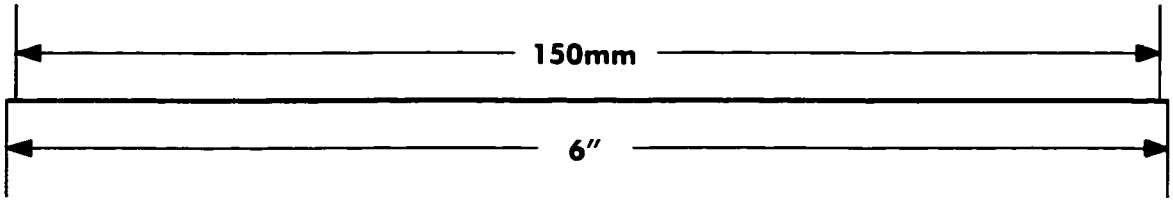
— "Skepticism about Meaning: Indeterminacy, Normativity, and the Rule-Following Paradox", unpublished.

Words and Objections: Essays on the Work of W.V. Quine, D. Davidson and J. Hintikka eds., Dordrecht 1969.

José L. Zalabardo, "Kripke's Normativity Argument" *Canadian Journal of Philosophy*, forthcoming.

TEST TARGET (QA-3)





APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989