Between Quine's Disquotationalism and

Horwich's Minimalism

Richard Wei Tzu Hou

A thesis submitted in fulfillment of the requirement

for the degree of Doctor of Philosophy

Department of Philosophy

The University of Sydney

May 2006

Abstract

Many criticisms of the prevalent deflationary theories of truth stem from some misunderstanding. Clarification can be found from considering Quine's reasoning on the disquotational feature of the truth predicate. Quine's disquotationalism and Horwich's minimalism are similar theses with respect to the concept of truth, though the difference between the choices of the primary truth bearers and the divergence in their accounts of meaning and reference are striking.

Chapter Two is devoted to making plain Quine's reasoning regarding the disquotational concept of truth, and to constructing a disquotational theory of truth. Also in this chapter, the topic of how to enhance the deductive power of this theory is discussed. The following chapter aims to square Quine's theses of inscrutability of reference and ontological relativity, with an account of the disquotational schema of reference. Whether or not a disquotational schema of reference and all its instances can be seen as providing a genuine reference scheme, as claimed by Horwich and most deflationists, is also discussed. In Chapter Four, after an introduction of Horwich's minimalist conception of truth, there are a number of issues considered, in particular Horwich's use-theoretic account of meaning and compositionality, along with the divergence between his account of meaning and Quine's. The final chapter, Chapter Five, provides a thorough analysis of three important factors regarding the disquotational theory and the minimal theory of truth. Among them, the first factor discussed is what sort of equivalence relation occurs within each instance of the disquotational schema or each axiom of the equivalence schema. Following this, there is an analysis of in what way the disquotationalist and the minimalist can explain all general facts involving truth. The last factor involves considering the proper ascription of the disquotational or the minimal truth predicate. Along with the analysis of these three factors, the issue regarding which theory of truth is preferable is elaborated.

Table of Contents

Acknowledgement

Chapter I

Introduction

Deflationism is probably the most noteworthy semantic theory of the past twenty odd years. It is still influential today. Among hundreds, if not thousands, of articles, discussing, criticising, or defending deflationism, some important issues are not as clear as they should be. This is, I argue, because of a pervasive tendency to over-complicate the key ideas. More complex versions of key aspects of the theory are subject to misunderstanding and refutation. The simpler versions are not.

This thesis argues that this over-complication has caused problems in dealing with at least eight questions that arise for deflationary theories. These include:

1. What is a proper formulation of a disquotational schema of truth?

2. What is the modal status of disquotation sentences or the axioms of the minimal theory of propositional truth?

3. How sensible are the presuppositions involved in asking these questions about the modal status of sentences or axioms?

4. How should we give the deflationary theory of truth the deductive power it needs so as to adequately explain all general facts involving truth?

5. What is the deflationary conception of reference? Is it feasible to explicate the deflationary conception of truth in terms of the disquotational concepts of reference and satisfaction?

6. What is an appropriate account of propositions for use in a minimalist account of

truth?

7. What sort of equivalence relation occurs within disquotational sentences or minimalist axioms of truth?

8. Between disquotationalism, taking sentences as the primary truth bearers, and minimalism, taking propositions as the primary truth bears, which is the more promising approach to a deflationary theory of truth?

These questions will be discussed only as they arise for the cases of disquotationalism and minimalism. If I am right the sources of many of the perceived problems of these theories will be exposed. At very least the theories will be clearer, and perhaps the reader will find them more plausible. However, since this dissertation is so confined, other deflationary theories of truth such as the prosentential theory will not be discussed. Hereafter, when I use the term 'deflationism' I use it to refer to some versions of disquotationalism or minimalism. There are of course many versions of deflationism—even versions of disquotationalism or minimalism—which I do not discuss; but my task will be achieved if I clear some of the obstacles to acceptance of the versions I do discuss, and this can be done without a more exhaustive survey.

Apart from this agenda, this thesis makes another claim. It claims that one can read Quine as more or less tacitly accepting a disquotationalist or minimalist theory of truth. This is not a view which is widely held. Although the general idea of disquotationalism and some deflationist theories is inspired by Quine's viewpoint on truth, few seriously see Quine as adhering to a deflationist position regarding the concept of truth in particular, or the semantic theory in general. Among other reasons, the most important one is that Quine

specifically expresses his support for a Davidsonian truth conditional approach to explicating the concepts of truth and meaning. However, discussion regarding this issue will not be addressed until the very end of this dissertation.

Except in contexts where a deflationist or an anti-deflationist introduces the idea of considering the truth predicate as a device of *semantic ascent*, Quine's reasoning is hardly mentioned. Put aside for the moment whether or not Quine is really a deflationist, and disregard his viewpoints on truth that is not disquotational. Nevertheless, it is still the case that it is Quine who sets up clear and distinct guidelines on a proper understanding of disquotationalism. In Chapter Two, I will explain what Quine's reasoning is, and how it gives us a clear perspective of disquotationalism. Another important task in this chapter is to analyse in what sense a disquotational theory of truth can be seen as providing enough deductive power to explain some general facts involving truth. It will be shown that to explicate truth in terms of the inductive axioms of reference and satisfaction is a dubious move.

This leads us to Chapter Three. Along with the deflationary theory of truth, most deflationists accept a disquotational theory of reference (and satisfaction). However, to consider a disquotational theory of reference as providing a genuine reference scheme for a given language, Quine's thesis of ontological relativity seems to be fallacious. In this chapter I show that it is not in fact fallacious. After that, I will present my interpretation of the role played by a disquotational schema of reference in Quine's disquotationalism. Finally I will argue that a disquotational schema of reference cannot be seen as providing us a genuine reference scheme.

Then we turn to Horwich's semantic theory. In Chapter Four, after a brief introduction to Horwich's views on truth, along with the discussion of whether or not minimalism can be finitely formulated, the focus is on one rationale of the minimalist theory of propositional truth—namely, a viable account of proposition. The purpose of this chapter is to show that Horwich's rejection of Quine's indeterminacy thesis fails, that Horwich's use-theoretic account of meaning is unsatisfactory, and that his account of compositionality not only fails to be deflationary, but is also based on some uncritically accepted assumptions such as that different languages share the same synthetic features.

Based on the outcome of the preceding three chapters, I make some important clarifications in Chapter Five; specifically I enunciate three important factors for a successful deflationary theory of truth. The first of these factors concerns the functional role of the truth predicate to do the semantic ascent required to help express infinite conjunctions and disjunctions. The second of these concerns the adequacy of the explanatory role in terms of which all facts involving truth are explained along with other suitable explanatory resources. The last of these concerns the proper ascription of the truth predicate. I then show that, although Horwich's standpoint on the theory of meaning is sharply demarcated from Quine's, some basic ideas of his minimalism are not very different from those behind my formulation of Quine's disquotationalism. The first section of the chapter, then, concerns the first factor above. In it I analyse what sort of equivalence relation occurs within disquotation sentences and the axioms of the minimalist theory of truth. The second section is devoted to the analysis of how disquotationalism and minimalism can account for all general facts involving truth. Then

in the third section, the proper application of the disquotational or minimalist truth predicate will be given. Along with the study of these three factors, I make an assessment of Quine's disquotationalism versus Horwich's minimalism. Finally, in the last section I consider some not very coherent attitudes of Quine, toward the matter of truth.

There are two additional matters that need to be addressed. In Chapter Three I concluded that a disquotational schema of reference is not supposed to be seen as providing a genuine reference scheme. It should be noted that the consequence of this observation is not pursued in this dissertation. Furthermore, the issue of how to make sense of the disquotational and the minimalist accounts of truth conditions is not discussed, for this issue does not undergird any difference between disquotationalism and minimalism and is therefore tangential to this thesis. The other matter is rather minor. It is about the relation between a deflationary account of truth and any given theory of meaning, insofar as it is not a truth conditional one. In both Chapter Two and Chapter Five, after I draw the conclusion that a proper understanding of the disquotational schema of truth and its instances should not involve any particular way of interpreting the content of the sentence mentioned and used in each instance of the schema, I will not further pursue how to sort out an account of meaning that is suitable for a disquotational theory of truth. For I rather think that this dissertation will not benefit from any conclusion reached by our study of this issue. Accordingly, these two issues are left for future study.

Chapter II

Quine's Disquotational Theory of Truth

and Semantic Ascent[1]

Deflationism is among the most prominent theories of truth. Two of the most important deflationary theories of truth are disquotationalism, inspired by Quine, and Horwich's minimalism. Both of these can be seen to have developed from Tarski's Convention T, though offering quite distinct theories that diverge in significant ways.[2]

Quine's famous theory about translation—that translation is indeterminate—is one of the most significant and influential theories of meaning. It is opposed to the more intuitive account of meaning that takes intensional notions such as analyticity and synonymy as playing central roles in theories of meaning and/or content. This approach also has an effect on his account of truth, namely, that his disquotational theory of truth is a revision of Tarski's ingenious T-schema. Although one of the major contributions of Tarski's semantic account of truth is to explicate the semantic conception of truth in terms of the analysis of satisfaction and reference, this is what Quine's disquotational account of truth rejects. The importance of this disquotationalist account is not just that it offers a new

---

[1] A draft of this chapter has been presented in a seminar held at the Institute of European and American Studies, Academia Sinica, Taiwan in 2004.
[2] The simplest version of Convention T is the set of bi-conditionals with the form (T):
 (T) *X* is true if and only if *p*,
     where '*X*' is replaced by the name of any sentence of the object language and '*p*' by the sentence itself or by a suitable translation within the metalanguage.
Convention T plays an important role in Tarski's theory of truth, that is, in the material adequacy of any given definition of truth. More precisely, any definition of truth (*D*) is materially adequate if and only if it implies all the T-biconditionals of the object language and implies that only sentences of the object language are true in the same language.

way of thinking with regard to the notion of truth, but it also fills in the final touch of Quine's general conception of semantics. Another crucial part of this account is Quine's interpretation of Tarski's semantic theory of truth. His interpretation shows that the truth predicate has a special function both as grammatical device and in philosophical contemplation as *semantic ascent* (and *descent*), by which we substitute the talk of sentences for the talk of reality. However, Quine does not offer a precise formulation of his disquotational theory of truth. He also fails to specify to a desirable extent how this theory behaves if it is more desirably strong in deducing some generalisations.

This chapter thus consists of three sections. The first section devotes a brief introduction to Tarski's Convention T. The other important part of this section is to explicate Quine's disagreement with Tarski's theory—the disagreement that set him on a divergent road. The following section aims at clarifying Quine's account of truth. Under a proper characterisation of deflationism, a major task of this section is to roughly provide a basic formulation of disquotationalism. Further, how the truth predicate helps express infinite conjunctions and disjunctions, and how strong the deductive power of the disquotational theory of truth is, are two questions in need of answers. In order to answer these questions, a discussion with regard to the problem of conservativeness of disquotational theories over logic or the base theories forming them, is given in the third section. In addition, in view of both Quine's earlier and latter remark on his account of truth, a specification of whether or not his theory of truth is substantive rather than deflationary will not be discussed until the final section of Chapter Five.

Here I must emphasise that I will not discuss disquotationalism in terms of

counter-factuality or related notions in this chapter. This is why many accounts of disquotationalism following this line, Field's quasi-disquotationalism, for instance, are not discussed until Chapter Five. Counterfactuals and possible world semantics seem to owe their clarity to a decent theory of truth but not the other way round. So it is probably a good idea not to give an extensional analysis of disquotationalism in an intensional atmosphere, at least not now.

1. Tarski's Paradigm on Truth

Many philosophers have felt the urgent need to deal with semantic paradoxes, especially the Liar Paradox. Tarski tells us that for any semantic closed language, if the usual laws of logic hold then there will be inconsistent applications in the application of some that language's semantic terms.[3] The formulation of a semantic account of the intuitive concept of truth that eliminates such semantic paradoxes is the primary task of Tarski's theory.[4] Another way to see it is, according to Field (1972), to reduce the concept of truth to other semantic concepts and then to reduce these concepts to, or to explicate them in terms of, physical or non-semantical terms. Since the purpose of my discussion is to account for Quine's interpretation of Tarski's paradigm, only the general idea of Tarski's theory is in need of explication. Following that, my version of Quine's interpretation can be addressed.

A sketch of Tarski's semantic conception of truth

In order to establish a semantic theory of truth without the threat of semantic paradoxes,

---

[3] That is, it can account for its own semantics.
[4] A semantic account of truth explicates the intuitive concept of truth in terms of semantic terms and first order standard logic plus set theory.

Tarski restricts the attribution of truthfulness to sentences of a language whose semantics

is to be constructed in another richer language. This is the famous distinction between

object- and meta-languages. Tarski's insight is that, if a language contains its own

semantic vocabularies and the usual laws of logic hold in it, the semantic theory of this

language that is constructed within the language itself, will be inconsistent. Tarski

presupposes that the object language is a set of expressions (of the first order arithmetic,

say,) with meanings. Thus if a semantic theory of truth for that given language is to be

constructed, a list of the object language containing all primitive expressions, as well as

the rules of definition by means of which new expressions are introduced in terms of

primitive ones, has to be laid down. A set of formation rules determining which sequences

of symbols from its alphabets are sentences among all primitive expressions has to be

specified. Meanwhile, the deductive apparatus of this language has to be given by setting

down (by fiat) how to infer some sentences from other sentences.[5] Further, besides the

complete devices of first-order logic and set theory, the metalanguage must include

expressions referring to those in the object language. The metalanguage is to be capable

of describing structural features and relationships of expressions of the object language,

that is, following Tarski's terms, the morphology of the object language. It has to include

expressions of the object language or their translations as well. According to the above

characterisation, object languages and metalanguages are expanded languages and

metalanguages in particular are with full logical and set-theoretic machinery.

Since the object language is about reality, a sentence, say, "Snow is white"[6] of the object

---

[5] If the object- and the meta-languages are both artificial, the deductive apparatus of the object language
has to be given by setting down (by fiat), firstly, that certain sentences of the language are to be axioms and,
secondly, the rules of inference by means of which the theorems are deduced from the axioms.
[6] Throughout this thesis I use double quotation marks around sentences, and single quotation marks around

language is true if and only if snow *is* white. This relation can be captured by the following metalanguage sentence: "Snow is white" is true if and only if snow is white. This is the truth paradigm—a paradigm embraced by Quine. These paradigmatic sentences (T-sentences) such as "'Snow is white' is true if and only if snow is white" instantiate Tarski's truth schema (Convention T),

(T) X is true if and only if p,

> where 'X' is replaced by the name of any sentence of the object language and 'p' by
>
> the sentence itself or by its suitable translation within the metalanguage.

Further, Tarski claims that any successful definition of the concept of truth is required to entail every instance of (T). In other words, given the successful definition of truth, we can prove a T-sentence for every sentence of the object language. This is the requirement of material adequacy, by which a definition of truth can be shown to get the extension of the truth predicate right. That is, this requirement can guarantee the defined concept to be the concept of truth rather than, say, the concept of falsity. Later Tarski states this in a looser way:

> Statements of this form [, the instances of (T),] can be regarded as partial definitions of
> the concept of truth. They explain in precise way, and in conformity with common usage,
> the sense of all special expressions of the type: *the sentence x is true*. (Tarski 1956*a*,
> p.404)

For Tarski, what is defined in (T) is not a general concept of truth but a *particular* concept of truth relative to a given language. That is, for every given object language L, there is a unique T-schema:

(T$_L$) X is true$_L$ if and only if p,

> by replacing 'X' with the name of any sentence of L and 'p' with the sentence itself

words and phrases.

or its suitable translation within the metalanguage.

Moreover, it is Tarski who precisely shows in his formulation of the proper way to define the concept of truth, that there is no significant way to provide a satisfactory account of the general concept of truth. However, even for a particular concept of truth relative to a given language, the metalanguage sentences such as "'Snow is white' is true if and only if snow is white" do not jointly constitute a complete analysis of the concept of truth. The complication of expanded languages under consideration, such as the possibility of the infinite number of sentences, makes the enumeration of all partial definitions or, in other words, all T-sentences impractical.

Although the number of sentences might be infinite, the number of the constituent parts of sentences, that is, names or predicates, is not. This prompts Tarski to further form an analysis of the semantic structure of sentences in terms of the semantic concepts of satisfaction and reference. By defining these two semantic concepts recursively at first, he provides a recursive definition of the semantic concept of truth. With the full set-theoretic machinery, the recursive definitions of satisfaction and reference can be converted into explicit definitions[7]. Following this, Tarski successfully forms an explicit definition of truth (for a given object language).

It is the general idea behind this procedure that renders it successful. The corresponding

---

[7] Briefly, there are normally at least three ways to define something. First, if the number of the terms or the concepts or the like in the definiendum is finite, then it can be defined *enumerativly*—that is, by providing a complete list of all the objects defined or all the expressions containing the objects defined. Or, if the number is infinite, the definiendum can be defined *recursively*—that is, by defining the possession of a property by a term x in terms of its possession by all the terms preceding x in a series, together with a definition of its possession by the first term of the series. Finally, all recursive definitions can be turned into explicit ones in terms of set theory and other suitable explanatory resources. Please cf. some instances in the following.

satisfaction schema with one free variable is as follows:

For all *a*, *a* satisfies the sentential function *x* if and only if *p*.

The instantiation of this schema with regard to the sentential function '*x* is white' can be put as follows:

For all *a*, *a* satisfies the sentential function '*x* is white' if and only if *a* is white.

Utilising the set-theoretic machinery of classes and sequences, the general schema for sentential functions (or open sentences) with more than one free variable can be characterised in a similar fashion as follows:

For all infinite sequences of classes *f*, *f* satisfies the sentential function *x*, if and only if

p,

where '*x*' is to be replaced by the name of any sentential function and '*p*' the function itself (with its free variables 'v_i', 'v_j' …replaced by 'f_i', 'f_j' …).

For instance, the sentential function consisting of 'is taller than', collocated alphabetically with *i*th and *j*th variables on both sides of 'is taller than', is satisfied by a given sequence if and only if the *i*th element in the sequence is taller than the *j*th element in the same sequence. To the extreme extent, any sentence in which all variables are bound by universal quantifier(s) is satisfied by any sequence with no regard to what its elements are; otherwise it is not satisfied by any of them. It is simply all or none. An explicit definition of truth is then given in terms of being satisfied by all sequences:

For every close sentence *a*, *a* is a true sentence if and only if *a* is a sentence and for every sequence of objects *f*, *f* satisfies *a*.

In terms of the account characterised above, it seems that Tarski successfully reduces the explication of the concept of truth to another semantic concept of satisfaction, specifying

the circumstances in which any singly given sentence can be said to be true or false—that is, satisfied by all sequences or else not at all. Whether or not a sentence is satisfied by all sequences is not the task of the semantic accounts of the concepts of truth or satisfaction. It depends on empirical science. However, Tarski's theory is interpreted in many different ways. In the following, I will discuss two major criticisms related to proper formulations of disquotational theories of truth respectively.

Etchemendy's criticism and Heck's response

One of the most enlightening observations comes from Etchemendy (1988). If we follow Tarski's procedure to define the concept of truth explicitly, then there is no way our semantic theory has any empirical significance. As a result, the definition does not capture the intuitive concept of truth. Tarski's explicit definitions of satisfaction and reference are enumerative. This does not seem to be an unreasonable strategy by itself since Tarski holds that the number of names and predicates of the object language is finite. However, according to Etchemendy, explicit definitions as such are stipulative. It follows that all T-sentences turn out to be logical truths, or at least necessary truths. This means a T-sentence, say, "'Snow is white' is true if and only if snow is white" does not have empirical content.[8] The definition does not give us any idea of the empirical content that the concept of truth in question might have. But explicit definitions of semantic concepts are indispensable for Tarski in order to guarantee the consistency of the semantic theory constructed in the metalanguage. Therefore, doing empirical semantics and providing a consistent semantic theory are not possible to achieve at the same time.

---

[8] Putnam, Soames and many others also hold a similar view.

In order to do empirical semantics, we should rather take the concept of truth as a

primitive, and formulate an axiomatic theory of truth for a given language. Every clause

of the recursive definitions of satisfaction, reference, and truth, are turned into axioms,

providing an extra axiom that the truth predicate of the given language, say, L—that is,

true$_L$—holds for all and only true sentences of L.[9] This axiomatic theory of truth also

satisfies the requirement of material adequacy. The difference is that T-sentences are no

longer logical or necessary truths because they are only derivable from those axioms, and

these axioms in turn are not without empirical content. Davidson argued in "The Structure

and Content of Truth" (1990) that Tarski's definitions of satisfaction and reference are not

necessarily stipulative. It depends on our interpretation of his intention. On the other hand,

Davidson does agree that we should drop the final step of Tarski's procedure of forming

explicit definitions, because Tarski's definitions only ensure that the extension of the truth

predicates is right, without further explicating what these predicates have in common; that

is, his definitions do not help us specify the meaning.[10] Nevertheless, the guarantee of

consistency of the target semantic theory must be abandoned if the final step of turning all

recursive definitions into explicit ones is dropped.

Adding to our axiomatic theory the above-mentioned new axiom that the truth predicate

of a given language L, namely, true$_L$, holds for all and only true sentences of it, is not a

threat to consistency, for the axiom gives no new properties to the truth predicate.[11] For

---

[9] This extra axiom is very likely to be indispensable for every deflationary theory of truth in one way or
another. For instance, Horwich concedes that, due to Gupta's observation, the extra axiom "(x) (x is true →
x is a proposition) is needed (Horwich 1998b, p. 23, note 7). This is implicitly acknowledged hereafter and
will not be brought up again.

[10] It is also clear why Davidson prefers to see the concept of truth as explanatorily prior to the concept of
reference, because he has no objections to Quine's inscrutability of reference thesis. This thesis, on the other
hand, is one of the major reasons for Quine's departing from Tarski's in accounting for the concept of truth.

[11] However, according to Davidson, adding this new axiom is also pointless, "For we can just as well
regard Tarski's truth predicate 's is true$_L$' as having the properties of our real-life predicate 's is true in L', as

Heck, on the other hand, there is something more. He further holds that there are parallel

axioms of satisfaction and reference. Heck calls these three new axioms the connecting

principles. These connecting principles reflect the requirement of material adequacy, and

the requirement that all the truth predicates of various successful definitions of the

concept of truth relative to a given language, are applicable to the same set of sentences

(Heck 1997, p. 541).

Etchemendy credits Tarski with the techniques he uses in his definition; that is,

> Getting from a Tarskian definition of truth to a substantive account of the semantic
> properties of the object language may involve as little as the reintroduction of a primitive
> notion of truth. (Etchemendy 1988, pp.59-60)
> [I]f we define a set TRUE using the standard recursive definition, then the claim that all
> and only true sentences of the language are members of TRUE is logically equivalent to
> [the axiomatic theory of truth]. (ibid., p. 59)

For Heck, this only suggests, "all the axioms of the axiomatic theory can be proven from

the explicit definition of truth and these new axioms, the connecting principles" (Heck

1997, p. 541). For instance, the axiom of truth,

true(A) if and only if (A is a sentence and $\forall \sigma.\text{sat}(A, \sigma)$),

can be derived from the explicit definition of truth,

A∈TRUE if and only if A∈SAT(A, σ)

and the connecting principles of truth,

true(A) if and only if A∈TRUE

and of satisfaction,

$\forall \sigma.\text{sat}(A, \sigma)$ if and only if A∈SAT(A, σ).

The consistency of the axiomatic theory is therefore not under threat. Heck points out:

---

long as those properties do not create inconsistencies" (Davidson 1990, p. 292).

> Since the connecting principles are little more than restatements of the conditions of
> material adequacy, it would be at best uncharitable not to credit Tarski with having
> proved as much. (p. 547).

The magic touch is of course the two new axioms added by Heck. He expands the

requirement of material adequacy to include the concepts of satisfaction and reference, so

that what are defined in the explicit definitions of reference and satisfaction are indeed the

intuitive concepts of reference and satisfaction. In addition to this, the distinction between

object-and meta-languages allows Heck to guarantee the consistency of the semantic

theory brought by the explicit definitions. Meanwhile, the criticism Etchemendy made is

resolved. Tarski's explicit definitions are not stipulative. For Heck,

> What [Tarski] has done is clearly to separate the mathematical from the empirical aspects
> of semantics: the mathematical part is wholly absorbed into the definition of truth; the
> empirical, into the claim of material adequacy.[12] (p. 543)

But this is problematic. First of all, consider Gupta's instance (2002), if a definition of

true-in-*L* appears to be "*x* is true in *L* if and only if *x* is not true in *L*", then it satisfies the

material adequacy requirement trivially because it also implies all T-sentences due to its

inconsistency.[13] Or we can give an inconsistent theory of truth with the following

biconditional:

  true(A) if and only if [A is a sentence and ($\forall\sigma$.sat(A, σ) & $\neg\forall\sigma$.sat(A, σ))].

This theory is also materially adequate. The connecting principle of truth thus can be false.

Since the connecting principle of truth is little more than the restatement of material

adequacy conditions, it is a material biconditional. This means it can be true when both of

its antecedent and consequent are false. Surely that an inconsistent theory plus the

---

[12] Heck thinks what really does the trick of avoiding semantic paradoxes is the distinction between object-
and meta-languages. But there is a good reason why appealing to the explicit definitions is necessary, viz.,
recursive definitions cannot prevent self-referential sentences.
[13] Gupta's instance is meant to reveal the fact that some restrictions have to be laid down for Convention T.
But it seems to also be relevant to the present discussion.

connecting principle, or really, anything, imply another inconsistent theory does not teach us any valuable lesson. The connecting principle of truth is too weak to derive Heck's conclusion.

Further, Tarski sets up two conditions for a definition of truth to be acceptable, that is, that it is formally correct and materially adequate. To make good sense of the motive for setting up two conditions instead of just the material adequacy condition, implies that there are definitions satisfying only the material adequacy condition. Otherwise, any definition satisfying the condition of material adequacy turns out to be formally correct. On Tarski's definition, however, there may be formally incorrect but materially adequate definitions of truth and they are implied by the explicit definition of truth and the connecting principle of truth as described by Heck. For instance, "$\underline{x}$ is true if and only if $\underline{x}$ is true or $\underline{x}$ is not true, and x" is a formally incorrect definition but it implies all T-sentences with the form "$\underline{x}$ is true if and only if x", where '$\underline{x}$' represents the name of a given sentence and 'x' the sentence itself.

For Gupta (2002), it turns out that Convention T is not even a satisfactory standard for evaluating the extensional adequacy of the definitions of truth, if the extensional adequacy of the definitions of truth amounts to their material adequacy. Gupta's argument rests on a very minimal assumption that, for any truth definition D for a language L, there is at least a true sentence p of L, say, "John Howard is the prime minister of Australia in 2004," that is not logically implied by the necessary apparatus, called AD, needed by D to derive consequences. Moreover, 'true in L' does not occur in both p and AD. Suppose D, "For all objects z, z is true in L if and only if $\varphi(z)$" (Where $\varphi(z)$ means that z is true of φ),

is an extensionally adequate definition of truth. Then D', "For all objects z, z is true in L if and only if (if p then $\varphi(z)$))" is also an extensionally adequate definition. It follows, however, that D' in conjunction with AD fails to imply all T-sentences of L. Suppose, for the sake of *reductio*, D' and AD implied all T-sentences, we can infer that they imply the following: not every sentence of L is true. But definition D' is particular in the way that it, AD, and not-p logically imply that every sentence of L is true. This suggests that they are logically inconsistent. Hence, D' and AD logically imply p. Since AD does not imply p, there is at least one interpretation on which all members of AD are true but p is false. If we assign the set of all objects to 'true in L', then we establish an interpretation on which D' and all the members of AD are true but p is false. This means D' and AD do not logically imply p. But this contradicts the previous conclusion that D' and AD logically imply p. So the *reductio* completes. Material adequacy is not sufficient for an account of truth.[14]

I agree with Davidson, however, that Tarski's theory does not enlighten us about the meaning of the concept of truth, so that the final step that converts the recursive definitions into explicit ones has to be dropped. Moreover, it is not that Etchemendy's analysis is preferable to Heck's, but rather that according to Quine's thesis of inscrutability of reference, the concept of truth is more basic than the concepts of reference and satisfaction. The thesis of inscrutable reference teaches us that the reference

---

[14] The current discussion is also relevant in connection with an acceptable formulation of disquotationalism. Disquotationalism in its simplest form takes all the T-sentences as its axioms. The undesirability of this formulation is its lack of deductive power. Some disquotationalists such as Halbach, propose to include all those axioms that are formed by the original clauses of the recursive definitions of satisfaction and reference in the disquotational theory, so as to help derive some important generalisations. If Heck is right, then this sort of axiomatic theory is, absurdly, deducible from the explicit definition of truth with the help of the corresponding connecting principle.

scheme of any given language or theory is inscrutable.[15] That is to say, there are no

empirical pieces of evidence that confirm any particular reference scheme over the others.

Among others, a reference scheme is equally good for a given language or a given theory

if it can preserve the distribution of the truth-values of all their expressions. If truth set

limits on the accounts of reference and satisfaction, it does not seem right to hold them

explanatorily more basic. A similar point will be made in the following discussion of

Field's important criticism of Tarski's theory.

Moreover, in spite of all the clarity and efficiency, Quine disagrees with the approach of

explicating the concept of truth in terms of the analysis of the sentential structure of any

given sentence. The immediate reason for Tarski to define the concept of truth in

accordance with the recursive definition of the concept of satisfaction is the impossibility

of forming a recursive definition of sentential truth outright. In "Truth and Disquotation"

(Quine 1976), Quine demonstrates that constructing the truth definition in the light of the

notion of infinite sequence and the recursive definition of satisfaction, is incidental to the

apparatus of quantifications and variables employed by Tarski in the meta-language. In a

different kind of logic in which the work of quantifiers and variables are managed by

constants, it can be shown, Quine argues, that the complication resulting from other

semantic concepts is dispensable (p. 309).

Field's physicalist criticism and the explanatory priority

As a physicalist, Field regards Tarski's theory as a project to reduce semantical terms to

non-semantical or physical ones. The first reduction appearing in Tarski's theory is from

---

[15] One of the consequences of abandoning the analytic/synthetic sentences distinction is the continuum of
language and theory.

the semantic concept of truth to the semantic concepts of satisfaction and reference. The

second step is to reduce these semantic concepts to non-semantical or physical concepts.[16]

And it seems that Tarski successfully employs non-semantic terms such as 'sequence' to

prepare the second reduction. Field's criticism is that Tarski's theory fails to provide a

non-trivial second reduction. He argues in "Tarski's Theory of Truth" (Field 1972) that

Tarski's account of truth has no significant difference in explication from a similar

account in which the reference scheme of the object language is in terms of *primitive*

denotation. That is, "every name *primitively denotes* what it denotes; every predicate and

every function symbol *primitively denotes* what it applies to or is fulfilled by; and no

complex expression primitively denotes anything" (p. 6). Field's point is that this

indifference makes the second reduction trivial.[17] As a physicalist, it seems natural for

Field to criticise on the ground that a non-trivial reduction is necessary for a successful

physicalist account of truth. (Notice though, that the same criticism could be made by a

non-physicalist who disagrees with the physicalist account.) This criticism helps to make

sense of why Field concludes in part that a failure of this second reduction suggests the

abandonment of the concepts of truth and reference. For such concepts are not, on this

account, qualified according to the criteria set up by physicalism (p. 25). In the

"Postscript" (2001) of "Tarski's Theory of Truth", Field still maintains a refined version

of his criticism. He does, however, concede at the end that given his interpretation of

Quine's account, should it turn out to be correct, then Field's own physicalist account of

truth is mistaken. For if Quine's account is correct, then the truth predicate is just a device

for expressing generalisations—that is, the truth predicate has its special merit as

---

[16] Field makes it clear in his paper that it is not necessary to require a reduction in a strict sense. An
approximate reduction will do.
[17] Many philosophers arrive at different conclusion. Devitt, for instance, holds that the list-like definitions
are deflationary. "Tarski's definition tells us a lot about 'true-in-<u>L</u>'," Devitt says, "It tells us nothing about
truth-in-<u>L</u> because it is implicitly committed to the view that there is nothing to tell." (2001)

abbreviating infinite conjunctions or disjunctions of sentences.[18]

According to Romanos (1983), however, this is not the only problem that Field's theory faces. Tarski's semantic account of truth results from attributing, or failing to attribute, truthfulness to sentences, in terms of the analysis of the sentential structure of the object language sentences as well as the list-like definitions of reference and satisfaction. Field's criticism rests on the point that Tarski fails to provide a satisfactory explanation of the second aspect of this process. Romanos correctly observes that the consequence of this failure—that is, the abandonment of the concepts of truth and reference because of their not being qualified according to the criteria set up by physicalism—is acceptable only if the original assumption that the analysis of the referential relation plays a decisive role in explicating the concept of truth, is itself without a shred of doubt (p. 170). But if one contests that the concept of truth should not be explicated in terms of the referential relation, then Field's criticism is no longer decisive.

Romanos continues his criticism by citing Field in order to show that Field employs Quine's theses of inscrutability of reference, under-determination of theories, and ontological relativity, in order to support the view that gluing our conceptual scheme onto reality from the outside is impossible (Romanos 1983, p. 168; Field 1972, p. 24). For Quine, these theses seem to further entail the implausibility of asking for causal theories of reference to nail language to reality (Quine 1960, p. 24). But according to Field the causal connection between language and (extralinguistic) reality is still required in order to avoid our conceptual scheme breaking down from the inside (Field 1972, p. 24). It

---

[18] Please see the following section for a more detailed elaboration.

seems to Romanos that Field is in danger of being inconsistent. However, Field in fact

disagrees with Quine. It is not clear with which thesis of Quine's Field disagrees with in

Field 1972. Is it the implausibility of asking for causal theories of reference or is it one of

the theses mentioned above? He does agree that our reflection of the accounts of primitive

reference and truth is not explanatorily prior to our scientific investigation.

In "Quine and the Correspondence Theory" (Field 1974), written two years later, Field

denies Quine's thesis of ontological relativity and proposes an account of partial

signification to remove the obstacle set up by the thesis of inscrutability of reference.[19] If

this position is established, Field's attempt to find a causal theory of reference is not that

puzzling. In the next chapter, it will be shown that Field's criticism of the thesis of

ontological relativity is a consequence of his misinterpreting Quine's ontological

relativity thesis.

This discussion centres around which concept is explanatorily more basic, reference or

truth. It is clear why Quine takes the concept of truth to be more basic to the concepts of

satisfaction and reference. Davidson pushes this point slightly further. Normally our

understanding of sentences of a given language, presumably infinite in number, is

dependent on our prior understanding of the words, finite in number, that form these

sentences. Thus the semantic properties of words seem to be conceptually prior to those

of sentences because the latter owes their comprehensibility to the former. According to

Davidson, it is a mistake to draw such a conclusion about the conceptual priority of the

---

[19] Although there is no empirical ground to choose one translation manual over another and thus relativise
the reference scheme of a given theory or language to another, for Field, it can be held that a general term x,
for instance, of a given language *partially* signifies the sets of objects of all competing manuals. A
simplified example can be put like this: the term *gavagai* partially signifies the set of rabbits (relative to a
translation manual M_1) and the set of undetached rabbit parts (relative to M_2) but fully signifies neither.

semantic properties of words over sentences, from a platitude about how we can possibly understand all sentences. That is,

> This mistake is to confuse the order of explanation of why the theory is correct. The theory is correct because it yields correct T-sentences; its correctness is tested against our grasp of the concept of truth as applied to sentences. Since T-sentences say nothing whatever about reference, satisfaction, or expressions that are not sentences, the test of the correctness of the theory is independent of intuitions concerning these concepts. Once we have the theory, though we can explain the truth of sentences on the basis of their structure and the semantic properties of the parts. (Davidson 1990, p. 300)

Quine and Davidson, of course, do not preclude there being a role played by a theory of reference in the explication of the truth of sentences. Still, the difference between the reference schemes suggests a difference between the translation manuals, and one of the reasons for selecting one reference scheme over the other is its capacity for preserving the distribution of true and false sentences in the interpreted language. Truth preserving sets limits on the selection of the domains of discourse. To regard the theory of the referential relation between the constituents of sentences and the reality as explanatorily basic is nothing but a misconception.

Additionally, Tarski's semantic analysis is either inconsistent with, or at least, is a peculiar companion for Quine's thesis of ontological relativity. The thesis is, put simply:

> What makes sense is to say not what the objects of a theory are, absolutely speaking, but how one theory of objects is interpretable or reinterpretable in another. (Quine 1969, p. 50)

To construct an account of sentence truth in terms of satisfaction conditions is to base the analysis of the truth of a sentence on a function of the referents of its component terms (Romanos 1983, p.145). This view raises concerns about the relativity of ontology in a sense that a primitive concept of truth does not (ibid.). If the explicit definitions of

satisfaction and reference are stipulative, some concerns with regard to the relativity of ontology have to be raised. For these stipulative definitions do offer a way to talk about what the objects are absolutely.[20] Romanos puts this well:

> If truth defined by Tarski's procedure is essentially a function of the features of a given theory's universe of discourse—the objects comprising its ontology—then how can our understanding of 'true' as applied to sentences of the theory be any less relative than our description of the theory's universe. (ibid., p. 159)

2. Formulating Quine's Disquotational Theory

Intuitively speaking, to say that a sentence is true is to say that it has a certain relation to reality, and that it is in virtue of this relation that it is true. This is the correspondence theory of truth in a very loose sense. What seems to be widely accepted by many philosophers is that the correspondence theory of truth is about the correspondence relation between the meaning of a given sentence, or the proposition expressed by it, and some fact of the matters or state of affairs. Still, for the correspondence theory to be more philosophically interesting it must be supported by a theory of meaning strong enough to allow two sentences, whether belonging to the same language or not, to bear some synonymous or equivalent relation. It is absurd to claim that the basic unit of meaning is not sentences and at the same time to hold that the uncritically accepted synonymous relations between sentences from different languages make sense.[21] It is clear then that the advantage of the correspondence theory of truth, under the normal understanding of it,

---

[20] Worse, there is indeed a reference scheme for every language in an absolute sense if we accept Heck's interpretation of the materially adequate conditions that includes the consideration of reference scheme in the requirement of material adequacy.

[21] Eventually the consequent will strike back to undermine this very assumption. To hold that there is a genuine discrepancy between the synonymy relations for a language and for different languages or between sentences of mine and sentences belonging to a language that is other than my mother tongue, is indeed to object to the public feature of languages—that is, the promotion of the private language account. Worse, this assumption has to offer some sort of explanation to demarcate the learning processes of my mother tongue and the second language, which is definitely not an easy job.

lies in its capacity to provide a satisfactory explication of how sentences from different languages which share the same truth value, can also share the same *meaning*. This explication depends heavily on the foggy distinction between facts or the like, and how sentences articulate them. According to this account what correspond to reality are of course not sentences themselves but their meanings. Thus, for instance, to say that a sentence "Snow is white" is true is to say that its meaning that snow is white corresponds to the fact that snow is white.

Quine thinks that it is here that we encounter the inexplicable part of the correspondence theory. To what extent can the notion of sentential meaning, the resultant notion of a relation of synonymy between sentences, and the notion of a fact of the matter, be clarified? One might propose to use the notion of a proposition to explicate the notion of sentential meaning. The sentential meaning of "Snow is white," for instance, is the proposition that snow is white. Similarly, the meaning of "Der Schnee ist weiss" is also this very same proposition. Truthfulness should be attributed to propositions instead of sentences. Therefore, sentences themselves become strings of physical signs of different notations in which propositions, their meanings, are couched. The intuitive account of the correspondence theory in the last paragraph suggests a two-step analysis of the notion of truth. That is, to say that a sentence "Snow is white" is true, is to say that its meaning that snow is white corresponds to the fact that snow is white. The first step is from sentences to their meanings and the second step is from their meanings to the facts to which they correspond. Utilising the notion of a proposition, we no longer seem to recourse to the first step in this process. The problem with this account seems to be that it is open to Quine's criticism of the myth of the museum. For it seems that propositions are

uncritically accepted as exhibits, and sentences from different languages are their labels. More importantly, accepting propositions as abstract entities introduces the problem of how to individuate such propositions.[22] The sentence "Snow is white" does not individuate the proposition that snow is white because it is just one of (infinitely) many notations of this proposition. Each notation of the proposition is nothing but a string of physical signs. The problem of individuation leads us back to the original consideration of equivalent or synonymous relations between sentences.[23] Quine's main criticism on this approach is as follows: "[T]he appropriate equivalence relation makes no objective sense at the level of sentences" (Quine 1970, p. 3). This criticism will be discussed in more detail later.

As described above, an adequate account of the notions of sentential meaning and synonymous or equivalence relations between sentences is vital to a successful explication of the correspondence theory of truth. Quine's criticism, if established, has a severe impact upon any such explication. It does not just show that the correspondence theory fails to provide an acceptable account of equivalence or synonymous relations. It indicates the dim chances of getting a plausible account of sentential meaning. It is not compelling to both hold that the equivalence or synonymous relations do not make objective sense at the sentential level, whilst also holding that the notion of meaning does make objective sense at the level of sentences. Consequently, the correspondence theory of truth loses ground.

---

[22] Horwich's minimal theory of truth takes propositions as the primary truth bearers. In Chapter Five, where the evaluation of this theory is given, however, Quine's objection to the notion of a proposition does not play any essential role.
[23] This, on the other hand, seems to shed some doubt on the rationale of introducing the notion of a proposition to play such an essential role in the explication of the notions of truth and sentential meaning.

It is rather puzzling why Tarski would declare that his theory is a kind of correspondence theory of truth, despite the fact that for Tarski there is no correspondence relation between the meaning of sentences on the one hand, and facts or states of affairs on the other. Davidson made it clear in "The Folly of Trying to Define Truth" (Davidson 1996):

> But facts or states of affairs have never been shown to play a useful role in semantics, and one of the strongest arguments for Tarski's definitions is that in them nothing plays the role of facts or states of affairs…Tarski's truth definitions make no use of the idea that a sentence 'corresponds' to anything at all. (p. 267)

Indeed, the material adequacy condition ensures that every definition of truth satisfying this requirement is successful. But no part of this analysis has anything to do with corresponding to facts or states of affairs. The analysis of the sentential structure of any given sentence does not even appear in the formulation of this requirement. Moreover, if one construes Tarski's position as taking the satisfaction function to play the role of the so-called correspondence relation, just as Davidson admits that he once held this position[24], then the correspondence theory is, as Davidson later puts it[25], thoroughly unintelligible. According to Tarski's definition, every closed sentence is true if and only if it satisfies every sequence. It follows that every true sentence corresponds to exactly the same thing, the *big* fact. This makes the Tarskian conception of the correspondence theory useless.

An introduction of Quine's reasoning on disquotational conception of truth

The intuitive correspondence theory does seem right in the sense that truth should hinge on reality rather than language. The employment of the notion of a proposition is one prominent way to account for this point. Quine agrees that the truthfulness of sentences

---

[24] Davidson 1990, p. 302, foot note 36.
[25] Ibid., p. 298. Also Davidson, 1996, p. 268.

depends on reality to make them so (Quine 1970, p. 10). For the truth of singly given

sentence he states:

> In speaking of the truth of a given sentence there is only indirection; we do better simply
> to say the sentence and so speak not about language but about the world. So long as we
> are speaking only of the truth of singly given sentence, the perfect theory of truth is what
> Wilfrid Sellars has called a disappearance theory of truth.[26] (ibid., p. 11)

According to the correspondence theory of truth, in speaking of the truthfulness of a

given sentence we are saying that the meaning of the sentence corresponds to a fact, a

state of affairs, or the like. Even putting aside the fact that the concept of a

correspondence relation is unclear, it will be difficult to specify the notions of sentential

meanings and facts or the like without utilising any linguistic framework. The only refuge

for the correspondence theory is the direct relation between sentences and experience, at

least in the broad sense. Quine's criticism on this matter shows that a new way of

thinking about the notion of truth is required. And Quine's proposal on the truthfulness of

any singly given sentence is similar to Sellars' disappearance theory of truth. The passage

cited above is probably the most important reason why many interpret Quine's

disquotationalism as well as most disquotational theories of truth, as no different to the

redundancy theory of truth in the cases of singly given sentences. This will be discussed

in Chapter Five. However, this passage is a bit misleading anyway.

This is not the whole story though. Quine is not an advocate of the redundancy theory of

truth. His position, on the contrary, reveals the unsatisfactoriness of this theory. For Quine,

the truth predicate is not redundant and thus has its special function in languages. He

---

[26] The truth predicate is needed for sentences that are not given, such as 'What Smith said is true' (Quine 1990, p. 80). In the cases similar to this, the truth predicate is not applied to a quotation but a pronoun or bound variable (ibid.). However, the disquotational schema, which will be formulated in a latter discussion, is still applicable to these contexts through some technical paraphrasing. For instance, Paul Horwich in *Truth* (p. 21).

expresses his point as follows:

> Truth hinges on reality… Where the truth predicate has its utility is just those places
> where, though still concerned with reality, we are impelled by certain technical
> complications to mention sentences. Here the truth predicate serves, as it were, to point
> through the sentence to reality; it serves as a reminder that though sentences are
> mentioned, reality is still the whole point. (Quine 1970, p. 11)

Certain 'technical complications' are about the necessary grammatical devices by which
we are allowed to generalise on sentences with the same feature or to generalise on
sentences in the case of blind ascription such as "What Tarski said on April first, 1953 at
11 a.m. is true". It is always easy to generalise on sentences like "Quine is mortal,"
"Ramsey is mortal," and so on, to form a generalisation "All men are mortal" without any
mention of truth. It is in contrast not that simple to obtain a generalisation on sentences
such as "Snow is white or snow is not white" and "Grass is green or grass is not green"
without employing some dummy symbol and the truth predicate in the generalising
sentence "Every sentence of the form 'p or not p' is true" (ibid.). In other words, the truth
predicate might be a device for abbreviating or expressing the (possible) infinite
conjunctions or disjunctions.[27] The truth predicate is thus the grammatical device for
semantic ascent— we can talk about reality by talking about sentences instead.

There is a difference between generalising "All men are mortal" from "Quine is mortal"
and generalising "Every sentence of the form 'p or not p' is true" from "Snow is white or
Snow is not white". In the first case what are changed from instance to instance are names.
It can thus be read as "x is mortal for all men x" in which x is the sort of things that
'Quine' is a name of. In the second case it would, similarly, be read as "p or not p for all

---

[27] Although Quine does not say this explicitly, it is compatible with his account. Moreover, not every
infinite conjunction is expressible in terms of the truth predicate. This will be briefly discussed in the third
section and more completely in the first section of Chapter Five.

things *p* of the sort that sentences are names of". But sentences are not names. So it is inevitable that we will need to utilise the truth predicate to allow semantic ascent to talk about sentences instead of names. There seems to be another option though. It can be claimed that sentences are the names of propositions. So the target generalising sentence can be read as "*p* or not *p* for all propositions *p*". Apart from all those problems discussed above, propositions in this particular context do not bring any benefit that the truth predicate cannot provide (Quine 1970, pp. 11-12). The special function played by the truth predicate in semantic ascent is unmistakable in the following message:

> This ascent to a linguistic plane of reference is only a momentary retreat from the world, for the utility of the truth predicate is precisely the cancellation of linguistic reference. The truth predicate is a reminder that, despite a technical ascent to talk of sentences, our eye is on the world. (ibid., p.12)

For Quine, Tarski's paradigmatic schema on all T-sentences explicates this characteristic explicitly. There is a major difference between Tarskian and disquotational T-sentences. For Tarski, T-sentences include instances such as "Morning star is the star last seen in the morning" is true if and only if Venus is the star last seen in the morning. For the right hand side sentence is an extensional equivalent of the quoted sentence. But disquotationalism does not allow instances like this. Henceforth in this chapter and the following, I will discuss T-sentences, or better, disquotation sentences, in this straight disquotational sense. For a disquotation sentence—"Snow is white" is true if and only if snow is white—the quotation is a name of a sentence containing a name. Calling "Snow is white" true is meant to call snow white. The truth predicate is a devise of *disquotation*. In "Notes on the Theory of Reference" (Quine 1953*a*) Quine proposes that the theory of truth should be an extra part parasitic to the language for which the theory is constructed.

Although Quine does not say it explicitly, the content of the quoted sentence and the sentence itself has to be without difference. It is not necessary for us to specify *the* (empirical) content of the given sentence as long as we restrict the interpretation or, in other words, the translation manual of the target language and the disquotational theory, such that it is the same. On the other hand, in a case when we are affirming an infinite set of sentences, then talking about reality outright is out of the question. Indeed the truth predicate has its characteristic as a function of semantic ascent to restore objective reference. It is because of this characteristic we can *affirm* an infinite conjunction by affirming the corresponding generalisation.

A preliminary set-up of a proper formulation

The first question to consider is whether or not the disquotational theory respects Tarski's distinction between object- and meta-languages. And the answer is no. This, however, seems to be an advantage rather than a flaw. Many disquotationalists, Halbach, for instance, take the aforementioned distinction as a problem rather than a requirement of a successful theory of truth. Tarski's proof that the semantic theory of any semantically closed language within which the theory is constructed is destined to be inconsistent, only shows that the theory of truth of a target language cannot be constructed within the same language. It does not show that the truth theory has to be constructed in a much stronger meta-language. To resort to a stronger meta-language is not always achievable—it depends on how strong the object language is. In some cases, this might cause a problem.

If the disquotational theory of truth is not constructed in a meta-language—that is, a meta-language relative to an object language, then the relation between the target

language, call it the base language, and its theory of truth has to be specified.[28] Call the

language that forms the truth theory the truth language. Then the main difference between

the truth language and the base language is that the former has the additional truth

predicate.[29] There is of course another constraint that we must impose. The base

language does not contain any vague, ambiguous, indexical, or self-referential

sentences.[30] As long as this is a general problem for every deflationary theory, it is not a

particular problem for my formulation of Quine's disquotational theory.[31] Since I do not

adopt Tarski's theory of explicit definitions of truth, I should further characterise the

standard disquotational theory as embracing the disquotational schema "'*p*' is true if and

only if *p*" as an axiom schema, and all its instances as axioms. Whether or not there

should be some additional axioms is left to the discussion in the following section. It is

clear then, that the truth predicate has as its syntactical function cancelling the quotation

marks and cashing out exactly the quoted sentence itself.[32] Although Quine does not

specifically carry out his account in this way, I think there should be no significant

disputes over this procedure. However, there is at least one other way to formulate the

disquotational theory—namely, a disquotational *definition* of true sentences.

---

[28] The terminology 'base language' is borrowed from Halbach.
[29] This is consistent with Quine's own words. In the subsection discussing Quine's reasoning regarding the disquotational concept of truth, I showed that Quine, in his paper "Notes on the Theory of Reference", considered the theory of truth as an extra part, parasitic on the target language.
[30] The main task of this dissertation is to first set up a formulation of Quine's disquotationalism and a brief introduction of Horwich's minimalism, and then to examine their performance in accommodating some general issues for deflationary theories. For further discussions of this constraint, please cf. Chapter Five.
[31] After all, all deflationists are inclined to see their theories of truth as applicable to sentences of languages that are free from the threat posed by these problems. Besides, it is a bonus if a theory of truth is capable of handling these problems nicely. But this alone is not a good reason to prefer one theory over the others.
[32] As indicated before, the quoted and disquoted sentences have the same interpretation. Moreover, the disquotational theory cannot just consist in some axioms instantiating the disquotational schema. This will be further clarified in the following.

Disquotational definitions of true sentences

The disquotational schema cannot be turned into a definition simply because objectual

quantifications do not quantify into quotation marks. However, substitution

quantifications can. The problem of utilising a substitution quantification in the

disquotational schema "$\prod p$('p' is true if and only if p)", where the symbol '$\prod$' represents

the universal substitution quantification, is that, normally, a substitution is successful only

when the resultant sentence is *true*. I should spend some extra space to emphasis this, for

some philosophers seem to underestimate the consequence of utilising substitution

quantifications in explicating truth. Alston, for instance, in *A Realist Conception of Truth*

uses universal substitution quantification in its traditional sense to formulate a general

account of propositional truth—that is,

$\prod p$(The proposition that p is true *iff p*) (p. 28).[33]

Further, he is fully aware of the criticism of utilising it in defining truth. However, he

thinks that this is absolutely fine. The three points he made are as follows: first, universal

substitution quantification is only *used* but not *mentioned* in the schema, so it is not clear

that truth must be mentioned to give someone an understanding of this schema; second,

and more fundamental according to Alston, this schema is not meant to be a *definition* of

truth, and hence there is no danger of the kind whereby a definition is invalidated by

circularity; third, and equally important according to him, this schema does not aim to be

reductive, to reduce 'true' to terms at a more fundamental level (p. 29). What Alston

misses, it seems, is that whether or not this schema is meant to be a definition of truth is

not so relevant to the circularity criticism. Even if this schema is not a definition, the *truth*

theory in which it is embedded is nevertheless circular. So one should be more prudent

---

[33] I change the symbol of universal substitution quantifier he used to suit ours.

when using substitution quantifications to explain the concept of truth.

But if the main purpose of the truth predicate is to abbreviate infinite conjunctions or disjunctions, then substitution quantifications can be taken as explicating this particular feature. In other words, substitution quantifications can be used not to quantify over the left hand side of the disquotational schema but the right hand side. In "The Deflationary Concept of Truth" (Field 1986), for instance, Field suggests that the disquotationalist should utilise substitution quantifications in their formulation of truth to capture the (possibly infinite) list of true sentences. The explication of the substitution quantifiers, following this line of reasoning, is as follows:

(SQ) The notation '$(\sum p)(\ldots p \ldots)$' encodes the disjunction of the sentences (open or
   closed) that result from replacing the free occurrences of 'p' in '$(\ldots p \ldots)$' by e;
   The notation '$(\prod p)(\ldots p \ldots)$' encodes the conjunction of the sentences (open or
   closed) that result from replacing the free occurrences of 'p' in '$(\ldots p \ldots)$' by e. (p.
   99)

If this is acceptable, then it seems possible to attain a disquotational *definition* of truth. Inspired by Field's reasoning on substitution quantifications, David in *Correspondence and Disquotation* (David 1994) provides a thorough analysis of the possible candidates of a disquotational definition.

David begins with the modified version of Tarski's schema (T) or, in other words, Convention T—that is, 'p' is true if and only if p—and proposes a simplified definition of disquotational sentence truth by replacing the 'if and only if' in (T) by definition sign '$=_{DF}$'. The definition is as follows:

(D*) '*p*' is a true sentence =_{DF} *p*. (David 1994, p. 64)

This simplified version has quite a few problems, such as the difficulty of forming a reading of it in terms of ordinary terms. Among the problems, the most important two are raised by Tarski. (D*) is not a *complete* definition of truth, according to Tarski, because it characterises the phrase 'true sentence' only within certain types of context. The idea is that the truth predicate is not just designed to attach to quotations of sentences, but to attach in some other contexts to a specific sentence picked out by a definite description, or to sentences with universal or existential generalisations over them. In other words, this definition cannot account for blind ascription or what Quine calls oblique generalisations[34]. These include sentences such as: "The sentence written by Tarski is a true sentence", "Every sentence in Tarski's book is a true sentence," or "The sentences with the feature '*p* or not *p*' is true". (D*) is not *eliminative* either, according to Tarski, since one cannot eliminate the phrase 'true sentence' from all contexts in which it occurs on the basis of the definition alone.[35] David argues correctly that the deficiency of (D*) is that it only provides us conditional information about truth. One knows how to employ the notion of disquotational truth on every given occasion only if one knows (D*) *and* all other relevant facts that help us sort out which sentences are picked out by which descriptions and which sentences are in the range of which quantificational phrases (p. 72). This is a fatal deficiency for a disquotational definition of truth because it would render the concept of truth substantive. The definition of disquotational sentence truth, then, has to employ substitution quantifications and the identity function to resolve these

---

[34] That is, generalisations on sentences with the same feature.
[35] Etchemendy (1988, pp. 53 f.) points out a formal reason for preferring eliminative definitions to noneliminative ones—a reason which David thinks must play a role in Tarski's consideration (David 1994, p. 71). The reason is that an eliminative definition of a term would automatically ensure consistency relative to the theory on the basis of which the term is defined. That is, an eliminative definition cannot lead to paradoxes, unless they are already derivable without the defined term, because this definition allows us to do away with the definiendum in favour of its definiens.

problems. The major definition according to David is the following:

(D) x is a true sentence $=_{DF}$ for some p, x is identical with 'p', and p (p. 74); or

   x is a true sentence $=_{DF}$ $(\sum p)(x = $ 'p' & p) (p. 89) .

The corresponding generalisation implied by (D) according to David is (GD):

(GD) (x) (x is a true sentence $\leftrightarrow$ $(\sum p)(x = $ 'p' & p)). (ibid.)

David then proves that schema (T) is entailed by (GD).

Do sentence types as the truth bearers create a modal problem?

A question for disquotational theories of truth in general is to what exactly is 'true'

applicable? Is it sentence types or sentence tokens? Many, if not most, advocates of

disquotationalism incline to choose types over tokens.[36] And David thinks that this

causes a modal problem for the instances of the disquotational schema. But I have to

detour a bit because, unlike others, Quine takes sentence tokens as the primary bearers of

truth.

As discussed at the beginning of this section, it is untenable to develop accounts of the

notion of sentential meaning, or of a fact of the matter, without the involvement of any

given linguistic framework. Equally, it is only by reference to visible or audible sentences

that we may comprehensibly say which sentence meaning or proposition we have in mind

when uttering or writing a sentence (Quine 1990*b*, p. 77). For Quine, the truth vehicles

are *eternal sentences*—that is, eternalised declarative sentences—in which indexical

terms are supplanted by names, addresses or other identifying particulars as needed, and

the tenses of sentences are dropped by instead using exact time or related predicates such

---

[36] Except Field. Because of his adherence to physicalism, for him it is sentence tokens to be applied by the
truth predicate.

as 'earlier than' to fill in (p. 78). So these eternal sentences do not include most utterances in everyday discourse that are deemed to be either true or false, because in general such utterances are not thus refined.

The only linguistic expressions which physically exist are those uttered or inscribed. These utterances or inscriptions are *tokens* of sentences or other linguistic expressions. The issue here is whether eternal sentences, as the vehicles of truth, are taken as tokens or types of sentences. Sentence types are inevitably abstract; if this is not the case there is no way to explain the possibility of linguistic behaviour in the first place. The abstract character of sentence types does not suggest that sentence types amount to the sets of sentence tokens. Otherwise the types of never uttered or inscribed sentence tokens are the empty sets, and thus these types are identical to each other. Still, the abstract character of sentence types should not worry us too much, for this is a general feature of every semantic theory of linguistic behaviour. However, one of the important features of a sentence type is the sameness in the physical appearances of its tokens, regardless of their meaning or meanings. For the one and the same physical string of sounds or signs can be two distinct tokens of different languages, or different evolving stages of one language. Consequently, according to Quine (1970, p. 14), truth or falsity should be ascribed to eternalised sentence tokens rather than types of a particular language at a particular time. But this is not very promising. Sentence tokens have physical existence. It follows that a sentence is never true if it has never been uttered or inscribed. Further, 'A' and 'B' might be both uttered or inscribed true sentence tokens, but their conjunction 'A and B' might fail to be a true sentence token if no one ever uttered or inscribed it. Worse, "Snow is white" appears on the two sides of its disquotation sentence "'Snow is white' is true if and

40

only if snow is white" as two distinct tokens. Hence it does not seem right to claim that we substitute talk of a token for talk of another token used to report or describe reality. So it is sentence types rather than tokens that are the truth bearers. The eternalisation of sentence types can also be dropped as long as we preclude vague, ambiguous, indexical, self-referential, or non-truth-apt sentences as part of the target language. Fixing an interpretation of the target language can help us to set aside the above-mentioned features of sentence types.

Recall that according to David, the disquotational schema (T) implied by the disquotational definition of truth (D) has a serious problem with regards to its modal status. All the instances of (T) should be necessary.[37] But according to the correspondence theory they are contingent. Consider, for instance:

(1) 'Snow is white' is true iff snow is white.

(1) is not necessarily true, because 'snow is white' could have a content other than the one that snow is white. If 'snow is white' had its content as snow is green and snow were indeed green, then the sentence 'snow is white' would have been true, but (1) would not have been true. On the other hand, if snow were white, just like it actually is, but 'snow is white' had its content as snow is green, then 'snow is white' and (1) would have both been false. The problem is that the disquotational schema cannot get the counterfactuals of its instances right. The gravity of this objection consists in the fact that the contingency of (1) indicates the contingency of the schema (T) and thus the contingency of the definition

---

[37] The plausibility of intensional concepts is sensitive to authentic Quinean philosophers. However, even Quine considers the relation between 'p is true', where p is the name of 'p', and 'p' should be something very similar to necessary equivalence. He calls it cognitive equivalence and Field follows him. There is no need to involve Quine's rejection of the concept of necessity here. It will suffice to consider the possibility that the relation between 'p is true' and 'p' could be derived from both the usual modal account and a separate one that might be favoured by Quine.

that implies it. The only possibility of deriving the contingency of the schema (T) from a

necessary truth, according to the correspondence theorist, is to rest on a contingent

premise concerning the contingent relations between sentences and their contents or the

like. The correspondence or representation theory of sentence-truth can provide a good

explanation of the contingency of the schema (T), whereas the disquotational theory has

certain difficulties meeting this challenge.

One way to meet the challenge, according to David, is to contest the assumption that

every genuine definition has to be necessary. Disquotationalists might argue that the

above-mentioned requirement is only plausible when applied to those definitions that

attempt to reveal the *nature* of their definiendum. Thus this requirement is not applicable

to the disquotational definition of sentence truth. This, however, cannot be the final

answer from the disquotationalist, since she still does not offer any account of the modal

status of her definition. If the corresponding generalisation entailed by (D) is just a

material biconditional, then the disquotationalist owes us an explanation of why (D) is

better or more preferable than the following definition of sentence truth: *x* is a true

sentence $=_{DF} (\sum p)(x =$ 'p' & p & (there are pigs or summers in Austria are much warmer

than in Tucson)) (David 1994, pp. 131-2). The answer is not too difficult to get, of course,

for the truth predicate is a device only for abbreviating *relevant* lists. But (D) still should

have an interpretation stronger than just coextensionality.

According to David, the solution can be found in McGee's *Vagueness, Truth, and

Paradox*. Let M be my *idiolect* that by definition is the set of sentences that are

intelligible to me. The reformulation of (1), it can be claimed, is necessarily true:

(2) 'Snow is white' is true in M iff snow is white.

David then claims: "*[I]n the language that I actually speak*, it is necessary that 'Snow is white' is true iff snow is white. But this does not entail that there isn't also a reading of (1) under which it is contingent" (David 1994, p.133). In other words, if I actually speak a different language, say, N, it could be necessary that 'Snow is white' is true in N iff snow is green. The contingency of (T) then rests on the contingent fact that I speak the language that I actually speak.

David's most important criticism can be found in his argument, *foreign intruder*. Here is the quote from David:

> My idiolect, M, contains a large part of English. If any language different from M contains a sentence that is identical to 'Snow is white' but is true iff all whales are fish, then this sentence will be a foreign intruder. It will come out true, according to (D), even though it is untrue. (David 1994, p. 138)

The following quote makes more apparent David's interpretation of disquotationalism. He claims:

> [I]f any language different from M contains a sentence that is identical to 'Snow is green' but is true iff all whales are mammals, then this sentence will also be a foreign intruder. *It will come out untrue, according to (D)*, even though it is true. (ibid., emphasis mine)

Many philosophers agree with David on this point. Brendel, for instance, in Brendel 2000 sides with David and views this as the consequence of failing to respect the distinction between object- and meta-languages. Following this, a disquotationalist does not have a counterpart of the translation relation in Tarski's Convention T.

There are three important features of David's interpretation of disquotationalism. First of all, what David means by saying that something in quotation is *identical* to something

disquoted, is that they are identical *sentence types* only. Second, in terms of David's understanding of the disquotational schema, the right hand side and the left hand side are disjointed. Sentences on the right side are those in *use* to report or describe what happens in the world, and those quoted on the left hand side are just sentence types relative to a given language. Third, as a consequence of the first and the second, when the meaning or content of the left hand side sentence varies from one interpretation to another or from one language to another, the meaning or content of the right hand side sentence is fixed because the whole schema or definition is formulated in a particular language. Given these three features, the quoted sentence "Snow is green" in David's argument actually has as its meaning or content, that all whales are mammals. So it is actually true, but turns out to be untrue according to the disquotational definition of truth.

To make his interpretation of disquotationalism more thorough, David also considers Field's formulation of the disquotational theory of truth in "The Deflationary Conception of Truth" (Field 1986). In that article, Field suggests that it is not feasible to apply the disquotational truth predicate to sentences that one does not understand. After all, our understanding of the disquotational truth of a given sentence is nothing more than the understanding of that sentence itself. It makes no sense to talk disquotationally about the truthfulness of a sentence that one does not understand. It follows that the truth predicate is not applicable to *every* language, but only my own idiolect. This leads to the revision of (D):

(DM) x is a true M-sentence $=_{DF} (\sum p)_M(x = \text{'p'} \& p)$. (David 1994, p. 141)

However, according to David, this newly revised (DM), providing restrictions both on the range of the variable x and the substitution class of 'p', cannot overcome the foreign

intruder problem. The reason is that (DM) characterises a *conjunctive* predicate—namely, the predicate 'x is true and x is a M-sentence' (p. 142). This conjunctive predicate indicates that sentences can be true but not M-sentences, and thus generates the gap allowing the foreign intruders to sneak in. The only remaining alternative for construing the disquotationalist position, David suggests, is to interpret the disquotational definition as defining a *simple*, non-conjunctive truth predicate for the sentences of my idiolect—that is, it is a truth predicate that is *undefined* for sentences outside my idiolect. The newly revised definition is:

(DM) x is true$_M$ =$_{DF}$ $(\sum p)_M(x = $ 'p' & p). (ibid.)

This should do the trick, because sentences from the intersection of two idiolects, say, M and N in the case of the foreign intruder problem, turn out to be both true$_M$ and not true$_N$, or vice versa, and this is not a problem at all.

David provides several reasons to believe that this newly refined version does no better than its predecessors. Most seriously, the foreign intruder problem does not go away, because it is still possible for the truth predicate 'true$_M$' to account for some N-sentences, which are not supposed to be accounted for by it. The only way to circumvent the foreign intruder problem, David argues, is to exclude potential foreign intruders entering the overlap area, and do so without invoking any substantive resources such as a representational theory of content. Again, the main issue in the foreign intruder argument is that, without the help of any theory of content, there is no decent way to employ the truth predicate disquotationally. This of course suggests that the disquotationalist cannot account for the strong modal status of schema (T) and its instances.

The interpretation and the resultant criticism of disquotationalism made by David are incorrect and misleading. Most astonishingly, if the main function of the truth predicate is to allow semantic ascent—that is, to substitute talk of sentences for talk of reality—then *by David's definition*, when one substitutes talk of, say, "Snow is white" for talk of snow's being white, one can always substitute talk of the sentence "Snow is white" that has *arbitrary* content, for talk of snow's being white. If David is right, then it is really a mystery why the disquotationalist would have proposed something similar to this in the very first place. It is not just that on this interpretation the disquotational schema of truth does not form an explicit definition of truth, but it is a schema for sentence types that disregard what meaning or content they might have. It is plainly wrong to disjoint the two sides of the biconditionals. The schema has to be formulated relative to a given language with a fixed interpretation manual. That is,

(D$_Q$) '*p*' is true iff *p*,

> where the first '*p*' is replaced by any sentence of the base language and the second
>
> '*p*' by the same sentence with the same interpretation.

Any sentence that is incidentally identical to a given sentence type does not cause any problem. Speakers of a given language would take it as a sentence of the language they speak. If this creates any trouble, then it is an empirical matter and will eventually be discovered through practical reasoning. Further characterisation of this formulation is best seen from the context of explaining the major features or functions of this theory, and this will not be discussed until Chapter Five.

Quine, in "Truth and Disquotation" (Quine 1976), gives us the exact reason to believe that (D$_Q$) is the proper formulation of the disquotational schema. In a different kind of logic,

46

Quine explains, in which the work of quantifiers and variables are managed by constants, the complication resulting from the semantic concepts of reference and satisfaction is dispensable (p. 309). Since the point of discussing this part of Quine's reasoning is to emphasise that the mentioned and used sentences in the instances of (D$_Q$) are exactly the same sentence with the same interpretation, I will leave out the technicality of Quine's account. Employing a language with full set-theoretic strength and Quine's predicate-functor logic, which is not stronger than the standard logic of quantification and identity, we can get a good explication of the disquotational schema.[38] According to Quine, the definition of truth for a given object language is settled in a metalanguage containing this object language, the definiens (, that is, the designation sign 'Δ') and a notation for naming the expressions of the object language.[39] Paraphrasing Tarski's truth schema with a dummy symbol 'p' as follows:

(Ds) 'p' is true if and only if p,

and forming a corresponding designation schema also with a dummy symbol as follows:

(Des) $\Delta('p') = p$,

it is clear that $\Delta('p')$ is equal to the claim that 'p' is true according to (Ds) and (Des), where the dummy symbol 'p' represents a sentence. The intermediate concept inductively defined here is not satisfaction, but designation. This inductive definition is destined to make every equation of (Des) with any one name replacing the dummy symbol, come out true. It turns out that truth is a special case of the designation schema, where again the dummy symbol represents a sentence (pp. 309-310).[40]

---

[38] According to Quine, the language employed is in connection with Schönfinkel's combinatory logic.
[39] I argued previously that the distinction between an object language and a metalanguage is not essential to my formulation of disquotationalism. I maintain this distinction here for the sake of convenience only.
[40] In Quine 1995, the disquotational character of the truth predicate can be seen as a special case for the denotation of no-place predicate. "In the zero case, there being nothing for the no-place predicate to be true *of*, denotation reduces simply to truth outright." (p. 65)

This is quite similar to what Quine claimed in Quine 1953*a*. In that paper, he claimed that the theory of reference should be characterised as follows: '*p*' is true-in-*L* of every *p* and nothing else; '*p*' names-in-*L p* and nothing else (in which '*L*' is not a free variable ranging over all languages). The difference between Quine's notion of disquotation and Tarski's is that, in Quine's account, after the last metalinguistic sign has been disposed of, just the formula of the object language quoted in the disquotational truth schema (or the designation one) remains. In Tarski's account, what remains is equivalent to the quoted formula.[41] This suggests that the stronger disquotational property under Quine's account is not preserved under the recursive definition of satisfaction. We should also consider whether or not any version of designation deserves that name, insofar as it fulfils (Des). For instance, unlike the designation account at hand in which Δ('S') is unwound into S, another account unwinds Δ('S') into S_1 because S and S_1 happen to be the same object, the same function. The virtue of the disquotational account, Quine argues, is that it exactly unwinds Δ('S') into S and Δ('S_1') into S_1 respectively.[42] Since Quine is only interested in accounting for the truth of interpreted languages, the quoted sentences 'S' and 'S_1' and those disquoted are exactly the same sentences with the same interpretation. Again, what are unwound are not sentence tokens but sentence types. Otherwise, the whole point of disquotation is missed.

---

[41] The word "disquotation" here is only meant to be the procedure of removing the quotation marks in the T-schema and the disquotational schema. However, it is arguable whether or not Tarski's theory is with the feature of disquotationalism. For instance, Devitt, Romanos and many others think it is.

[42] The usual way to define logical particles depends on truth and falsity. For instance, negation in the logical sense can be defined in terms of two conditionals: <*p*> is true → <not *p*> is false; <*p*> is false → <not *p*> is truth (Following Horwich, we distinguish 'not' from 'not', which is in its ordinary sense.) It seems that, unless there is another method to avoid this, logical particles might cause some trouble of circularity. Nevertheless, Paul Horwich considers a way out in *Truth*. For instance, negation can be defined not in terms of a truth table, but in terms of a set of theses: $p \rightarrow$ not not *p*; not *p* → not *p*; *p* or not p.

(D), (DM) and (D$_M$) are of course not in harmony with my formulation of the

disquotational theory. Besides mine, there are two objections made by Field and Halbach.

Field in Field 1994 warns us that, if the sort of interpretations like David's has general

applicability, then the truth predicate no longer serves the purpose of abbreviating infinite

conjunctions or disjunctions, or expressing blind ascription. Halbach also makes a correct

observation in Halbach 1999 and Halbach 2000 on this sort of interpretation, which he

calls Levy's argument. That is, there are no necessary truths in terms of such an

interpretation. "Any sentence," he continues to argue in Halbach 2001, "even a tautology,

could have meant something else and thus be false; therefore no sentence is necessarily

true."

There is another problem for David's (D), (DM) and (D$_M$). That is, do we have

substitution quantifications in ordinary languages? Field, for instance, suggests in Field

1986 as follows:

> [I]f 'true' is simply a device of infinite conjunction, then we have a serious need for a
> predicate of truth only because (or, only *if*) we don't have a substitution quantifier in
> English.[43] (p. 58)

Grover, Camp, and Belnap might disagree with this assumption, claiming that the

substitution quantifier in English is actually expressed by the truth predicate (serving as a

'prosentence' according to their terminology). David (1994) claims that we can have a

weaker or a stronger position toward this question. The weaker position is that, although

the formulation of the disquotational definition of truth is not entirely based on ordinary

language, the definition plus the metalinguistic explication of substitution quantifiers

---

[43] Field mentions this in his footnote number 4. However, there are some different opinions regarding the
assumption that the sole purpose of the truth predicate is a device of infinite conjunctions or
disjunctions—for instance, Gupta in "Minimalism".

makes it clear enough to understand the disquotational definition of truth. Besides, if there is a device in an ordinary language that can abbreviate the infinite disjunctions or conjunctions, the truth predicate is dispensable anyway. The reason why the truth predicate still has a role is the very reason why we cannot give an account of truth purely in terms of ordinary terms. The stronger position, on the other hand, is that the metalinguistic explication of substitution quantifications can be managed to become part of our ordinary language. That is, the truth predicate is dispensable as far as we can *in principle* learn substitution quantifications to abbreviate the infinite list (p. 109). The weaker position, it seems, is more comprehensible than the stronger one since the intuitive understanding of the truth predicate is clearer than the understanding of substitution quantifications. This is because there seems to be no intuitive way that we can formulate substitution quantifications without the dummy symbol 'p'. To assume that the metalinguistic explication of substitution quantifications is learnable *in principle* and thus that the truth predicate can be replaced by substitution quantifications in an ordinary language seem less promising.[44]

Moreover, Quine does not accept substitution quantification. Since there is no restriction on the grammatical category of the admissible substituends, the substituted expressions are not just for sentences but for any symbol or sequence of symbols in a sentence as well. In *Philosophy of Logic*, Quine makes the distinction between objectual quantification and substitution quantification. In standard logic, names are redundant—that is, 'F*a'* can

---

[44] The issue on which David concentrates is this: How can the disquotational definition of truth be comprehensible in ordinary terms? What concerns me here is the role played by substitution quantifications in the explication of the disquotational theory of truth. On this particular concern, it seems that David's viewpoint also helps make some clarification.

always be paraphrased into '(∃x)(Ax & Fx)'.[45] So the singular terms contained in objectual quantification are variables only. It is another story for substitution quantification. Names are essential to this non-standard quantification. Consequently, it is deviant when the domain of discourse is rich. For, there are nameless things in the discourse even if we grant it an infinitude of names (p. 92). Furthermore, according to Quine, substitution quantification for sentences is defined in terms of the interchangeability of true substituted expressions, as well as the interchangeability of false expressions, in all contexts *salva veritate*. Therefore, it is circular to employ substitution quantification in formulating an explicit disquotational definition of truth. For these two reasons, it is not possible for Quine to accept David's various definitions.

According to Gupta (1993) David's definitions are also not right because "The sentence with the form '*φ* or not *φ*' is true" and the infinite conjunction it expresses are not sense equivalence. We do not need to know the meaning of every disjunct to know that the sentence with the form '*φ* or not *φ*' is true. Nevertheless this seems to be a problem for the disquotational theories of truth in general. It raises two questions: 1) What does it mean to say that the truth predicate *expresses* the infinite conjunctions and disjunctions in certain cases? 2) How much deductive power is desirable according to the disquotational theory?

3. Conservativeness and Substantiveness

If the disquotational theory consists only of the disquotational schema, (D_Q), as axiom schema and all of its instances as axioms, it will be too weak to account for all the facts with regard to the concept of truth of the target language. For instance, it implies all the

---

[45] Or better, we can use '(∃x) (Ax)' and '(x)(y)(Ax & Ay & (x = y))' together to indicate that predicate A is only true of one thing (Quine 1970, p.25).

T-sentences of the instances of "The sentence with the form 'p or not p' is true" but not this very sentence itself. Moreover, all the instances of the disquotational schema are held, but not the schema itself. So, it seems, the theory of truth needs to include something else.

The reason I choose to focus on the issue of the conservativeness of the disquotational theories, is to show that loosening my previous formulation of disquotationalism to include additional axioms for the purpose of enforcing deductive power, is a risky matter. But first let us turn to question one given at the end of the last section. Namely, what does it mean for a disquotationalist to say that the truth predicate helps *express* infinite conjunctions or disjunctions in the cases of blind ascriptions or sentences with the same feature, such as "Every sentence with the form 'p or not p' is true"?

Expressibility

First of all, not *every* infinite conjunction or disjunction is expressible with the help of the truth predicate. For instance, the infinite conjunction of sentences that have less than ten words is not expressible in terms of the truth predicate. Although it is most likely false, we can hardly find any reason to see how the truth predicate helps express this generalisation. The reason is clear. The instances of it are not derivable from it. So we may say that the truth predicate only helps express an oblique generalisation over sentences with the same *syntactic* feature. Therefore, the truth predicate does not express all infinite conjunctions or disjunctions.

However, an infinite conjunction of sentences with the same syntactic feature is not necessarily expressible in terms of the truth predicate. All sentences, for instance, with the

feature 'if *p* or not *p*, then p' together form an infinite conjunction not expressible in terms

of the truth predicate, for the truth-value of every instance of it depends on the truth-value

of 'p'. It follows that some contradictory instances such as "If snow is white or snow is

not white, then snow is white" and "If snow is not white or snow is white, then snow is

not white" are derivable. Therefore, the expressibility of infinite conjunctions or

disjunctions in terms of the truth predicate is not subject to the condition of their having

the same syntactic feature only, but is also subject to the condition of their instances

having consistent truth-value.[46] But this second condition is not obtainable because if

there is not sufficient deductive power, then an axiomatic disquotational theory in which

only instances of the disquotational schema are axioms does not even fix the extension of

the truth predicate. Halbach (1999) correctly observes that, if a model-theoretic definition

of truth is presupposed, the formula "(x) $(\varphi(x) \rightarrow Tx)$" indeed expresses the infinite

conjunction over a set of sentences specified by the function 'φ'. That is, "(x) $(\varphi(x) \rightarrow Tx)$"

holds in a model if and only if the infinite conjunction '$\varphi(x)$' holds in the same model

(Halbach 1999, p. 7). Without doubt, this solution is not available for the

disquotationalist.

Another option suggested by Halbach is *provability*. The second condition to be satisfied

can be thus characterised as the following: an infinite conjunction or disjunction is

expressible in terms of the truth predicate if and only if the valid or provable translation

of it in the truth-theoretic language is provable (Halbach, 1999, p. 9). To be precise,

suppose the target language is a first-order language *L₀* plus a recursively enumerable

theory *T₀* that is formulated in classical logic and is with some deductive power not

---

[46] In Chapter Five, what sorts of oblique generalisations are expressible in terms of the truth predicate is
important to my argument by which Gupta's sense equivalence criticism is rejected.

exceeding first-order logic. And further suppose a language L_1 as L_0 together with the disquotational truth predicate, and a disquotational theory of truth T_1 as T_0 along with the additional T-sentences of L_0 as axioms. Hence, an infinite conjunction can be expressed in terms of the truth predicate in T_1 if and only if it can be proved in T_0. If an infinite conjunction "$(x) (\varphi(x))$" can be proved in T_0, then $(x) (\varphi(x) \rightarrow Tx)$ can be proved in T_1.[47] But does saying that an infinite conjunction is provable by all its conjuncts and some inferential rules in T_0 not amount to saying that the corresponding generalising sentence is derivable from all its instances and some inferential rules in T_1? This is untenable because the disquotational theory T_1 consisting of T-sentences as axioms and T_0 does not have the deductive power exceeding first-order logic. Tarski (1956b) is probably the first one who notices this. Without the help of some infinitary rule such as ω-rule, the provability of infinite conjunctions in T_0 or the derivability of the generalising sentences in T_1 is out of our reach. Moreover, it is not sure at this stage that T_1 has the capacity to prove that any sentence is not true if its negation is true, or that all propositional tautologies are true. And *any* infinite conjunction or disjunction can only be expressed by attaching the truth predicate to the corresponding generalising sentence. In other words, "all sentences with the syntactic feature 'p and not p' are true," for instance, is also a legitimate generalising sentence for the infinite conjunction "$(x) (\varphi(x))$". These consequences immediately follow when it is shown that T_1 does not fix the extension of the truth predicate.

But this seems to be rather unbearable. With what sort of additional apparatus can disquotational theories similar to T_1 make this disadvantage disappear? This leads us to one of the main topics of this section, namely, conservativeness.

_____

[47] Halbach (1999) offers a detailed argument showing that $(\varphi(x) \rightarrow Tx)$ is not provable in T_1.

Conservativeness

Shapiro (1998) and several others make a forcible attack on disquotationalism, claiming

that it will be inevitable for disquotationalism to imply some substantive consequences. It

will be an easy job to prove this if disquotationalists concede to the constraint that all

disquotational theories have to be conservative over logic. For a trivial argument offered

by Halbach (2001) simply shows that in pure identity logic, where "$(x)\,(x = x)$" is

provable and "$(x)\,(x \neq x)$" is refutable, "$(x)\,(x = x) \neq (x)\,(x \neq x)$" and "$\exists x \exists y(x \neq y)$" are

consequences of T_1.[48] Claiming that the above argument employs the identity axioms to

which the truth predicate is already applied is not a shelter for disquotationalists because

this makes every newly introduced symbol conservative trivially. No rules or axioms of

logic can in principle be extended to the expanded language with any new symbol (p.

179). So, standard disquotational theories such as T_1 imply that there are at least two

objects. But this is not surprising, for standard disquotational theories already commit to

the abstract objects—namely, sentence types—according to my analysis in the last section.

At any rate, conservativeness over logic is a formidable Lorelei for disquotationalists. In

the following, I turn to another kind of conservativeness, conservativeness over the base

theory formulated by the target language.

If the base theory is Peano arithmetic, the conservativeness of standard disquotational

theories is quite straightforward according to Halbach (2001). The base theory together

with T-sentences does not license the derivation of any arithmetical sentence that is not

provable in Peano arithmetic. Further, if rules of inference and axioms of first-order

---

[48] T_1 is of course presupposed to have the deductive power of pure identity logic here.

propositional logic are allowed to apply to sentences with the truth predicate, the standard

disquotational theory T_1 is capable of fixing the extension of the truth predicate. It can

also be proved that a sentence is not true if its negation is true, and that a sentence with

the syntactic form of any propositional tautology is true. Here is one demonstration.

For a given T-sentence,

(1) $T\underline{x} \leftrightarrow x$, where $\underline{x}$ is the name of x,

giving that the rule of *modus tollens* in propositional logic be applicable to sentences with

the truth predicate, we can get

(2) $\neg T\underline{x} \leftrightarrow \neg x$

from '$\neg x$'.

Replace '$\underline{x}$' and 'x' in (1) by '$\underline{\neg x}$' and '$\neg x$', we get

(3) $T\underline{\neg x} \leftrightarrow \neg x$.

From (2) and (3), (4) follows:

(4) $\neg T\underline{x} \leftrightarrow T\underline{\neg x}$.

So any given sentence is not true as long as its negation is true.

Giving (5): $T\underline{x} \leftrightarrow T\underline{x}$,

(6) $T\underline{x} \lor \neg T\underline{x}$

follows. And (7) can be derived from (1), (2), and (6):

(7) $x \lor \neg x$.

Giving the T-sentence of (7), we get

(8) $T\underline{x \lor \neg x} \leftrightarrow x \lor \neg x$.

From (7) and (8):

(9) $T\underline{x \lor \neg x}$

follows. So any sentence with the syntactic form of any propositional tautology can be

proved to be true in a similar fashion. Any sentence, it was proved, cannot be both true

and not true.[49] However, what can be proved are sentences with some tautological

syntactic feature but not the corresponding generalisation because of the lack of an

infinitary rule of inference. So the applicability of all inferential rules of first order logic

to disquotation sentences is the default position of the disquotational theory regarding

(D_Q).[50] Furthermore, the disquotational theory of truth is not just an axiomatisation of

truth but a theory *about* the axiomatisation of truth within which all uses of the truth

predicate can be explained by its axioms and other truth-free explanatory resources.[51]

As stated previously, not all generalisations should be proved by the theory of truth. For

some, "all the consequences of the general theory of relativity are true" is among them.

Most importantly, Halbach (2001) correctly makes the claim that some mathematical

generalisations should be left not proved by the theory of truth. For instance, if all

mathematical generalisations are provable in terms of the disquotational theory and the

base theory of Peano arithmetic, substantive consequences such as the consistency of ZF

plus the continuum hypothesis follow.

For most disquotationalist accounts by which the disquotational theory of truth is

characterised as consisting of T-sentences of the target language as axioms and the

disquotational schema as an axiom schema only, the generalisations of different sets of

sentences with the same syntactic feature is to play a role of reasoning similar to the role

---

[49] Fixing the extension of the disquotational truth predicate seems to be extremely important. Otherwise the disquotational theory fulfils the material adequacy condition trivially by implying all T-sentences of the target language sentences but not the NT-sentences "$\neg T\underline{x} \leftrightarrow \neg x$".

[50] This is also Horwich's default position.

[51] In Horwich 1998*b*, Horwich emphasises the methodological point that not all the facts regarding truth are about truth only. Therefore, it is perfectly proper to make use of theories about those matters to explain these sorts of facts (p. 7). I am with Horwich on this point.

played by infinite conjunctions or disjunctions. Whether or not these conjunctions or disjunctions are derivable from the truth theory and other suitable explanatory resources is not that important.[52] Equally, whether or not "All sentences with the form 'p or not p' are true," for instance, is derivable from all its instances is less important. What matters is what sort of sentences they imply; in the case of "All sentences with the form 'p or not p' are true," what is important is that it implies all the sentences "ψ or not ψ," where 'ψ' stands for a sentence of the target language. If we are inclined to strengthen our disquotational theory, it seems quite natural to add some further inductive axioms similar to those resulting from the axiomatisation of Tarski's recursive definition of satisfaction. For instance, Halbach claims the following to be one of the inductive axioms: For all closed terms t: Pt is true if and only if Pval(t) (for all one-predicate symbols P and similarly for predicate symbols with more places; val(t) of a term is defined in terms of a reference scheme).[53] The exact content of these new axioms depends on the base theory formulated in terms of the target language.

Halbach (2001) and Field (1999) are two advocates of this expansion of the disquotational theory of truth. They both try to defend disquotationalism against the attack issued by Shapiro (1998) that the disquotational theories of truth embracing the above-mentioned inductive axioms are not conservative over the base theory. Halbach (2001) concludes,

> If the deflationist understands his claim that truth is not a substantial notion as implying that his truth theory has no substantial consequences, he commits a mistake. (p. 188)
> [The deflationist] claimed only that truth serves the purpose of generalisation, and he does well to parse this purpose in a strong way such that his truth theory also proves

---

[52] Please cf. the second section of Chapter Five for a more elaborate discussion.
[53] The last clause in the brackets is a paraphrase. Halbach chooses Peano arithmetic as the base theory so that val(t) of a term is defined within it.

certain generalisations. That it incidentally also yields certain 'substantial' consequences is merely a secondary effect. (p. 189)

But be careful, certain 'substantial' consequences might be very unwelcome. In the following I would like to show two reasons for not embracing the additional inductive axioms of reference and satisfaction. One reason relates to the need for an accurate formulation of Quine's disquotationalism. The other relates to disquotationalism in general.

It was repeatedly emphasised in the first two sections that Quine could hardly agree with the explanatory order of Tarski's semantic theory of truth. More precisely, any given reference scheme is acceptable only if it can preserve the distribution of *truth*. Hence, the notion of truth is explanatorily more basic to the notion of reference. Additionally, Horwich dissents from the idea of bringing inductive axioms of reference and satisfaction into a deflationary theory of truth, for the following reasons.[54] First, he thinks that only the axiom schema, but not inductive axioms, of reference and satisfaction can make sense—that is, they have to be infinite too; second, many kinds of propositions or sentences cannot have their truth explained on the basis of the referents of their parts—for instance, statements of probability or counterfactual conditionals; and most importantly to his minimalism, propositions constructed from the primitive concepts that are not expressed in our language are not explainable (Horwich 1998*b*, pp. 136-137).

The second reason is that my objective regarding the issue of conservativeness and substantiveness is different from Field's and Halbach's. Field tries to find an acceptable

---

[54] These reasons actually come from Horwich's disapproval of truth conditional semantics, but it also seems suitable to see them as objections to bringing inductive axioms of reference and satisfaction into the disquotational account of truth.

way to embrace the inductive axioms without being accused of inflating the disquotational theory of truth. Halbach, on the other hand, handsomely admits that the revised version does have some substantial consequences. But this does not amount to an inflation of disquotationalism for him. Field and Halbach use Peano arithmetic as the base theory in both of their arguments, for the disquotational theory of truth is destined to fail if it cannot survive in giving an explanation of the truths of mathematical theories. This is fine, but my interests are different. I would like to show that embracing the inductive axioms might cause some problems in accounting for the truths of natural languages.[55]

Inevitably in Halbach and Field's reasoning, a reference scheme is involved in explaining the truths of any given natural language. Even if the inductive axioms do not cause any inflation, introducing a reference scheme into the account might. Take English as the target language, and consider any given belief system formulated in English that is comprehensive enough to include as many as empirical truths as the base theory. Then the reference scheme introduced is not supposed to have any empirical, semantical, and ontological features to lend any substantive flavour to the disquotational theory of truth. For instance, this base theory should include some empty names such as 'the King of France Louis XVIII' or 'phlogiston'. To spare disquotationalism unnecessary inflation, substantive theories of reference such as a causal theory have to give way to some deflationary reference theories. Among them, Tarski's is no worse than any others. There are three ways to explain empty names. First, there are no empty names in the reference scheme. That is, the reference scheme does not enlighten us with regard to empty names.

---

[55] Halbach admits in Halbach 2001 that he is taking the advantage of employing Peano arithmetic as the base theory by not being in need of providing any consideration with regard to a selection of the theory of reference.

They then disappear from the list of T-sentences. This is untenable because empty names appear essentially in many sentences. Second, empty names behave just like others, *as if* they refer to something—that is, 'the King of France Louis XVIII', for instance, *vacuously* refers to the King of France Louis XVIII.[56] This is also not practical because any name behaving in this way cannot be used in the inductive axioms. Again, the corresponding T-sentences are not implied by this reference scheme and the inductive axioms. Finally, empty names refer to nothing.[57] Hence "The King of France Louis XVIII is bald" is pre-determined to be satisfied by no sequences of objects. But this does not give us the needed piece of information to allow us to determine whether or not "The King of France Louis XVIII is bald or the King of France Louis XVIII is not bald" is an instance of the infinite conjunction expressed by "All sentences with the form '*p* or not *p*' are true." If no sequences of objects satisfy "The King of France Louis XVIII is bald" purely because 'the King of France Louis XVIII' is an empty name, then "The King of France Louis XVIII is not bald" is also not satisfied by any sequence of objects. 'The King of France Louis XVIII' is just a dramatic case. Many theoretical terms face a similar problem. And these problems of reference should not be the business of any disquotational theory of truth in the very first place. So committing to the inductive axioms causes serious problems to which the standard disquotational theories are not committed. This will be further clarified in the following chapter.

Final Remarks

As a result, Quine's disquotationalism is better off not to embrace the inductive axioms of

---

[56] Cf. the following chapter: "Inscrutability of Reference, Ontological Relativity and a Disquotational Scheme of Reference".

[57] If singular terms are paraphrased as one-place predicates, it is important not to consider them to be true of the class of all things that it is true of. Otherwise empty names fallaciously turn out to be coextensive.

reference and satisfaction. However, Quine's disquotational theory is not just about a theory of truth but also about a theory of reference in general. 'Semantic ascent' also applies to singular terms, general terms, and predicates. Hence the explication of Quine's disquotationalism is not complete until Quine's theory of reference is subjected to thorough investigation.

Chapter III

Inscrutability of Reference, Ontological Relativity and a

Disquotational Schema of Reference

The central issue of this chapter is the role played by a disquotational schema of reference in Quine's ontological theory. In terms of the analysis of this issue, it is also important to see whether or not we should incorporate inductive axioms of reference and satisfaction into the disquotational theory of truth. Although I must inevitably discuss Quine's ontological theory, it is not my intention to evaluate its significance and plausibility.[1] On the contrary, by concentrating on one important argument against the thesis of ontological relativity, my analysis will illustrate the essential nature of a disquotational schema of reference within Quine's ontological theory. Based on this, the whole picture of Quine's disquotational theory of truth should be further clarified.

To begin with, two related theses of Quine will be briefly introduced, namely, the thesis of inscrutability of reference and the thesis of ontological relativity. The thesis of

---

[1] For instance, Fogelin in "Quine's Limited Naturalism" criticises Quine's indeterminacy semantics as incompatible with his naturalism. Although it seems implausible, I am not tempted to include the related discussion in this chapter. Moreover, Fodor in *The Elm and the Expert* offers an argument against the thesis of inscrutability of reference and Georgalis in "Reference Remains Inscrutable" tries to tackle it. Both of them seem to show some insufficiency in a proper understanding of the target thesis. This can be exemplified in the following. Fodor claims: "[G]iven the data Inf [ormant] provides, the linguists know that Inf hold 'there's a rabbit' true in a situation iff he holds 'there's an urp [, that is, undetached proper rabbit part,]' true in that same situation'" (Fodor 1994, p. 63). But how can the informant hold true both 'there's a rabbit' and 'there's an urp'? Georgalis thinks that Fodor's argument begs the question. For him, Fodor simply presupposes that we have to interpret certain general terms in the standard interpretations of them while the non-standard interpretation of them assumed for the *reductio* is gone. And this causes a disagreement in the truth-values of certain sentences between linguists. However, this is exactly what the *reductio* is about. According to the target thesis, linguists will provide incompatible reference schemes even if the totality of empirical truth is preserved. Fodor intends to show, in terms of his argument, that incompatible reference schemes will inevitable cause disagreements in the truth-values of certain sentences between linguists, and Georgalis seems to miss the point. Again, this is not my primary task of this chapter.

inscrutability of reference is as follows: if the reference of the terms of an informant's

language is not uniquely decidable by a field linguist purely on empirical grounds, then

what these schemes say about the reference of the terms of the translated language is

therefore not an empirical feature of those terms. That is, the linguist can provide more

than one translation manual within which the reference schemes of the terms can be

radically different from each other, but still be compatible with all (possible) pieces of

empirical evidence. It follows that the choice of one particular reference scheme is

arbitrary or, in other words, these terms do not uniquely determinate a reference scheme.

What of the thesis of ontological relativity? One simplified version of ontological

relativity can be put in this way: "What makes sense is to say not what the objects of a

theory are, absolutely speaking, but how one theory of objects is interpretable or

reinterpretable in another" (Quine 1968, p. 50). It was held in Chapter Two that the main

reason not to incorporate the inductive axioms with regard to reference and satisfaction

into Quine's disquotational theory of truth is the explanatory priority of the concept of

truth. This claim, however, is plunged into turbulence when both of the aforementioned

theses are under fire.

Hartry Field considers the thesis of ontological relativity to be incompatible with the

correspondence theory of truth. If a correspondence theory of truth is about the

explanation of the notion of truth in terms of the relation between words on the one hand

and the extralinguistic objects that they are about on the other, then the only noteworthy

correspondence, following the thesis of ontological relativity, is the one between the

words of one theory and the words of another (Field 1974, p. 199). Field's ambition is to

show that, granted the thesis of inscrutability of reference, the correspondence theory of

truth still prevails (perhaps in a more general sense, however). Donald Davidson accepts the thesis of inscrutability of reference and a relativised concept of ontology, but he agrees with Field's argument against Quine's thesis of ontological relativity, criticising it as falling short of providing an intelligible way to justify itself. Roger Gibson, Jr. on the other hand, censures both Field and Davidson's misinterpretation of Quine's theories. One important feature of his interpretation of Quine's ontological relativity thesis is that speakers of every language are normally uncritical of their own languages and thus make available a way to end the ultimate relativisation of ontology by maintaining a word-world correspondence relation between words and extralinguistic objects.

This chapter has five sections, in which the following five tasks will be undertaken. First, I will elucidate Quine's reasoning with respect to relativising ontology. Second, I will show that Field's argument in favour of a correspondence theory of truth is invalid. Third, I will show that Gibson's interpretation of Quine's ontological relativity thesis is incorrect. Fourth, I will show that a disquotational schema of reference of our mother tongue is essential to enable us to continue common sensically doing science and living daily. Finally, I will show that a disquotational schema of reference does not and cannot provide any support to the reasoning that, in order to redeem the lack of deductive power of the disquotational theory of truth, one should explicate the disquotational conception of truth in terms of the inductive axioms of reference and satisfaction. And more notably all instances of the disquotational schema of reference do not constitute a genuine reference scheme.[2]

_____

[2] The phrase 'the disquotational schema of reference' is used throughout this chapter in its general sense. That is, it includes satisfaction as well. Any point made in relation to the disquotational schema of reference is equally applicable to the disquotational schema of satisfaction. This is the default position and will not be emphasised hereafter.

1. Quine's Manoeuvre

In the quoted passage about ontological relativity, the point seems to be that, although it is improper to ask what terms of a theory (or language) refer to in an absolute sense, it does make sense to ask this question in a relativised way. More plainly, the only sensible way to talk about reference is to relativise it to a background language or a background theory. The simplest and probably least questionable manner of explicating ontological relativity might be in accordance with Quine's account of proxy function—that is, for any given theory we can provide a function which maps its universe on to the universe of a background theory. This mapping can be a one-to-one or many-to-one relation, depending on our purpose in utilising it. And the universe of a background theory is, if not identical to, then normally broader than the object one. In the case of a many-to-one relation, we use this relation as a means of ontological reduction. According to Quine, it is not necessarily worrying that we must grant the ontological profligacy of the excessive universe of the background theory in order to operate the reduction. This is because the worry is no more pressing than in the case of *reductio ad absurdum*, where we assume a falsehood so as to disprove it (Quine 1968, p. 58). So the excessive universe is eventually dispensable. In the case of a one-to-one relation, a given theory is always reducible to a background theory, according to Quine, and it can also form an argument in favour of the thesis of inscrutability of reference. Put simply, for any given theory in which a name refers to an object x according to its reference scheme, one can produce a proxy function, say, ø to map x on to $ø(x)$ in the reference scheme of the background theory; similarly, if a predicate refers to (is true of) each thing x such that $F(x)$ in the object theory, it refers to each thing x in the background theory such that $Fø(x)$; along with an appropriate

characterisation of a relation like satisfaction in, presumably, each case, the truth

conditions assigned by the background theory are equivalent to those assigned by the first

theory, thus the inscrutability of reference. However, it seems to be one step further to

hold the thesis of ontological relativity in terms of the thesis of inscrutability of reference.

If there are neither semantical nor physical facts to determine the choice of reference

scheme uniquely, then the only significant way to talk about the reference of any given

theory (or language) is to relativise it to another theory (or language) no matter how

arbitrary the choice is. The consequence of doing so is the risk of infinite regress from

one theory or language to another. For Quine, we end the infinite relativisation of

ontology in practice by acquiescing in our own language and taking its words at face

value (ibid., p. 49). Let us call this Quine's manoeuvre. With respect to this, both Field

and Davidson hold an opposing position.

The proxy function argument is the one that really indicates the mystery of reference,

since proxy function can produce more dramatic cases than those in radical translation.

By this I mean that normally, in translating an informant's language we would not use a

reference scheme that maps itself on to, for instance, the cosmic complement of its

objects. In translating a remote language it is more likely to be a limited case in which the

mission of the field linguist is to find a way to understand or interpret her informant's

language in a more *ordinary* vein. However, it is different in the case of ontological

relativity since it is in a more *theoretical* vein. That is, in the ordinary vein the field

linguist works hard to reduce the amount of competing manuals, while in the theoretical

vein it is supposed to show that the interpretation of the ontology of any given language

or theory could be varied drastically but that the sentence-to-sentence structure is still

preserved. Probably because of this difference, some philosophers such as Field are

prompted to assume that providing a solution to the inscrutability of reference in the case

of radical translation is sufficient to answer the mystery of reference issued from the

thesis of ontological relativity. In a similar sense, perhaps, Gibson does not make a clear

distinction between these two cases, either. Field's reasoning is only justified when the

thesis of ontological relativity can be successfully renounced. It will be shown, however,

in the following discussion, that a disquotational interpretation of Quine's manoeuvre can

save him from the criticism, made by Field, that he commits the fallacy of the myth of the

museum, whilst also maintaining the thesis of ontological relativity. But let us first shed

some light on Field's argument.

2. Field's Argument against Ontological Relativity

If the idea of signification[3] is to be explained relative to a translation manual, Field

argues, then this relativised way of talking about the reference scheme of a given theory

or language is useless. There is no difference between saying that a term T used in one

language signifies the set of rabbits, relative to a translation manual M, and saying that M

translates T as 'rabbit', if the only sensible way to talk about the reference scheme of a

theory or language is to interpret or reinterpret it in another theory or language. More

precisely,

(1) for every predicate T, set $\{x \mid Fx\}$, and manual M, T signifies $\{x \mid Fx\}$ relative to M

    if and only if M maps T onto 'F' (Field 1974, p. 203).

This is not satisfactory because we cannot quantify *into* quotation marks in an objectual

fashion. According to Field, we can make some adjustment to avoid this problem by

---

[3] Quine uses 'denotation'.

paraphrasing (1) into the following:

(2) for every predicate T, set $\{x \mid Fx\}$, and manual M, T signifies $\{x \mid Fx\}$ relative to M if

and only if M maps T onto some term which signifies $\{x \mid Fx\}$ (ibid.).

Field's objection to (2) is that the relativised explication of any given reference scheme

has to rely on unrelativised semantic notions applied to our own language. The

assumption that the ontology of our own language, unlike others, is fixed, is what Field

cannot accept. He thinks that this makes Quine a victim of the myth of the museum (p.

204). That is, if the ontology of our own language is fixed and the relativisation ends in

our mother tongue, all the reference schemes of different theories or languages seem to be

fixed eventually by relativising them to our language. It is, however, hard to see the

difference between the myth of the museum and this account in which the exhibits are the

referents, and referential words of different languages are just like expository labels for

different language speaking visitors. Davidson agrees with Field on this point and renders

a similar argument (Davidson 1979, pp. 230-1), which will not be discussed.

Moreover, Field cannot accept the arbitrariness of the choice between the competing

translation manuals due to the fact that there are no semantical or physical facts to show

the adequacy of choosing between any of them. Instead, he proposes a reformed account

of certain semantical notions. Although there is no empirical ground for choosing one

translation manual over another to relativise the reference scheme of a given theory or

language to another, it can be held that a general term x of a given language *partially*

signifies the sets of objects of all competing manuals. A simplified example can be put as

follows: the term *gavagai* partially signifies the set of rabbits (relative to a translation

manual M_1) and the set of undetached rabbit parts (relative to M_2) but fully signifies either.

The problem for Field is that the indeterminacy of reference scheme occurs hand in hand with the translation of the identity predicate. Generally this is not a problem for Quinean semantics because the translation of 'gavagai', say, is fixed as far as the translation of the identity predicate is. In other words, 'gavagai' can be translated as signifying the set of rabbits, if the identity predicate is translated in certain *relevant* ways. However, this is a difficulty for Field's account, because of the random combinations of the translations of 'gavagai' and the identity predicate. To solve this problem, Field claims that it is inevitable to that some sort of correspondence relations be introduced to define *relevance* in terms of them. It follows that these causal relations between terms and extralinguistic objects can provide a way for Quine to avoid committing the myth of the museum fallacy.

Davidson on the other hand correctly criticises Field's reformed account on the grounds that although in a standard theory of truths, semantic roles are assigned to the parts of a sentence to explain what makes it true, Field confuses the issue by overemphasising the importance of the independent account of the semantic properties of reference, naming, signification, and so on (Davidson 1979, p. 236). The confusion is the failure to notice the difference between explaining truth, given the theory, and providing evidence that the theory is true of a speaker or community (ibid.). In addition to the assumption that there are causal connections between terms of a language and extralinguistic objects, it can be imagined that these connections do not play any role in the explanation of the use of these terms in sentences. Consequently, these connections are irrelevant to an account of that particular language. Establishing a reference scheme of a given theory or language relies on evidence concerning the adequacy of a theory of truth at the sentential level.

Further, it seems that Field's account does not provide a reasonable solution to the alleged failure of the thesis of ontological relativity. This account is disputable in the case of radical translation because the main purpose of it is, after all, to reduce the number of the competing translation manuals.[4] It is different in the case of the thesis of inscrutability of reference. According to this thesis, the reference scheme of any given theory or language is *in principle* indeterminate. The proxy function argument shows that the term 'rabbit', for instance, can be reinterpreted by permuting the universe to map it on to, say, the cosmic complement of rabbits. We can even permute the universe of our language or theory and map it on to itself. It follows that there could be many, if not infinitely many, reference schemes of English. 'Rabbit' would perhaps signify many things that you can imagine as far as the truth conditions assigned by the new reference scheme are equivalent to the old one. It can be claimed that, in some, if not in most cases, some general terms of a given theory or language might be boiled down to partially signifying exactly the same sets of objects. For instance, the general terms 'rabbit' and 'undetached rabbit parts' and 'rabbit stages' all partially signify the set of rabbits, the set of undetached rabbit parts, and the set of rabbit stages. This is not incompatible with Field's account. Unless Field can prove that this is in principle impossible, I can hardly see how his account can work and how to establish any *causal, correspondence* relations between words and objects.

3. On Gibson's Criticism of Field's and Davidson's Arguments

Whether Field's and Davidson's arguments are conclusive or not, there are at least two

---

[4] In the 'Postscript' of Field 1974, Field claims:

> Even though Quine was not sympathetic to a causal theory of reference, he did often grant that the stimulus meanings of terms could be used to rule out alternative reference schemes. (p. 219)

Still, I think, this is in the case of radical translation rather than ontological relativity.

ways to account for Quine's manoeuvre. One can make a dichotomy between one's

mother tongue and translated languages, and thus hold that the relativisation of ontology

only happens in translating (/interpreting) another language. Or, one can find a way to

account for Quine's manoeuvre without violating the thesis of ontological relativity. I will

discuss these respectively in this and the following sections.

Gibson has a strong counterview to Field's accusation that Quine commits the myth of the

museum fallacy. Since Davidson only opposes the part of Field's argumentation in favour

of the correspondence theory of truth, Gibson's deprecation of Field's position applies

equally to Davidson's view. In the following discussion I will first make clear Gibson's

argument, and then indicate my disagreement with it afterwards.

Gibson thinks that Field and Davidson failed to notice in Quine the significant distinction

in the thesis of inscrutability of reference and the thesis of ontological relativity in the

epistemological vein—in which there is no fact of the matter on the denotation of the

terms of a particular language—and in the ontological vein—in which the denotation of

the terms of one's uncritically accepted language is a matter of fact (Gibson 1988, p. 143).

The idea is that we must conduct our epistemological investigation under an ontological

framework. And this ontological framework is nothing but our scientific enterprise. There

is, however, another problem regarding the choice between competing scientific theories

that are empirically equivalent to, but logically incompatible with each other, insofar as

this might undermine Gibson's account. That is, the thesis of under-determination is now

the problem.[5] It is very likely that incompatible theories formulated in the same language

---

[5] The thesis of under-determination in the broadest and probably most appropriate sense is: providing all
possible evidence or observation data, there are many, if not infinitely many, theories that are empirically

might have different reference schemes. Thus the aftermath of ontological relativity hits

Gibson's ontological framework after all. It is not surprising that Gibson does not think it

this way, since he and Dagfinn Føllesdal would like to persuade Quine to adopt his

sectarian position regarding competing comprehensive scientific theories. That is, "one of

two systems of the world must be deemed false if we know them to be empirically

equivalent" (Quine 1986*b*, p. 156). So we have to pin down our (scientific) theory no

matter what the reason is, and then we have *the* reference scheme in terms of the theory

we pick. But isn't the idea of ontological relativity to show that, for any given theory or

language, its universe can be permuted to map it on to something else or even itself? To

this, Gibson argues that we uncritical English speakers can of course *go critical*, refusing

to take 'rabbit', say, at its face value, but we need not do this. On the other hand, the field

linguist has to go critical in the case of radical translation because she does not have any

other choice (Gibson 1988, p. 144). Whether or not our *willingness* to go critical to our

home language has any essential power of overturning the presumably ultimate regress of

relativisation seems to be a minor point to him.

Gibson's interpretation of Quine's manoeuvre—that is, ending the regression of

background languages is, in practice, by acquiescing in our mother tongue and taking its

words at face value—is that the reference scheme is *fixed* by acquiescing to homophonic

translation in the home language (Gibson 1988, p. 139). This is of course in an

ontological vein and for Gibson it follows that what the terms of one's uncritically

accepted language denote *is* a matter of fact (p. 143). Indeed, Gibson even claims that

some, if not all, are behavioural facts. The uncritical speakers of English would probably

---

equivalent but logically incompatible. The explication of the notion of empirical equivalence is that theories
are empirically equivalent if and only if they imply exactly the same set of observational consequences.

assent to the query "Does 'rabbit' refer to rabbits?" while our indigenous friend does not

stand a chance to assent to the query "Does 'gavagai' refer to rabbits?"

This is a rather bizarre way to argue, because what Gibson shows is merely the fact that

the indigenous informant does not acquire the ability to speak English. It seems to be

more sensible to ask our field linguist whether or not she agrees that 'gavagai' refers to

rabbit. She is in any rate the expert who provides translation manuals. And following the

same reasoning of sectarianism, she can choose one of the competing manuals and thus

decide the reference scheme of her informant's language. This is probably very

controversial, because there seems to be a difference between the choice of a scientific

theory, which has facts of the matters to back it up, and the choice of a translation manual,

which does not. But this does not seem to cause any trouble if we adopt Gibson's

interpretation anyway.[6] If, presumably, she assents to our query "Does 'gavagai' refer to

rabbits?" what next? Is this a *behavioural fact* that can show the incorrectness of both of

the theses of inscrutability of reference and ontological relativity?[7] Surely not, and not

because this is not a behavioural fact, but because it is illegitimate. The major dispute so

far on either the thesis of inscrutability of reference or the thesis of ontological relativity

is that there is no way to determine how to use the semantic terms 'refer' or the like

uniquely. Therefore, "Does 'rabbit' (or 'gavagai') refer to rabbits?" is not a legitimate

question, because the only sensible way to ask this question is if there are some related

semantic facts to help us single out the usage of 'refer'. But according to Gibson, there

are no semantic facts.

---

[6] For the simplicity sake of argument, I will ignore this problem.
[7] Gibson is a full-blooded Quinean, so he surely does not accept the existence of semantic facts.

Moreover, even if uncritical English speakers do assent to the query "Does 'rabbit' refer to rabbits?" if the counterfactual consideration is dropped, it is dubious that they will assent to or dissent from this sort of queries uniformly. For instance, some, if not most, uncritical English speakers might assent to the query "Does 'Atlantis' refer to Atlantis?" or "Does 'phlogiston' refer to phlogiston?" Gibson does not show us that uncritical speakers of English would uniformly assent to every query of the form "Does 'n' refers to n(s)". Worse, if some uncritical English speakers do assent to the query "Does 'phlogiston' refer to phlogiston?" what should Gibson explain this *word-world* correspondence relation? Therefore, there is no such demarcation as Gibson expects to be between staying uncritical or going critical to one's home language. In Section Four, I will show that what Gibson shows us is not a word-world correspondence reference scheme of a given language constituted by taking words at their face value, but a disquotational schema of reference that is essential to our daily communication as well as scientific discourse.

To be exact, the gravity of Quine's indeterminacy thesis is the denial of semantic facts that can *absolutely* determine certain semantic features such as in the case of radical translation, and all the related physical facts can do no better on this point. Field's position is that, even if we grant the assumption that there are no semantic facts, the thesis of ontological relativity does not follow. Consequently, the reference scheme of a given theory or language is neither explicable in the absolute sense of a word-world relation nor in the relativised sense of a word-word relation. This predicament can be solved, Field claims, by introducing the causal relations between words and extralinguistic objects and thus maintaining a more general sense of word-world correspondence relations between

them. Surprisingly, Gibson is eager to retain this word-world correspondence relation[8] and notices that the point Quine made in "Reply to Eleven Essays" (Quine 1981*a*) is so deserving of putting out the welcome mat—namely, "I use the word ['refer'] to relate linguistic expressions to objects. *Real* objects." (p. 243). Gibson's main position is as follows:

> However, so long as linguists *A* and *B* continue to conform unquestionably in their usages of 'rabbit' and 'undetached rabbit parts', just so long do 'rabbit' and 'undetached rabbit parts' refer, respectively, to rabbit and undetached rabbit parts. And this…pair of relations, unlike the pair of relations 'gavagai'/'rabbit' and 'gavagai'/'undetached rabbit parts', is *not* of the word-word type but of the word-world type: 'rabbit'/rabbit, 'undetached rabbit parts'/undetached rabbit parts. Quine's point is that these…word-world relations, too, can be disturbed, but they need not be. And so long as they remain undisturbed, it makes perfectly good sense to say that 'rabbit' refers to (i.e. denotes) rabbits and 'undetached rabbit parts' refers to (i.e. denotes) undetached rabbit parts. (Gibson 1988, p. 140)

I will render a different interpretation of Quine's manoeuvre. But I would like to first point out how misleading Gibson's explication is.

If, presumably, there are word-world correspondence relations between words and extralinguistic objects, *real* objects, in our own language, and others can only have a relativised word-word relations in their interpreting reference schemes of my language, then it seems to make *perfectly good sense* to claim that this establishes a solid case of private language. After all, nobody is in principle able to grasp my understanding of my own language, except me.

Of course, there is an important difference that might cause some trouble to my

---

[8] Gibson certainly does not accept Field's correspondence account of truth, but holds that the thesis of ontological relativity does not preclude a correspondence theory of truth, at least in some general sense of 'correspondence' such as Davidsonian (Gibson 1988, p. 140). This is not entirely wrong. But I can hardly accept Gibson's reasoning.

criticism—namely, I use 'my own language' instead of 'our home language'. This is, however, exactly the point. If it is the case that the reference scheme of *our* language is fixed so that 'rabbit' only refers to real rabbits, then so is the reference scheme of *my* language. But how do I know that other people who speak English uncritically do use 'rabbit' to refer to rabbits? If there are no semantic facts about this then there must be physical or, probably, behavioural facts to help me figure it out. It nevertheless follows that there should also be physical or behavioural facts to pin down the reference scheme of any foreign language. As Quine says in "Ontological Relativity":

> If it is to make sense to say even of oneself that one is referring to rabbits and formulas and not to rabbit stages and Gödel numbers, then it should make sense equally to say it of someone else. After all…there is no private language. (p. 47)

My argument is clear. If the reference scheme of my language can be fixed by a set of word-world correspondence relations as Gibson expects it to be, there must be either semantic or physical/behavioural facts to support that. And similarly my fellow English speakers should equally be able to fix their reference scheme. In fact, we should have exactly the same reference scheme in which 'rabbit' refers to real rabbits in a word-world correspondence sense. Since Gibson holds that there are no semantic facts, this *fix* of ontology has to depend on physical or behavioural facts. Just what kind of facts, then, fixes the ontology of a home language but not the one of a remote language? Certainly not Gibson's *behavioural fact* that our jungle informant *fails* to reply affirmatively to our query "Does 'gavagai' refer to rabbits?" As I said before, our informant fails to respond otherwise rather because he takes our question as gibberish. The difference between home and remote languages in this sense should only be a matter of *degrees* but not of *kinds*. If Gibson's position implies that one's home language is privileged over other translated

languages, it follows, however, that all related indeterminacy theses fail. But this means that there are semantic facts, and this is contrary to Gibson's original supposition. Therefore, there is no way to fix the reference scheme of our language as well as our informant's. If Gibson insists on the point that the difference is a matter of kinds, then he has to respond to the question that what kind or kinds of physical/behavioural facts can fix our ontology but not our informant's. Moreover, he also has to give a satisfactory answer to a more general question: how is it that I can determine my fellow speaker's reference scheme? and what is the difference between this determination and the one in the case of radical translation? If the answers are not satisfactory, especially to the second question, Gibson has to face the charge, even from Quine, of conceding to the existence of private languages.

Quine also makes an analogy between staying aboard/rocking the boat and his manoeuvre (ending the regression of background languages is, in practice, by acquiescing in our mother tongue and taking its words at face value)/contemplating a permutational mapping of our language on itself or undertaking translation in "Things and Their Places in Theories" (Quine 1981*b*, p. 20). Gibson's reading of this analogy is that reference does not go inscrutable unless we rock the boat. Otherwise there is a fact of the matter to questions like "Does 'rabbit' refer to rabbits?" in our home language (Gibson 1988, p. 145). That 'fact of the matter' is supposedly semantic rather than physical or behavioural. What I am not sure is whether or not Quine himself is absolutely consistent on the issues of inscrutability of reference and ontological relativity. But it is almost certain that Gibson's account, which assigns every referential term of our own language a word-world correspondence relation to a specific object or set of objects, causes a discrepancy

78

between the relativisation of reference scheme in the case of mother tongue and the relativisation of the reference scheme in the case of remote languages. This is exactly what Field argues against. And this is just what we cannot accept. The staying aboard/rocking the boat analogy should have a more sensible interpretation.

4. Disquotational Schema of Reference and Semantic Ascent of Terms

Davidson approves of Field's criticism that there is no intelligible way to justify the concept of ontological relativity in either an absolute or a relativised sense. For him, there is no intelligible way to stride the first step in a so-called relativisation of ontology. The original reason to relativise the reference scheme of a given language or theory to another is that no empirical evidence can fix its reference scheme. That very problem, however, materialises again in the interpreting language or theory. This occurs accordingly in every attempt to relativise any given language or theory. It is then hard to see why the ontological regress should have gone on in the first place. Except for this, Davidson does agree that it makes sense to say about a relativised concept of ontology. His reasons are twofold. On the one hand, all acceptable reference schemes of a given language for a speaker or a community have a common feature that they ascribe equivalent truth conditions to all sentences. On the other hand, these schemes are distinguishable from each other not on empirical grounds, but because they employ predicates that cannot have the same extension (Davidson 1979, p. 238). This relativised concept of ontology, however, only helps us answer questions *about* reference but not reference itself (p. 239). This is Davidson's standpoint on the thesis of ontological relativity.

My interpretation of Quine's manoeuvre is not incompatible with Davidson's position on

the thesis of ontological relativity. I think that Davidson makes a more feasible and more prudent interpretation than Gibson does. Quine does not offer us an account of word-world correspondence relations between words of our home language and extralinguistic objects. When he says that the use of 'refer' is to relate linguistic expressions to real objects, he does not pin down the reference scheme of English in which 'rabbit' refers to *real* rabbits rather than *real* undetached rabbit parts or anything else. This is apparent in Quine's reply to Paul A. Roth (Quine 1986*a*) in *The Philosophy of W. V. Quine*:

> Within the home language, reference is best seen (I now hold) as unproblematic but trivial, on a par with Tarski's truth paradigm. Thus 'London' designates London (whatever *that* is) and 'rabbit' denotes the rabbits (whatever *they* are). (p. 460)

Those *italics* are the key to my interpretation.[9] In Quine 1953*a*, Quine holds that, except the truth paradigm, Tarskian paradigms on 'is true of' (that is, 'denote' or 'signify') and 'name' relative to a given language are also sustainable (pp. 134-6). They can be characterised as follows:

(1) '__' is true-in-*L* of every __ thing and nothing else;

(2) '__' names-in-*L* __ and nothing else.

(1) holds when any one general term is written in the two blanks and (2) holds when any name is written in the two blanks. According to these three paradigms, I think we should interpret Quine's manoeuvre as *semantic ascent* of terms—as offering a disquotational schema of reference to one's own language. In "Quine and the Correspondence Theory" (1974), Field thinks that our disquotational schema of reference is incompatible with indeterminacy of translation. In the "Postscript" (2001) of this paper, he disagrees with himself on this point. However, he does not change his mind on his criticism to Quine's

---

[9] The same point is also shown in "Truth and Disquotation". Please cf. the previous chapter.

ontological relativity thesis. It follows that he at least does not go in for this disquotationalist approach.

Some commentators overlook the significance of this point. For instance, Hylton in "Quine on Reference and Ontology" (Hylton 2004) maintains a similar point to Gibson's. After his citation of Quine 1986*a*, (cited above in the previous paragraph) he says:

> The point here, I think, is that while we are speaking our familiar language (or a language whose translation into it is well established), the inscrutability of reference simply gets no grip. (p. 142)

But remember Quine says that 'London' designates London (whatever *that* is) and 'rabbit' denotes the rabbits (whatever *they* are) right after he says that reference within the home language should be best seen *on a par with Tarski's truth paradigm*. He sees T-sentences in a disquotational sense, so is that 'London' designates London (whatever *that* is) and 'rabbit' denotes the rabbits (whatever *they* are). Similar to (D$_Q$), of course, the term quoted and not quoted in an instance of the disquotational schema of reference have to be regarded as the same word type with the same interpretation.[10]

This is probably a bit reminiscent of Stephen Leeds' account of ontological relativity in "How to Think about Reference" (Leeds 1973). Leeds takes the thesis of ontological relativity to amount to a philosophically uninteresting triviality that (semantical) questions, whatever their subject matters are, must be asked relative to a background language (p. 488). He continues his argument in the following way:

> To speak of reference, we must *pretend* that certain questions about our language have been answered—these questions are questions about the reference of the terms in our language; the answer is given by the Tarski schema for reference in the background

---

[10] Horwich holds the same view. Please cf. the following section.

theory. (p. 490, my *italics*)

This pretending has its practical value as a result of its 'harmless convenience'. The

ontological talk, in which the subject matter is not the relations between language and the

world but some important relations such as intertranslatability held between languages, is

in principle possible without ever speaking of reference (pp. 491-2). This account of

ontological relativity is nevertheless incompatible with Quine's naturalism in that the

existence of objects depends on our ontological commitments and they must in turn reside

within a (scientific) theory. Surely the ontological talk is not about relations between

languages but between language and the world. Or better, it is about relations between

theories and the world.[11] Theory as a whole does owe its ontology to the world in the

sense that its observational consequences have to be compatible with the world observed.

It is only when theories are empirically equivalent, the ontological talk seems to be

blurred. To develop an interpretation of Quine's philosophy accordingly, by assuming that

there is an objective position to be occupied to claim the *truthfulness* of any theory, or that

the importance of ontology is not about the relations between language and the world but

between languages, is an utter misunderstanding. Except for this, Leeds' account does

offer a valuable observation that Tarskian paradigms have to play a role of great

magnitude in the interpretation of Quine's thesis of ontological relativity. But it is not a

matter of *pretending*, but of offering a disquotational schema of reference to make

possible ontological talk. And all the instances of this schema do not constitute any

genuine reference scheme that can possibly be given.

---

[11] Quine does have the inclination to use 'language' and 'theory' interchangeably. Another familiar
terminology is 'conceptual scheme' that causes a debate between Davidson and Quine. (That is, "On the
Very Idea of a Conceptual Scheme" and "On the Very Idea of a Third Dogma".) This is however not among
the main issues here.

Constructing a disquotational schema of reference is not doable in translating or interpreting a language other than the translator or interpreter's home language. For instance, "'Schnee' refers to snow" is not an instance of a disquotational schema of reference in English.[12] This is already apparent in the discussion of the designation schema in the previous chapter. Therefore, to take the disquotational schema of reference as the proper interpretation of Quine's manoeuvre is the only way to make sense of it.

A disquotational schema of reference is crucial to the thesis of ontological relativity and even the thesis of ontological commitments.[13] There are at least two aspects of the thesis of ontological relativity, and the disquotational schema of reference is essential to both. One way to see this is in terms of so called notational relativity—that is, the universe of our language or theory can be permuted to map it on to itself. In this aspect, the universe is kept intact. So the relativity comes from the different notational permutation between referential terms and the members of the given universe.[14] However, it is inevitable for the notational permutation to depend for its significance on a pre-established disquotational schema of reference. For instance, if 'snow' under the permutation refers to milk instead of snow, this can only make sense if 'snow' refers to snow and 'milk' refers to milk disquotationally. Otherwise we have difficulties in understanding what the word 'milk' stands for in the case of permutation. Another way to see ontological relativity is in terms of the inscrutability of reference in which the universe is varied, as pointed out in

---

[12] However, in terms of an adequate translation manual "'Schnee' refers to snow" might be conceived of indirectly.

[13] Again, I do not intend to defend Quine's viewpoint on ontological commitments. For instance, Azzouni in "On "On What There Is"" urges us that there might be some other criteria for recognizing what a discourse commits us to (CRD in his abbreviation). He might be right, but this does not concern me anyway.

[14] This notational permutation is not necessary on a global scale. And the permutation on a small scale reminds us of the reconstrual of predicates introduced by Quine in "On Empirically Equivalent Systems of the World".

the first section. For any given theory in which a general term, say, 'rabbit' refers to rabbits according to its reference scheme, one can produce a proxy function, say, ø to map rabbits on to ø(rabbits) or , say, undetached rabbit parts in the reference scheme of the background theory. Similarly, to make sense of this we have to rely on reading "'Rabbit' refers to rabbits" and "'Undetached rabbit parts' refers to undetached rabbit parts" disquotationally.

In the case of ontological commitments, the crucial nature of the disquotational schema of reference is even more apparent. When Quine says that the answer to the question of whether or not there is an F, say, depends on our decision to assign or not assign a value to the variable of an objectual quantification "$\exists x(Fx)$" in a regimentation of our scientific theory into the language of first-order logic, he certainly does not specify what 'F' refers to. In the above case, what 'F' refers to is of course F(s). That is, if 'F' amounts to 'rabbit' then 'F' refers to rabbits, or if 'F' amounts to 'undetached rabbit parts' then 'F' refers to undetached rabbit parts, and so on. Indeed, the thesis of ontological commitments tells us that the comprehensive scientific theory held by us commits us to certain sorts of objects, but what it does not tell us is what these objects really *are*. This is why in Quine 1995 Quine says:

> I conclude…that what matters for any objects, concrete or abstract, is not what they are but what they contribute to our overall theory of the world as neutral nodes in its logical structure. (pp. 74-75)

Without the disquotational schema of reference, none of this is possible.

In *Word and Object*, Quine holds that semantic ascent of terms usually happen when we are talking about some abstract objects such as miles or points. In the case of physical

objects such as wombats or kangaroos, this normally would not happen. The usefulness of semantic ascent of terms is as follows:

> The strategy of semantic ascent is that it carries the discussion [of what there is] into a domain where both parties are better agreed on the objects (viz., words) and on the main terms concerning them. (p. 272)

My borrowing of this strategy is not entirely unreasonable. In the same page, Quine states: "Semantic ascent, as I speak of it, applies anywhere." He also claims the following: "we have talked more of words than of objects even when most concerned to decide what there really is: what objects to admit on our own account" (p. 270). In speaking of Quine's manoeuvre and the thesis of ontological relativity, this strategy is essential. Without it the manoeuvre fails miserably. And this can be well grasped in my illustration of the disquotational schema of reference. We simply substitute talk of terms for talk of extralinguistic objects. What should not be confused with the disquotational schema of reference is that we do not render any fixed reference scheme of our mother tongue. Inscrutability of reference and ontological relativity are inevitable in doing empirical semantics. They represent the mystery of the semantic concept of reference. Conversely, the disquotational schema of reference and satisfaction are the reason why scientific research can still be carried on without being disturbed by the semantic mystery.

If the ontology of one language or theory has to be reinterpreted in a background language or theory so that it can be understood by others, and this process is infinite, then, presumably, one has to understand all these theories or languages in order to understand the original ontology. This makes the thesis of ontological relativity the most inflationary theory ever, and the most inflationist theory one can imagine.[15] It seems that the

---

[15] This is inspired by A. Gupta's criticism to P. Horwich's minimalistic theory of truth in "Minimalism" (collected in *Philosophical Perspectives*, Volume 7, Issue Language and Logic, p. 365).

disquotational schema of reference can provide a shelter for the thesis of ontological

relativity from this criticism. It is also crucial to scientific theories so that they can have a

significant way to talk about their reference schemes without the infinite regression of

reinterpretation.[16] This schema is also compatible with Quine's general view on ontology.

After all, he states:

> Structure is what matters to a theory… not the choice of its objects… The objects, or
> values of variables, serve merely as indices along the way, and we may permute or
> supplant them as we please as long as the sentence-to-sentence structure is preserved.
> (Quine 1981*b*, p. 20)

Therefore, the disquotational schema of reference can be seen as a part of the more

general account of disquotationalism.

In "Ontology and Ideology" (Quine 1951), Quine explains the distinction between

ontological and ideological commitments as follows:

> Given a theory, *one* interesting aspect of it into which we can inquire is its ontology: what
> entities are the variables of quantification to range over if the theory is to hold true?
> Another no less important aspect into which we can inquire is its *ideology* … what ideas
> can be expressed in it? (p. 14)

One of the values of the disquotational schema of reference is to help us axiomatise some

familiar mathematical theories without question begging (Quine 1960, pp. 272-273). In

other words, by going up a level we can separate the target theory from its ideology. For

instance, in axiomatisation of a geometrical theory, one is in trouble with not being sure

that one deduces some familiar formulas of the axiomatised theory purely from its axioms

without utilising any part of further geometrical knowledge. To avoid this unwelcome

embarrassment one can turn to the device of disinterpretation, acting as if one only

---

[16] Marian David in *Correspondence and Disquotation* offers a disquotational account of reference to
contribute to the recursive definition of disquotational truth (pp. 110-119). My interpretation, on the other
hand, shows that there is a distinct need to provide a disquotational schema of reference.

understands the logical vocabulary but not those peculiar terms with regard to the axiomatised geometrical system. The problem with this method, according to Quine, is that it leads us to the following consequence: "we never know what we are talking about, nor whether what we are saying is true."[17] Quine then claims that, with Frege's achievement of a full formalisation of logic, semantic ascent as a device against question-begging turns out to be available (ibid., p. 273). He then says:

> Given the deductive apparatus of logic in the form specified operations on notational forms, the question whether a given formula follows logically from given axioms reduces to the question whether the specified operations on notational forms are capable of leading to that formula from the axioms. (ibid.)

What concerns us in the axiomatisation of a mathematical theory is the danger that its ideology is always inseparable from the axiomatisation. Semantic ascent as described above shows a way to avoid this danger.

Another value is in the evaluation of a neoteric scientific theory (Quine 1990*b*, p. 81). Quine's example is Einstein's special and general theories of relativity by which our basic conceptions of space and time are challenged tremendously.[18] The evaluation of this new theory is cumbersome with or without the ideology of Newtonian physics. Without this particular ideology, the evaluation would not proceed in the first place. With the help of this ideology, however, the evaluation begs the question. It is just like taking the premise that kanji (which is both the pronunciation of 'Chinese character' in Chinese and in Japanese) is purely a cultural product of Japan, to evaluate whether or not Japanese kanji is under the influence of Chinese characters. In terms of the semantic ascent of terms, Quine then says:

---

[17] This is Quine's citation of Russell's *Mysticism and Logic and Other Essays* (p. 75) in *Word and Object* (p. 273).
[18] A related point is also made in Quine 1960, p. 272.

But by semantic ascent one could compare the new and old theories as symbolic structures, and so appreciate that the new theory organized the pertinent data more simply than the old. Simplicity of symbolic structures can be appreciated independently of those basic conceptions. (ibid.)

In other words, by taking the special and general theories of relativity and Newtonian physics as symbolic structures, one can compare the two with regard to which sentence-to-sentence structure is capable of organising empirical data more simply.

5. *Disquotational Schema of Reference and Disquotational Theory of Truth*

In the last chapter, I considered two reasons to reject the idea of incorporating inductive axioms of reference and satisfaction into the formulation of a disquotational theory of truth. The central issue there was what sort of theory of reference for a given natural language we can get to obviate the potential difficulties of our axiomatised disquotational theory of truth suggested by the incorporation.[19] The first reason is that the theory of reference involved has to deal with the explanatory order of semantic concepts, as indicated there that this is crucial to Quinean semantics. The second reason is that the theory of reference has to be without any substantive feature, so that the resultant inductive axioms are be able to properly imply nothing but all disquotation sentences of the truth language within which the disquotational theory of truth of the base language is constructed. It seems that, now, we have a candidate here, namely, the disquotational schema of reference.

Nelson is probably one of the philosophers who might object to this reasoning, but for the

---

[19] As stated there, the incorporation of inductive axioms only arrests our interest when this is about to give a formulation of disquotational theory of truth for any given natural language.

wrong reason. The main rationale for incorporating inductive axioms into the disquotational theory of truth is to render a way to explain *how* the semantic ascent of sentences helps express some infinite conjunctions or disjunctions with the same syntactic feature. The disquotational schema of reference is not supposed to amount to explicit definitions of the semantic concepts of reference and satisfaction. If it were, then it would have violated the explanatory order between semantic concepts. Since it is not, it follows that all disquotation sentences are material biconditionals. This, however, according to Nelson (1997), causes a problem of compatibility between the disquotational theory of truth and the thesis of ontological relativity. In the case of notational relativity in which the universe is not varied, suppose one produces a proxy function, say, ø to map *a* on to ø(*a*) in the reference scheme of the background theory and similarly to map a predication F(*a*), that F refers to (is true of) each thing *a* such that F(*a*) in the object theory, to Fø(*a*) in the background theory. The T-sentence "'F(*a*)' is true iff F(*a*)" is a material biconditional according to the above analysis. Nelson indicates that, in terms of ontological relativity, according to which F(*a*) iff Fø(*a*), we can substitute 'Fø(*a*)' for 'F(*a*)' in the T-sentence. The resultant biconditional is "'F(*a*)' is true iff Fø(*a*)" and it is certainly not a T-sentence. This is apparently not a problem to my formulation of the disquotational theory, (D$_Q$), in the last chapter, for, after substituting 'Fø(*a*)' for 'F(*a*)' in "'F(*a*)' is true iff F(*a*)", "'F(*a*)' is true iff Fø(*a*)" is not a disquotation sentence anyway. It is the conclusion of two premises—namely, the biconditional "F(*a*) iff Fø(*a*)", resulting from using the proxy function 'ø' to map F(*a*) on to Fø(*a*), and the disquotation sentence "'F(*a*)' is true iff F(*a*)".

The best way to see the disquotational schema of reference in general is to see it as based

on the pre-theoretic intuitive concepts of reference and satisfaction of any given natural

language. It does seem right, intuitively, that speakers of any language incline to

understand the pre-theoretic referential and satisfaction relations of their languages in the

disquotational sense. For instance, every English speaker normally accepts the

disquotational schema of reference with no regard to whether or not the referent, say, of

the legendary city of 'Atlantis' or the mythical creature 'Pegasus' exist. If there are words

such as 'rabbit' and 'London' in English as well as words like 'phlogiston' and 'Vulcan',

people are ready to accept that 'phlogiston' denotes phlogiston and 'Vulcan' refers to

Vulcan as far as they are ready to accept that 'rabbit' denotes rabbits and 'London' refers

to London. There are no ontological reflections involved here.

The problem of incorporating inductive axioms of reference and satisfaction in

formulating a disquotational account of the semantic concept of truth, is that it is difficult

to find a suitable theory of reference to account for the explanatory order and to enable

the inductive axioms to imply *all* disquotation sentences. Providing that the thesis of

ontological relativity holds, Halbach's idea of incorporating inductive axioms of reference

and satisfaction to enhance the deductive power of the disquotational theory of truth is

therefore not satisfactory. For no matter whether or not the disquotational schema of

reference goes finite or infinite, it is not a genuine reference scheme. Otherwise, it

contradicts the thesis of ontological relativity.[20] Even if we assume that the thesis of

ontological relativity is untenable, to consider a disquotational schema of reference as

providing a genuine reference scheme is problematic. This will be shown in the following

---

[20] McGee is another philosopher who is in favour of accounting for disquotational truth in terms of the
notions of reference and satisfaction. Although he does not accept Quine's indeterminacy theses, the
implausibility of his account will not be discussed, even briefly, until Chapter Five.

discussion of Horwich's reasoning. And the criticism to Horwich's account is equally

applicable to Halbach's.

Horwich is another important advocate of a disquotational theory of reference and

satisfaction. According to him, Quine's thesis of ontological relativity does not hold.

Hence he can consistently see the instances of the disquotational schema of reference (and

satisfaction) as providing a genuine reference scheme. That is,

> The rough idea is that our meaning what we do by the word "refers" consists in our
> disposition to accept sentences such as
> Tokens of *London* refer to London,
> Tokens of *the highest mountain* refer to the highest mountain,
> and so on. Since an *ambiguous* singular term, "*n*", when used as these examples to
> designate a referent, must be understood in just one of its senses, it will not (if any
> sense of "*n*") be true to say that *every* token of "*n*" (individuated phonologically) refers
> to *n*. This is why the expression-types deployed in disquotational reference
> specifications are articulated using *star* quotation marks—indicating that they are
> individuated…in the basis of meaning.
> Thus our meaning what we do by "refers" consists in our inclination to accept
> instances of the disquotational schema
> Tokens of *n* refer to *n*,
> where what are substituted for the two occurrences of "*n*" are understood in the same
> way.[21] (Horwich 1998*a*, p. 118)

In the following chapter I will show that Horwich's rejection of the thesis of ontological

relativity fails. So we could tentatively say that one can always find a proxy function ø to

map *n* of the disquotational reference scheme on to ø(*n*) in the reference scheme of the

background theory, and similarly to a predication F(*n*).

---

[21] Horwich does not use 'disposition' and 'inclination' indifferently. In a footnote (23) of "A Defense of
Minimalism" (Horwich 2001*a*), Horwich makes it clear that he distinguishes *inclinations* from *dispositions*
to accept instances of the equivalence schema because of consideration regarding the liar paradoxes.
However, it seems that he does use them indifferently here.

But we do not have to. Horwich's account has the difficulty of properly explaining all relevant facts regarding reference. Let me first make some clarification of his schema of reference. For Halbach's account, the problem is that it falls short of explaining some disquotation sentences such as "'Pegasus is a flying horse' is true if and only if Pegasus is a flying horse." Horwich's account, on the other hand, is immune to this criticism, since he does not attempt to explicate disquotational or minimalist truth in terms of the inductive axioms of reference and satisfaction. He is fully aware of this sort of problem, and hence he modifies his schema as follows:

(R) Tokens of *n* refer, if at all, to *n*.[22] (Horwich 1998*a*, p. 119)

He further provides some principles to help form a corresponding schema accounting for our attribution of reference to terms of foreign languages. To simplify my discussion, (R) can be restricted as only applying to referential words of a home language with a fixed interpretation, when indexical, vague, and ambiguous words along with other related matters are precluded.[23] An account of reference surely need not tell us which words refer and which do not. But it has to tell us why some words refer and why some do not. For instance, a causal theory of reference would explain that a word refers if and only if a causal relation between it and its referent can be established. Otherwise, the word does not refer. Since Horwich regards it as false that 'Atlantis' and 'the largest prime number' refer to Atlantis and the largest prime number, because of there being no such things to be referred to, the 'if at all' in (R) is to make the demarcation between words that refer and those that do not. The problem of Horwich's account arises from this additional condition

---

[22] The sentence "Pegasus is a flying horse" is a declarative sentence, and according to Horwich's equivalence schema, there is a corresponding proposition expressed by it. As a result, that the proposition that "Pegasus is a flying horse" is true if and only if Pegasus is a flying horse is an axiom of the equivalence schema of propositional truth. It is clear then why Horwich does not and would not accept Halbach's approach.

[23] Horwich's disquotational schema of reference and all its axioms are simplified in notation as "'*n*' refers to *n*" hereafter.

stood for by 'if at all'. For instance, what does the 'if at all' stands for in the case of

"'Phlogiston' refers, if at all, to phlogiston"?

Horwich's formal characterisation of (R) might be helpful; that is

(R1) (x) (Tokens of $*n*$ refer to x iff $n = x$). (ibid.)

This is similar to his formulation of the disquotational schema of reference in *Truth*; that

is:

(R2) (x) $(<d>$ refer to x iff $d = x$),

     where '$<d>$' stands for word types. (Horwich 1998*b*, p. 116)

Apparently, if (R1) is the formal characterisation of (R), then "Tokens of $*n*$ refer to x" at

the left hand side of (R1) is definitely meaning that tokens of $*n*$ refer *disquotationally* to

x. Otherwise, that tokens of $*$morning star$*$ refers to evening star if and only if morning

star $=$ evening star is an instance of (R1). But this is not an instance of (R).[24] The

deflationary account of reference, according to Horwich, admits no substantive relation of

reference. Although there is really a relation of reference, this relation is without any

underlying nature (Horwich 1998*a*, p. 123). And all the facts regarding reference can be

explained by all the axioms/instances of the disquotational schemata of reference and

satisfaction (ibid.). Thus the 'if at all' cannot be explicated as encapsulating within it that

tokens of $*n*$ (or $<d>$) bear some non-semantic relation r to something x. This is not just

inflationary and circular, but absurd. (R) is then saying that, if there is really an object to

which tokens of $*n*$ bear a substantive referential relation, tokens of $*n*$ refer to n. The

absurdity rests on the assumption that a substantive referential relation always boils down

to a disquotational referential relation, so the disquotational reference scheme *is* the only

---

[24] Moreover, 'if at all' in (R) cannot be read as that if tokens of $*n*$ *refer* they refer to n. For one should rather regard the account of whether or not tokens of $*n*$ *refer* as the account of reference, and it is far from clear that this account would be not substantive.

reference scheme. Therefore, what Horwich has in mind must be that some reference-free auxiliary explanatory resources can explain, not which word refers, but whether or not something that is referred exists. For instance, if there is a suitable explanation of the existence of tachyon, then 'tachyon' refers to tachyon. There is no suitable explanation of the existence of Atlantis, so there is nothing to be referred by 'Atlantis'. This, however, leads to three problems.

First, according to what criterion does one acquire the suitable explanatory resources by which what there really is, is explained? It cannot be our commonsense, for it is very unlikely that commonsense provides all the answers of what there is. It seems that we do not have other options but to regard *a* comprehensive scientific theory as the auxiliary explanatory resources. But it is more than apparent that different theories are very likely to commit us ontologically to different entities. For instance, for a disquotational schema of reference based on a comprehensive theory within which the theory of phlogiston is a branch, 'phlogiston' refers, and refers to phlogiston. So we probably have to resort to the notion of the division of linguistic labour to depend for the suitable selection of our comprehensive scientific theory on experts' professional knowledge.[25] This leads us to the second problem.

It seems that Horwich's deflationary account of reference, based essentially on the disquotational schema of reference and satisfaction, cannot explain that in some cases different theories are actually logically equivalent, if not identical, but formulated in

---

[25] In Horwich 1998*a*, Horwich borrows Putnam's idea of "the division of linguistic labour" to explain why the regularities of experts, with different specialised knowledge, are the paradigm samples of the regularities of the use of all kinds of words. So it seems quite reasonable to use this notion to explain why some words do not refer in terms of some experts' specialised knowledge of whether or not there is something to be referred to.

different languages, or even the same language. For the auxiliary explanatory resources are supposed to be born from *the* selected comprehensive theory. Suppose there are two comprehensive theories T_1 and T_2, and according to both of them there are electrons and molecules. Accordingly, that 'electron' refers to electrons and that 'molecule' refers to molecules are axioms of the disquotational theory of reference. Suppose it is further found out that T_1 and T_2 are logically incompatible in the sense that what T_1 says about electrons is about molecules in T_2 and what T_1 says about molecules is about electrons in T_2, and vice versa. In other words, Quine's reconstrual of predicates can render T_1 and T_2 to be logically equivalent.[26] Regarding a disquotational reference scheme as a genuine reference scheme seems to undermine our strategy of reconstrual of predicates. This strategy, according to Quine, does not only work for one pair of predicates but many (Quine 1975, p. 320). It seems really difficult for Horwich's disquotational account of reference to utilise the strategy of reconstrual of predicates to discern the logical equivalence of the two theories. Further, some cases are much less trivial than the reconstrual of 'electron' and 'molecule'. For Quine, the reconstrual of predicates is applicable to even Poincaré's instance that one formulation of cosmology takes space as finite, but depicts all objects as shrinking in proportion as they move away from centre, while another formulation takes space as infinite. This is not doable if one considers Horwich's disquotational schema of reference as providing a genuine reference scheme.

---

[26] For reconstrual of predicates, Quine says:
> [T]he general way of reconstruing an *n*-place predicate is by supplying an open sentence in *n* variables, not caring whether there happens to be a word in our language with the same extension as the open sentence.
> By a *reconstrual* of the predicates of our language, accordingly, let me mean any mapping of our lexicon of predicates into our open sentences (*n*-place predicates to *n*-variable sentences). (Quine 1975, p. 320)

In the case of 'electron' and 'molecule', the predicates 'molecule' and 'electron' can be mapped to the respective open sentences '*x* is an electron' and '*x* is an molecule'.

Finally, the third problem is more severe than the first two. Horwich's disquotational

account of reference cannot account for some very important facts of reference. In the

case of evaluating a neoteric scientific theory discussed toward the end of last section, one

cannot properly evaluate any neoteric theory if one always sticks to the original theory.

The instance raised by Field in "Theory Change and the Indeterminacy of Reference"

(Field 1973) is a good example showing that some important fact of reference is not

explicable in terms of Horwich's account of reference.[27] That is, for Newtonian physics

'mass' disquotationally refers to mass, but for Einsteinian physics 'mass' is an ambiguous

term referring either to rest mass or to relative mass. Or, consider Akiba's instance in

"Can Deflationism Allow for Hidden Indeterminacy?" (Akiba 2002):

> The pre-Newtonians' term 'heavier than' is to be considered indeterminate between *more
> weighty than* and *more massive than* from post-Newtonian viewpoint; the
> pre-Newtonians themselves did not have this distinction.[28] (p. 225)

In other words, there are examples of hidden indeterminacy of reference in

pre-Newtonian formulation of physics.[29] It is hard to see how a deflationary theory of

reference such as Horwich's can account for this problem. For Horwich, 'mass' refers, if

at all, to mass; 'rest mass' refers, if at all, to rest mass; and 'relative mass' refers, if at all,

to relative mass. But how to explain the referential indeterminacy of Newton's term

'mass'? It is very likely that some referential indeterminacy is hidden in our formulation

of *the* selected comprehensive theory too, and it will always be possible for a future

theory to contain hidden indeterminacy. Horwich's disquotational account of reference is

too deflationary to account for referential indeterminacy. A disquotational schema of

---

[27] What interests me is only Field's instance of the comparison between Newtonian 'mass'/mass and
Einsteinian 'mass'/mass but not his opinions on the issues behind this instance.
[28] This is also a problem in Horwich's use-theoretic account of meaning. I will discuss this in the following
chapter.
[29] A hidden indeterminacy by definition is not a manifest indeterminacy of a given language, such as
vagueness, relative to a particular time period. The hidden indeterminacy is also objective—that is, it is not
decidable in terms of our awareness of it.

reference is useful in scientific reasoning, but serious problems emerge when it is seen as providing a genuine reference scheme. So I can conclude that a disquotational schema of reference cannot provide Halbach's account the needed reference scheme for explicating the disquotational concept of truth in terms of the inductive axioms of reference and satisfaction. And it is a mistake for Horwich as well as for Halbach to see a disquotational account of reference as providing a genuine reference scheme.

Final Remarks

This temporarily puts an end to my analysis of Quine's disquotational theory of truth. Many issues are left without discussions until Chapter Five in order for a more integrated analysis. Next on the agenda is Paul Horwich's minimalism. If one commits oneself to the intelligibility of the concept of a proposition, and indeed the existence of propositions, it seems that minimalism is attractive. In the next chapter, I will summarise what minimalism consists in and, more importantly, I will discuss Horwich's use-theoretic theory of meaning according to which the legitimacy of a minimalist theory of *propositional* truth can be well established.

Chapter IV

On Horwich's Minimalist Theories of

Truth and Meaning

Reading the first chapter of Paul Horwich's *Truth* (Horwich 1998*b*), one can almost feel

the enthusiasm in his attempt to show how concise the elaboration of the minimal theory

of truth is. The chapter has only eight pages to expound to the minimal theory. Of course,

however, the theory is not that concise after all. After the first chapter, there are six more

to go. Although very many articles are devoted to the debate over deflationary theories of

truth, it is not the task of this chapter to take a stand on these matters. Instead, this chapter

aims to offer a brief introduction of minimalism. For minimalism has little to distinguish

it from disquotationalism, especially in terms of my analysis in the following chapter,

except that for a minimalist like Horwich, the primary truth bearers are propositions.

This difference is, however, certainly significant with respect to the problem of the

expressibility of the truth predicate discussed in Chapter Two. For his theory of truth to be

philosophical interesting, Horwich has to show us that it is possible to come up with an

account of propositions without any involvement of the concept of truth. In considering

this issue, this chapter consists of four sections. The first section is a brief introduction of

the minimal theory of truth. The following section relates to Horwich's criticism of

Quine's scepticism concerning meaning. The third section analyses and criticises the

central idea of Horwich's use-theoretic account of meaning. And the final section

considers his account of compositionality.

1. The Framework of the Minimal Theory of Truth

The formulation of Horwich's minimalist account is indeed simple. Horwich (1998*b*)

claims that for any declarative sentence

(1) *p*

there is an equivalent sentence

(1') The proposition *that p* is true,

and *the equivalence schema* is given when combining (1) and (1') with a logical

connective of biconditional—that is,

The proposition *that p* is true if and only if *p*.[1]

Or following Horwich's way by putting the dummy symbol '*p*' into angle brackets to

represent a proposition, we can convert the equivalent schema into

(E) <*p*> is true if and only if *p*.

Instances of (E), such as "<Snow is white> is true if and only if snow is white", constitute

not the axioms but the expressions of them of the minimal theory of truth, for according

to Horwich, the axioms are also propositions.[2] He further explains that axioms such like

"<<Snow is white> is true if and only if snow is white>" can be seen as applying the

propositional structure or, in other words, the propositional form of "<<*p*> is true if and

only if *p* >" to the proposition "<Snow is white>".[3] (cf. Horwich 1998*b*, pp. 17-20)

There are three important features of the minimal theory, or, rather, of deflationary

---

[1] Just what sort of equivalence relation holds between (1) and (1') is worth pursuing. This will be reserved
for the next chapter.
[2] However, in the following the distinction between the expression of an axiom and the axiom itself is
dropped for the sake of simplification in some discussions.
[3] This is based on what Horwich calls the Fregean Principle. Please see section four of this chapter. In
addition, Horwich takes the propositional forms such like "<<*p*> is true if and only if *p* >" as universal
propositions.

theories in general. First, every instance of the equivalence schema is an axiom of the

minimal theory. Second, the entire conceptual and theoretical role of truth is completely

explicable in terms of these axioms and the equivalence schema (E) (Horwich 1998*b*, p.

5).[4] Finally, all the facts involving truth and certain other matters can be also explained

on the basis of the minimal theory, along with appropriate theories of those other matters

(p. 7). Disquotationalism, on the other hand, will be totally unattractive to philosophers,

and even undeserving the name of deflationism, if it does not conform to these three

features. The major task of this chapter, then, is to specify the major difference between

minimalism and disquotationalism. Except for this difference between them, and the

peculiar consequences of their divergence, the problems regarding Horwich's minimalism

and Quine's disquotationalism are similar. This is why these problems will be reserved for

the next chapter.

But first let us examine whether or not it is possible to formulate the minimal theory

differently. Horwich apparently 'goes infinite' when explicating his minimal theory, and

he explains why this is inevitable in accounting for semantic concepts (cf. Horwich 1998*b*,

the fifth objection). Matthew McGrath thinks otherwise.[5] Taking propositions as objects,

we can use objectual quantifications to form finite theories of propositional truth. In

"Weak Deflationism" (McGrath 1997*b*), following Horwich's idea, McGrath thinks that

propositional forms or structures are functions from entities to propositions, and are

expressed by schematic sentences (p. 73). For instance, the propositional form $<p \ \& \ q>$ is

---

[4] For some general fact regarding truth it is inevitable to employ schemata to do the explication.
[5] McGrath's theory is a modification of what Ernest Sosa proposes in "The Truth of Modest Realism"
(*Philosophical Issues* 3, *Science and Knowledge*, pp. 77-95). The finite theory proposed by Sosa is as
follows:
For all propositions *P*, *P* is necessarily equivalent to the proposition that it is true.
The corresponding principle for deriving instances is
(PE) If $<p>$ entails $<q>$, then if *p*, then *q*.

a function, CONJ, which, given propositions P and Q, returns a proposition CONJ (P, Q) (ibid.). Similar functions include DISJ for disjunction, IFF for bi-conditional, and so on. Most importantly, the truth predicate is seen as a function, namely, Tr. Tr is to be the propositional function that takes an entity and returns the proposition with respect to it, that it is true. (E) then can be seen as a function E-prop, which is decomposable into the functions IFF and Tr in such a way that for all propositions P, E-prop(P) = IFF(Tr(P), P) (p. 74). McGrath's finite theory, *weak deflationism* is as follows:

(FTP) For all propositions P, P is materially equivalent to Tr(P).

In order to derive instances of (FTP), McGrath introduces a *principle of material implication*:

(PMI) If *<p>* materially implies *<q>*, then, if *p*, then *q*.

If this is doable, then the minimal theory should be reformulated in this fashion, providing some unproblematic premises, such as "<Snow is white> = EXP('Snow is white')" and "Tr('Snow is white') = <'Snow is white' is true>" (ibid.).[6] But is it?

Whether or not the material equivalence relation between 'P' and 'Tr(P)' is feasible is not what concerns me here.[7] What interests me is whether or not (FTP) along with (PMI) successfully formulate the minimal theory. According to Adam Kovach, however, this is not the case. In "Deflationism and the Derivation Game" (Kovach 1997), Kovach argues

---

[6] McGrath also provides finite theories of sentence, belief, and utterance. The basic idea is to use a function 'EXP' to explain the relation between a proposition and the sentence expressing it. The function, EXP, "takes an entity as argument and returns the proposition it expresses (McGrath 1997*b*, p. 74). According to McGrath, theories of utterance and belief truth can be accounted for in terms of the theory of sentence truth. Details of this will not be discussed in the following. However, there are some complications of this account. According to McGrath, weak deflationism inflates sentence, utterance, and belief truths while keeping propositional truths deflated. That is, the first three need an inflation of the theory of meaning to ensure the EXP function is doing its work properly. Putting aside these complications, however, there is a serious flaw in this account, which will be discussed shortly.

[7] This will be discussed in Chapter Five.

that (PMI) has to either be a generalisation governed by substitution quantifications or go infinite case by case, for it contains both name positions, that is, the pair of '<*p*>' and '<*q*>', and sentence positions, that is, the pair of '*p*' and '*q*'. But McGrath rejects the idea of formulating the finite theory in terms of substitution quantifications. So it is, after all, inescapable that the finite theory has to go infinite.

For instance, in explaining the capacity of his finite theory in accounting for the general fact regarding truth that, for all propositions *P*, *Q*, if *P* entails *Q* and *P* is true, then *Q* is true, McGrath gives us the following argument. Three further principles are needed. That is,

(P1) For all propositions *P*, *Q*, if *P* entails *Q*, then *P* materially implies *Q*;

and

(P2) For all propositions *P*, *Q*, CONJ(<*P* materially implies *Q* >, *P*) entails *Q*;

and

(P3) If CONJ(<*p*>, <*q*>) entails <*r*>, then if *p* and *q*, then *r*.[8]

The argument runs as follows:

1. For all *P*, *Q*, CONJ(<*P* entails *Q* >, <*P* is true>) entails

   CONJ(<*P* materially implies *Q* >, <*P* is true>)   Assm., (P1)

2. For all *P*, Tr(*P*) is necessarily equivalent to *P*   Assm., Weak Defl.

3. For all *P*, *Q*, CONJ(<*P* entails *Q* >, <*P* is true>) entails

   CONJ(<*P* materially implies *Q* >, *P*)   1, 2

4. For all *P*, *Q*, CONJ(<*P* entails *Q* >, <*P* is true>) entails *Q*   3, (P2)

5. For all *P*, *Q*, CONJ(<*P* entails *Q* >, <*P* is true>) entails <*Q* is true>   2, 4

---

[8] (P3) is a consequence of Sosa's (PE), according to McGrath. Please see footnote 5.

6. For all *P*, *Q*, if *P* entails *Q* and *P* is true, then *Q* is true   5, (P3)

The problem comes from the dubious (P3), which has the exact problem of (PMI). It is obvious that in (P1) and (P2) we have objectual universal quantifiers and not in (P3). So within (P3) the '*p*', '*q*', and '*r*' not in angle brackets are sentences rather than propositions. Therefore, the quantifier occurring in 6 has to be a substitution one. Similar mistakes appear in every argument provided by McGrath while utilising (PMI) or principles in a similar fashion. In a footnote, McGrath indicates that Horwich goes for a principle similar to (PMI) in proving that minimalism can account for some general facts regarding truth.[9] What McGrath fails to appreciate is that Horwich consistently goes infinite in explicating minimalism. Most importantly, Horwich actually does not use a principle similar to (PMI). It is nothing but a typographical error.

However, Kovach's observation is not news to McGrath. In his *Longer Notes* of "Weak Deflationism", number 2 to be exact, McGrath refers to Davidson who writes in "The Folly of Trying to Define Truth" as follows:

> Why, though, does Horwich not try generalizing his schema by quantifying over propositions? The answer should be: because then we would have to view ordinary sentences as singular terms *referring* to propositions, not as *expressing* propositions. (p. 273)

Davidson's claim, according to McGrath, is perhaps just that Horwich's schema (E) is not possible to generalise without utilizing sentential variables—that is, the eligible generalisation of (E) is "For all *p*, <*p*> is true if and only if *p*". McGrath believes that this claim is falsified by his finite theory of propositional truth.

McGrath's belief is, according to Kovach, ill-founded. For this McGrath argues in "Reply

---

[9] Please cf. the second section of the following chapter for Horwich's argument.

to Kovach" (McGrath 1997*a*) that (PMI) can either be taken as inferential principles or as providing support for an unstated generalisation. (PMI) can be read as inferential principles permitting, for instance, the inference that is allowed from formula of the form "<*p*> materially implies <*q*>" to the corresponding formula "If *p*, then *q*" (p. 582). Or, alternatively, (PMI) can be read, for instance, in the case that true conjunctions have true conjuncts, as providing support for the following generalisation.

(G) For all *P*, *Q*, if <CONJ(*P*, *Q*) is true> materially implies <*P* is true>, then if CONJ(*P*,

    Q) is true, then *P* is true. (ibid.)

The supposed clarification being accomplished by reading (PMI) as inferential principles or providing support to an unstated generalisation does not throw light on any prospect of clarifying (PMI). Reading (PMI) as inferential principles has exactly the same problem mentioned above, namely, appealing implicitly to substitution quantifications. In *Between Deflationism and Correspondence Theory* (McGrath 2000), McGrath explains what he means by reading (PMI) as a principle of inference.[10] In terms of (PMI), as an inferential principle, one can infer "If Hesperus is a planet, then there is a planet" from "The proposition that Hesperus is a planet entails the proposition that there is a planet" (McGrath 2000, p. 31). But this falls short of its explanatory purpose. To avoid the utilisation of substitution quantifications in (PMI), one has to go case by case to include a corresponding inferential principle for each inference.[11] On the other hand, the only sensible way to read (G) without any involvement of substitution quantifications, it seems, is that, for all *P*, *Q*, if <CONJ(*P*, *Q*) is true> materially implies <*P* is true>, then IF(Tr(CONJ(*P*, *Q*)), Tr(*P*)).[12] But this reading is still about schematic sentences and is

---

[10] McGrath actually uses Sosa's (PEI) in his explanation. But the point is the same.
[11] It seems that (PMI) under this consideration boils down to Horwich's criticism of substitution quantifications. Please cf. Horwich 1998*b*, pp. 25-26.
[12] 'IF' is also a propositional function for 'if-then'.

totally different from (PMI). According to McGrath's notation, '$<p>$' means a proposition in (PMI) and 'p' does not. Otherwise '$<p>$' should be taken as '$<<p>>$', if 'p' means a proposition rather than a sentence, but '$<<p>>$' does not even make sense. If this is a correct interpretation of (PMI), then "if snow is white then grass is green" should follow from "$<$Snow is white$>$ materially implies $<$Grass is green$>$". Similarly, "if $<$snow is white$>$ is true, then snow is white" follows from "$<<$Snow is white$>$ is true$>$ materially implies $<$snow is white$>$". But 'snow is white' and 'grass is green' not in angle brackets are sentences. In order to derive instances from (FTP), (PMI) has to utilise substitution quantifications to do the generalising work.

After showing that the minimal theory is in no position to go finite, it is time to move on. It is clear that minimalism and disquotationalism diverge in virtue of appealing to different truth bearers. Horwich rejects the idea of taking sentence types of a given language as truth bearers, on the basis that the sentence type of English, say, "I am hungry" is true on one occasion but false on another.[13] It is the content of a sentence type that is true or false. Curiously enough he does not rest his criticism on the usual grounds that the notion of a language is elusive, or that the relation between a sentence type and its content is contingent. As indicated in the second chapter, the disquotational schema within the additional truth language is the necessary device to disquote a given sentence type of an interpreted language to perform semantic ascent. Thus the two sentences are the same sentence with exactly the same empirical content. An explanation for how we can do this is required, but it is not reasonable to simply assume that the sentence in

---

[13] Horwich does not use an indexical instance to make this point. But the instance he uses to indicate that it is the content that is true or false, is expressed by a Quinean eternalised sentence type "Oscar was hungry at midday on 1 January 1988" (Horwich 1998*b*, p. 16). The problem of indexicality, it seems, is at least one of the reasons for him to have reservations about taking sentence types as truth bearers.

quotation marks and the sentence disquoted have different empirical contents. The reason that Horwich does not fall for this line of criticism probably has something to do with a similar consideration.[14] For it is also indispensable for him to identify what a particular proposition is, in terms of an interpretation of the sentence type of a given language expressing it. It just does not do to say that a linguistic expression expresses a particular proposition, by adding "the proposition that" in front of it. However, Horwich's criticism does not seem fair because indexicality is everybody's problem.

Perhaps it does not have to be fair anyway. Intuitively speaking, if there is an appropriate theory of proposition, minimalism does seem to be preferable to disquotationalism. It has two advantages over disquotationalism. First, minimalism can provide a more general concept of truth; one applicable to linguistic items (regardless even of the language they are in) that *express* one and the same proposition. Second, the expressibility problem explicated in Chapter Two seems to be solved if the primary truth bearers are propositions. Apparently, all instances of the generalising proposition can be derived from it, and vice versa. At least, this was held at one time by Horwich, in Horwich 1998*b*.[15] Of course in the present case, this theory of proposition has to be completely independent from any involvement of the concept of truth. Otherwise it is circular. But, on the other hand, we should not forget that the virtue of disquotationalism is its absolute separation from any theory of meaning. In other words, given *an* interpretation of all sentences of a given

---

[14] When explaining how to take utterances as truth bearers (Horwich 1998*b*, p. 100-101), Horwich uses the device of star marks to indicate that *I am hungry*, say, is a sentence type. He further derives a schema (Du),

(Du) $(u \in *p*) \rightarrow (u$ is true iff $p)$, where 'u' is an utterance, the first 'p' quoted by star marks is a sentence type, and the second 'p' is a sentence,

from the disquotational schema of sentence types. He then correctly claims that what is put in place of 'p' is given a uniform interpretation.

[15] Please cf. the second section of Chapter Five.

language, the concept of truth is fully characterised by a disquotational theory. Even Horwich's theory of truth has to appeal to an interpretation of a set of sentence types of a language to interlink which sentence type with which proposition. So it seems that minimalism and disquotationalism have their distinctive and different virtues with respect to how they treat the relation between sentence types and propositions expressed by them. Thus one can only say that, if Horwich's point of view in Horwich 1998*b* holds, the advantage of taking propositions as primary truth bearers is its reconciliation with the expressibility of infinite conjunctions or disjunctions in terms of the truth predicate. However, it will be shown in Chapter Five that this is incorrect.

A peculiar feature of minimalism follows from taking propositions as the truth bearers. Though minimalism is inspired by Tarski's truth paradigm, there is no such a hierarchy of languages similar to the case of disquotationalism, that is constituted at each stage by the target languages and the truth languages. Since every instance of the equivalence schema is also a proposition, each instance can also take the place of '*p*' in (E) and form a further axiom of the minimal theory of truth. This process goes on and on and on. This actually makes sense, for there are all kinds of natural languages but no language of propositions. Tarski's treatment of the liar paradox, however, is not available to Horwich. Horwich does not provide any solution to the Liar in *Truth* (1998*b*). But he certainly does not accept Tarski's solution, for this "smacks of overkill" (p. 41). Many instances of '*p*' in (E) are with the truth predicate. For instance, every instance of (E) can replace '*p*' and form another axiom of the minimal theory. This actually creates a pretty serious problem for him.[16] Discussion of this will be reserved for the next chapter.

---

[16] It causes a tension between Horwich's solution to the generalisation problem criticised by many others, for instance, Anil Gupta in "Minimalism" and "A Critique of Deflationism", and his solution to the Liar.

Although minimalism does not depend for its characterisation and analysis of the concept of truth on a tenable theory of proposition; it is not compatible with just any theory of proposition. For minimalism to be philosophically interesting, Horwich has to successfully provide a theory of proposition that is without any involvement of the concept of truth whatsoever. Or else one might incline to prefer disquotationalism to minimalism.

Horwich does have a response to this. He endorses the idea of *meaning as use*. Among others, Horwich counts Quine among the advocates of the use theory of meaning, though reluctantly so. The predominant difference between their views is that they reach entirely opposed conclusions. Based on his analysis of the informants' dispositions to linguistic behaviours under certain occasions, Quine establishes the thesis of indeterminacy of translation/interpretation according to which to seek a theory of proposition is simply out of the question. On the contrary, Horwich holds a use-theoretic account of word meaning and then argues that a proposition expressed by a sentence can be analysed purely in terms of the meanings of its sentential constituents and the way of combining them. It is not just that the use-theoretic account is independent of the concept of truth, but the analysis of compositionality—that is, the ways of combining the meanings of the sentential constituents—is also free from any complication caused by the concept of truth. More, Horwich claims that his account of compositionality is deflationary. It is compatible with *any* theory of meaning. His use-theoretic account of word meaning and his account of compositionality will be analysed and criticised respectively in the third and fourth sections. But first let us come to see, in the following section, how Horwich

can quell Quine's qualms.

2. Quelling Quine's Qualms or Quelled by Quine's Qualms?

In order to clarify Horwich's interpretation of the thesis of indeterminacy of translation, it is necessary to make plain the central idea of his use-theoretic account of word meaning. The account, according to Horwich, is inspired by Wittgenstein's thought that the meaning of a word is constituted by its use, namely, from the regularity governing the deployment, or, in other words, its overall use, in terms of the sentences in which it appears (Horwich 1998*a*, p. 3 and p. 6).[17] Put it in another way, the meaning constituting property of a word stems from the acceptance property *A(x)* that gives the circumstances in which certain specified sentences containing it are accepted (ibid. p.45).[18] Further, it is not just that any regularity governing our acceptance of all the sentences containing a given word will do: it also has to be explanatorily basic.

To surmount Quine's barrier, Horwich first distinguishes five different forms of sceptical views concerning meaning, namely (Horwich's caps), MEANING HOLISM, MEANING INDETERMINACY, MEANING INTERDEPENDENCE, MEANING DISTRIBUTION, and INDETERMINACY OF EVERYDAY PROPERTY.[19] MEANING HOLISM does not constitute a major part of Quine's thesis, and the last view mentioned is not the

---

[17] It is not clear what 'overall use' means. In criticising what he calls MEANING HOLISM, Horwich characterises this thesis as saying that, "the meaning of a term depends on *every single aspect of its overall use*" (Horwich 1998*a*, p. 61; my *italics*). Later in his summary of this position, he says, "the meaning of a word depends on its *overall* use" (ibid., p. 65). There seems to be a difference between 'every single aspect of its overall use' and 'its overall use'. The former seems to be suggesting that every speaker's mood, intention, or personal history is included in the elaboration of a given word's meaning. Whereas the latter does not seem very different from Horwich's own formulation of his use-theoretic account, due to the equivocality of 'depend on'. However, we should overlook this now and return to it later.

[18] There are many distinct features of this account, such as non-revisionism. But that is not in need of clarification in this chapter.

[19] The name of the last view does not come from Horwich. He does not give this view a name.

indeterminacy Quine has in mind. So what concerns me here is the second, third, and fourth.[20]

MEANING INDETERMINACY, according to Horwich's formulation, is the following thesis:

> [T]here may be no objective fact of the matter as to whether a given property of a word is, or is not, part of what constitutes the meaning of that word. For example, there is no objective fact as to whether a given acceptance property is explanatorily basic. (Horwich 1998*a*, p. 61)

The thesis of MEANING INTERDEPENDENCE can best be explicated by giving an example of Horwich—for a given pair of words, α and β, what constitutes the meanings of those words is a non-conjunctive relational fact, $R\alpha\beta$, according to which the meaning-constituting properties of α and β are $(\exists y)Rxy$ and $(\exists x)Rxy$ respectively (ibid.). Apparently, α or β having its meaning, depends on the other to mean what it means as characterised above. This is especially important to Horwich's account of word meaning, for it is vital for a word to have a certain meaning-constituting property, that other words in the sentences containing it have their own meanings.

MEANING DISTRIBUTION on the other hand is most essential to my discussion of Horwich's QQQ account because in terms of his interpretation, Quine's qualms mainly result from this thesis. It is useful to explain MEANING DISTRIBUTION by means of

---

[20] Although the last view will not be discussed hereafter, I would like to point out one minor inadvertency of Horwich's account. INDETERMINACY OF EVERYDAY PROPERTY, according to him, is a breeding ground for philosophical paradoxes. This is because certain predicates are *normally* used within some restricted domain. "[I]f it were extended beyond that domain, it would conflict with other use regularities. Philosophical paradoxes often originate in this way." (Horwich 1998*a*, p. 64) In a footnote of this point, Horwich raises the Liar as an instance—that is, our inclination to accept every instance of the equivalence schema occurs in *normal* contexts that preclude propositions involving the concept of truth. But according to the minimal theory of truth every instance of the equivalence schema can be substituted into the schema to form a further axiom. So the problem of the Liar could not possibly be about this point. This is probably just a slip anyway.

the above instance—namely, the meaning-constituting relational fact, *Rxy*, is symmetric

so that *Rxy = Ryx*. In the present case, *α* and *β* only have a joint meaning-constituting

property. It follows that some pairs or clusters of words might not have a

meaning-constituting property of their own (ibid. p. 62). Horwich then claims that, it

seems to him, this is what Quine argues in Chapter Two of *Word and Object* and

"Ontological Relativity". He further concludes that there is another alternative drawn by

MEANING DISTRIBUTION according to which the idea of word meaning is not

undermined by Quine's arguments. This is, however, really a surprising reading of *Word*

and Object chapter two.

In Chapter Nine of *Meaning*, "Quelling Quine's Qualms", Horwich proceeds with his

criticism on Quine's sceptical view regarding meaning. According to his interpretation,

Quine argues that translating a remote language is not to rely our evaluation of competing

translation manuals on their equating terms of a translated language with those of

translating languages on the basis of their having the same meaning. Rather, it is the

capacity of various manuals to preserve assertability that distinguishes adequate manuals

from inadequate ones. In other words, it is the pragmatic consideration that makes the

demarcation. And there are no grounds to tell which is correct and which is not, as far as

they all conform to pragmatic considerations. Horwich extends his formulation of Quine's

analysis a bit further. First, according to the pragmatic point of translation, any translation

manual is adequate if it preserves all the informants' dispositions to assent and dissent.

Second, based on the first point, it can be argued that there are many non-equivalent but

nevertheless adequate translation manuals by means of which the same term of the

translated language is mapped onto different non-synonymous terms of the translating

language. Third, Quine thereby rests his arguments on the conclusion that there are no such things as meaning entities (pp. 197-198).

In order to argue against Quine's qualms, Horwich starts with the inference from the second step to the third. In what sense, that is, can we say that there are no meaning entities by showing that there are multiple non-equivalent but adequate translation manuals? For the sake of carrying out his criticism, Horwich believes that he has successfully reformulated Quine's most important argument to include the following three premises (cf. Horwich 1998*a*, pp. 198-199).

(I) Two manuals with regard to their respective translations between a remote language *L* and English are both adequate but are incompatible with each other in the sense that according to one of them the translation of a foreign word *v* is '*e*' and according to the other it is '*e**'. '*e*' and '*e**' are not co-referential for English speakers.

(II) If a word has a meaning then an adequate translation of it must have the same meaning.

(III) If two English (context-insensitive) referring expressions have the same meaning then speakers of English regard them as co-referential.

According to these premises, it is easy to derive the conclusion that no words have a meaning. For, in terms of (I) and (II), '*e*' and '*e**' have the same meaning and, combining this with (III), '*e*' and '*e**' should be co-referential for speakers of English. But according to (I), '*e*' and '*e**' are not co-referential. Therefore *v* does not have a meaning.

Horwich's criticism has two major parts, against (I) and (II) respectively. He first argues that, even if (I) is granted, (II) is untenable. For the pragmatic consideration of *adequate*

translation manuals only ensures that each one of them is *useful* or *justified*. Although the

amount of adequate manuals could be many, only one of them is true and thus preserves

meaning (Horwich 1998*a*, p. 199). Modifying (II) into (II'), that if a word has a referent,

then an adequate translation of it must have the same referent, nevertheless does not help.

Horwich has a much stronger standpoint than I do on the disquotational schema of

reference. For him, that the translating terms, '*e*' and '*e**', refer disquotationally to *e* and

e* respectively is *perfectly determinate*. It does not follow that *v* can be regarded as

referring to both *e* and *e** because they are not co-referential. So (II') has to be rejected

too.

Horwich then turns to criticise the first premise. In terms of proxy function, which I

discussed in the third chapter, he thinks that Quine provides us unlimited instances of

translations of words. Let *f* be a 1-1 function that takes each physical object into its

'cosmic complement'. For each name '*a*', there is a corresponding singular term '*a**' such

that

$a^* = f(a),$

and for each primitive predicate '*G*', an additional predicate '*G**' such that

$\{x\}\{G^*[f(x)] \leftrightarrow G(x)\},$

and similarly for relational predicates. It follows that, for each atomic sentence '*G(a)*',

$G(a) \leftrightarrow G^*(a^*)$. And any extensional sentence '*p*' is thereby equivalent to the sentence

'*p**'.[21] For speakers of the same language, '*p*' and '*p**' are co-assertable. Suppose that

$D(x_1, x_2...)$ is the totality of assertability facts for a speaker's dispositions to assent and

dissent, $D(p_1, p_2...)$ is equivalent to $D(p_1^*, p_2^*...)$. With an additional assumption (A):

---

[21] I follow Horwich's formulation.

(A) Any manual of translation that preserves assertability is adequate,

one can suppose that, for a language H, which has sentences h_1, h_2…and so on, and a

language J, which has sentences j_1, j_2…and so on, a manual translating from H to J is

adequate according to (A) as far as it maps h_1 onto j_1, h_2 onto j_2, …, and so on. If (A) is

correct, then (I) can be derived from (A) and the consideration regarding proxy function.

In other words, given any manual M and a suitable proxy function, one can always

produce another manual M' that is as adequate as M but gives referentially different

words as the translations of words in the translated language.

(A) is untenable, according to Horwich. Assertability preservation is too narrow a

requirement for *adequate* translation manuals. If the pragmatic value of a translation

manual M is to help us communicate smoothly with our foreign informant in the way that

we substitute what we would have said with the fellow speakers of our language with the

corresponding foreign utterances in terms of M, then this manual does not assure us that

we will always produce correct predictions of what the informant will say under certain

circumstances. More importantly, we cannot predict the informant's *inferential*

behaviours.[22] Therefore, the collapse of premise (I) is the consequence of the rejection of

(A). Horwich then continues to construct what he thinks is a more reasonable account of

translation, based on his use-theoretic account of word meaning.

The first impression of Horwich's account of Quine's qualms, is the qualm that these are

really not Quine's qualms. Horwich's reformulation is actually *not* a reformulation of

Quine's argument. This reformulation, or more accurately this formulation, comes from

---

[22] Please cf. Horwich 1998*a*, pp. 202-206.

two claims of Quine—namely, reference is inscrutable, and adequate translation manuals are compatible with the totality of the informants' linguistic dispositions of assent and dissent. But these two claims do not jointly together form *one* argument. Further, to reformulate the alleged argument of Quine residing in *Word and Object* in terms of the consideration regarding proxy function is not chronologically correct. For the idea of proxy function does not even appear in *Word and Object*. In addition, Horwich does not seem to understand the idea of proxy function properly. Thereby his criticism of Quine's failure to predict the foreign informant's inferential behaviours is misleading. Moreover, it is dubious that Quine only takes assertability preservation as the one and only qualification of adequate translation manuals. Let us consider these points one by one.

Horwich's formulation is confused. Quine raises the *reductio* argument with regard to the inscrutability of reference at the beginning of section 12, Synonymy of Terms, of the second chapter of *Word and Object*. To make this argument plain I need to first introduce the definition of stimulus meaning. The definition of stimulus meaning comes from Quine's speculation on the learning process of a home language and the radical translation of a remote language—the stimulus meaning of a given utterance is the ordered pair of *affirmative* and *negative* stimulus meanings.[23] The affirmative stimulus meaning of a given sentence is the class of all stimulations, for a given speaker, which *would* prompt his assent to the sentence, and the negative stimulus meaning is the class of all the stimulations that *would* prompt his dissent from it.[24] So a translated sentence and a

---

[23] Horwich's use-theoretic account of word meaning heavily relies on the idea that to identify the meaning-constituting property of a word, is to find the acceptance property *A(x)* according to which *the circumstances in which certain specified sentences containing the word are accepted* are given. I assume that he does not have a strong opinion against the concept of stimulus meaning. At least it is not a serious problem in his formulation of Quine's argument.

[24] The word 'would' is used to enable the definition to include all possible verbal dispositions of the

translating one are stimulus synonymous if and only if they have the same stimulus

meaning.[25] The *reductio* argument is intended to show that, even if the stimulus

synonymous relation between the informant's one-word sentence "Gavagai" and our

translating one-word sentence "Rabbit" is granted, 'gavagai' and 'rabbit' as terms are not

thereby *intuitively* synonymous.[26] The argument runs as follows:

1. Stimulus synonymy between one-word sentences can guarantee the intuitive

   synonymous relation of the corresponding two terms. (the *reductio*)

2. If "Gavagai" and "Rabbit" are stimulus synonymous, then 'gavagai' and 'rabbit' are

   intuitively synonymous.

3. If 'gavagai' and 'rabbit' are intuitively synonymous, then they are co-referential.

4. It is possible to maintain the stimulus synonymous relation of "Gavagai" and "Rabbit"

   while the extension of 'gavagai' is varied from the referent of 'rabbit'.

5. Therefore, stimulus synonymy between one-word sentences cannot guarantee the

   intuitive synonymous relation of the corresponding terms.

Let me elaborate this point a bit further. Quine eventually arrives at a definition of

stimulus synonymy of terms, in terms of the definition of stimulus analytic sentences. The

definition of stimulus analytic sentences is as follows: sentences are int*ra*subjective or

int*er*subjective stimulus analytic respectively if and only if a speaker or all speakers of a

speech community would assent to it respectively, when given any stimulation.[27] Thus

---

speaker's assenting to, or dissenting from, the given sentence. There is, of course, a flavour of
counterfactuals in this explication and Quine is famous on denying that counterfactual conditionals have
any philosophical significance. However, this does not deprive him of legitimately employing such a
grammatical usage of the *subjunctive*. Quine makes it clear in the following passage: "its use [That is, the
use of 'would'] here is no worse than its use when we explains '*x* is soluble in water' as meaning that *x*
would dissolve if it were in water" (Quine 1960, p. 33).

[25] This is not quite right. Some socially widespread collateral information might cause some sort of
problem. Since this does not affect my elaboration, it is dropped for the sake of argumentation.

[26] It is not obvious how to count "Gavagai" as a one-word sentence. But, again, for the argument sake, this
issue is temporarily dropped.

[27] For Quine, stimulus analyticity comes in degrees. For a speech community within which all members

two general terms '*F*' and '*G*' are stimulus synonymous for a speaker or for speakers of the same speech community respectively if and only if "All *F*s are *G*s, and vice versa" is stimulus analytic to her or to them respectively. This also extends to cover singular terms—two singular terms '*a*' and '*b*' are stimulus synonymous for a speaker or for speakers of the same speech community respectively if and only if "*a* = *b*" is stimulus analytic to her or to them respectively. However, the concepts of stimulus synonymy of sentences, stimulus synonymy of terms, and stimulus analyticity, are not the candidates for reconstructing their intuitive counterparts (Quine 1960, p. 66). It is clear, then, that premise (II) of Horwich's formulation is an incorrect interpretation of Quine's theory. Quine never rests the adequacy of a translation manual on conferring the same meaning of the translating word to the translated one. They are only held to be stimulus synonymous.

It is noteworthy that what is crucial in the above argument is not the inscrutability of the reference regarding the *translating* words in different manuals. Instead, it is the referent of the *translated* word 'gavagai' that is inscrutable in terms of the stimulus meaning of "Gavagai". The stimulus synonymous relation between 'gavagai' and 'rabbit' as terms also does not pin down the reference of the former term. Further, the reference of 'gavagai' depends on the imagined field linguist's analytic hypothesis, by which the translation business of her is guided in the very first place. Part of the reason for introducing the argument of the inscrutability of reference is to be a preface of the discussion of analytic hypotheses. In discussing the phenomenon of the inscrutability of reference, Quine urges us that the stimulus meaning of "Gavagai" does not identify the

---

have certain specialised knowledge, there is a set of sentences that are stimulus analytic to those members but not to those that are not a member of that community.

reference of 'gavagai' as rabbits, undetached rabbit parts, or stages of rabbits. The reason is that what a gavagai is, involves quantitative but not qualitative identity. And quantitative identity is not just about the specification of the identity predicate, but it is also about the specification of the whole syntactic device, such as the singular and the plural, the copula, and the articles and pronouns, and so on. It is only possible to identify 'gavagai' as referring to a rabbit rather than undetached rabbit parts or stages of a rabbit, when the translation of certain syntactic device is determined. For instance, if the informant's sentence regarding identity is translated as "Is this the same rabbit as that?" then 'gavagai' is more likely to be translated as referring to rabbits.[28] On the other hand, 'gavagai' can also be translated as referring to undetached rabbit parts. For instance, we translate the same informant's sentence as "Are these the same undetached rabbit parts as those?" But these considerations have nothing to do with the intuitive meaning(s) of 'gavagai' and 'rabbit'. Of course, translating 'gavagai' as referring to rabbits, undetached rabbit parts, or rabbit stages is not entirely indeterminate in terms of empirical inquiry. For instance, if our linguist queries the indigenous informant on whether or not the gavagai is the same one of yesterday, in a sentence of the translated language of course, and if our informant's disposition is to assent to it, then at least our field linguist can know for sure that 'gavagai' does not refer to rabbit stages. For rabbit stages have their unique position in a space-time continuum. The point is rather that asking questions regarding reference can only make sense after the related syntactic device is determined. Again, premise (II) or the modified version of the argument is not Quine's.

There is one more thing I would like to say about the second premise of Horwich's

---

[28] The underlinings indicate the related syntactic device. The same applies to the second sentence with underlined words.

formulation. To say that, based on the thesis of the indeterminacy of translation, Quine attempts to refuse the abstract entities of meaning is a bit overstated. What Quine offered us from "Two Dogmas of Empiricism" through to *Word and Object*, is the sceptical view that there is no methodologically and epistemologically consistent way to account for the concept of meaning and the existence of meaning entities. If there is no consistent way to provide any tenable account regarding meaning, then this concept is the least thing we should care about while doing empirical semantics, much less the existence of meaning entities. In other words, to argue against meaning entities metaphysically is not what concerns Quine.

As indicated three paragraphs back, the idea of a proxy function does not appear in *Word and Object*. The argument of the inscrutability of reference is later promoted by many commentators as one of the most important arguments supporting the thesis of translation indeterminacy. But the point made originally in the second chapter of *Word and Object* is to be the curtain raiser of the considerations of analytic hypotheses. In order to tear down the alleged argument of Quine, however, it is essential for Horwich to formulate it in terms of the idea of a proxy function. According to Horwich's interpretation of this idea, a suitable proxy function that preserves truths of the original theory/language maps each name, each singular term, each general term, and all kinds of predicates onto their corresponding counterparts of the background theory/language. He criticises the idea of a proxy function, as the basis of the indeterminacy of translation, by saying,

> Thus, in appropriate circumstances I think and say "That is a rabbit"; but I will not think "The cosmic complement of that is the cosmic complement of a rabbit"—even though I might be disposed to work it out and assent it if queried. Consequently, if we use such facts about *actual* usage, together with a translation scheme, to generate expectations about what foreigners will say under similar circumstances, then only one of the various

But this is a fairly strange interpretation of the idea of a proxy function. In *Word and Object*, a given translation manual mapping 'gavagai' onto 'rabbit' does not base this mapping on their being stimulus synonymous as terms. It is actually because we premeditate to translate the syntactic device of the remote language in a corresponding way. If we did this differently, then 'gavagai' might be mapped onto, say, 'undetached rabbit parts'. Proxy functions with regard to relativisation of ontology are another matter, a highly related one perhaps. It is nonsense to say that 'rabbit' refers to rabbits in an absolute sense. It can only be asked, relative to a background theory *T*, to what does the term 'rabbit' of the object theory refer? Indeed, we will not think and say "*The cosmic complement of that* is the cosmic complement of a rabbit" if we think and say, in appropriate circumstances, "*That* is a rabbit". We will rather think and say "*That* is the cosmic complement of a rabbit (relative to a background theory *T*)." A suitable proxy function maps the universe of the object theory onto the universe of the background theory. Although the background theory contains new predicates of each of the corresponding old ones, the proxy function at hand is not designed for giving a mapping of them.

According to the above characterisation, even the stimulus synonymous relation of 'gavagai' and 'rabbit' as terms, for a bilingual or a bilingual community, does not fix their reference, for the stimulus meaning does not fix the reference of the given term or terms. Remember that the definition of stimulus synonymy of terms is: "All *F*s are *G*s, and vice versa" is stimulus analytic, "'All gavagais are stages of rabbits, and vice versa' is stimulus analytic" is as good as "'All gavagais are rabbits, and vice versa' is stimulus analytic" for

the sake of defining their stimulus synonymy.

Then, why does Horwich choose to criticise the idea of proxy functions as his major launching pad? Why does he interpret this idea in this way? My best guess is that it is to do with his theory of reference. Analogously to my interpretation of Quine's theory of reference in Chapter Three according to which the disquotational schema of reference is indispensable to Quine's theory, the disquotational schema of reference is essential to Horwich's theory too. The difference is that Horwich considers that the disquotational schema of reference and all its instances provide a genuine reference scheme. More precisely, his theory of reference is that nothing more lies behind the trivial instances of the disquotational schema of reference, (R). The disquotational schema of reference (R) is:

(R) Tokens of *n* refer, if at all, to *n*,

    where the star quotation marks indicate that the expression-type in quote is

    de-ambiguised to avoid the multiple referential relation of ambiguous

    expression-types (Horwich 1998*a*, p. 119).

He further derives a conditional schema (R_T) for foreign terms:

(R_T) *n* is the correct translation of $w \rightarrow (x)(w$ refers to x iff $x = n$).[29]

The reason Horwich believes that there is one and only one correct translation manual is revealed in the last sentence of our quotation two pages back, namely:

> Consequently, if we use such facts about *actual* usage, together with a translation scheme,
> to generate expectations about what foreigners will say under similar circumstances, then
> only *one* of the various assertibility-preserving schemes will be right. (Horwich 1998*a*, p.
> 205 my *italics*)

---

[29] (R_T) is acceptable so long as we change the word 'correct' to 'adequate' and add "relative to a manual *M*" after both the first appearance of '*w*' and the last appearance of '*n*'.

In other words, if presumably a list of all possible translation manuals in terms of suitable proxy functions is given, only one of them is correct because it *genuinely* matches what foreigners have in mind. But the idea of a proxy function is also applicable to home languages—that is, foreigners can also permute the universe of their language onto, say, itself. Considering his interpretation of the idea of a proxy function, Horwich seems to pick the wrong target. He should criticise the thesis of ontological relativity instead of the thesis of translation indeterminacy. For holding true both the disquotational schema of reference, in an absolute sense, and the thesis of ontological relativity is inconsistent.

Horwich has not finished his quelling of Quine's qualms yet. He further claims that the preservation of assertability is not a suitable criterion of the adequacy of translation manuals. Some, but of course not the *real* correct one, will fail to predict the foreigner's *inferential* behaviour. Regarding the case addressed in our quoted words, Horwich continues to say,

> For we infer h^* [That is, "The cosmic complement of that is the cosmic complement of a rabbit."] from h [That is, "That is a rabbit."], not the other way round. We first observe (or in some other way discover) that "That is a rabbit" is true, and then, given the definition of h^* in terms of h, we reason that "The cosmic complement of *that* is the cosmic complement of a rabbit" must also be true. Moreover, the fact that we reason in this way is a behavioural fact about us—a fact about our behaviour with respect to h and h^*. Consequently a fully adequate translation manual should preserve such behavioural relations, and not merely dispositions to assent and dissent. (Horwich 1998*a*, p. 206)

This is the strongest reason why Horwich conceives of being assertability preserving as not sufficient to identify the one and only *adequate* translation manual. I suspect this is the reason why, according to his classification, he attributes Quine's qualms to MEANING DISTRIBUTION. If, in terms of a suitable proxy function, 'rabbit' under a particular interpretation I_1 of a language L can be mapped onto 'the cosmic complement

of a rabbit' under another interpretation I_2 of the same language, and 'the cosmic complement of a rabbit' under I_1 can be mapped onto 'rabbit' under I_2, then there is a phenomenon of MEANING DISTRIBUTION between 'rabbit' and 'the cosmic complement of a rabbit' with respect to that proxy function. According to his use-theoretic account of word meaning, however, the predicate 'the cosmic complement of a rabbit' is a complex one whose meaning in turn has to be constructed in terms of the meaning of the simple predicate 'rabbit'. He therefore concludes that the phenomenon of MEANING DISTRIBUTION is nothing but an illusion. Interpreting Quine's qualms as MEANING DISTRIBUTION, however, is really an illusion. Considering the idea of a proxy function, Horwich seems to go backwards in constructing the view of MEANING DISTRIBUTION. In the light of his disquotational schema of reference, 'rabbit' refers to rabbit and 'the cosmic complement of a rabbit' refers to the cosmic complement of a rabbit. The meanings of 'rabbit' and 'the cosmic complement of a rabbit' are needed to make sure that these two instances of the disquotational schema come out right. Hence, it seems that their meanings also have a mapping relation similar to the one between the referent of 'rabbit' and the referent of 'the cosmic complement of a rabbit' in terms of a suitable proxy function. This, however, is not what Quine has in mind.

Still, I have to clarify Horwich's qualms over the adequacy criterion of translation manuals. Right after the cited words in the last paragraph, Horwich continues to argue,

> [A fully adequate translation manual] should translate h into a sentence of J from which J-speakers infer the sentence into which it translates h^*. That is, since we infer h^* from h, and adequate translation, T, should be such that J-speakers infer $T(h^*)$ from $T(h)$. But suppose that $T(h^*) = j^*$ and $T(h) = j$, whereas $T^*(h^*) = j$ and $T^*(h) = j^*$; and suppose that J-speakers infer j^* from j. Then T^* will not correctly predict the inferential relations, even though it may satisfy Quine's adequacy condition (A), preserving all the

assertibility facts. (Horwich 1998*a*, p. 206)

There are two problems of Horwich's argument. First, the intangible part of Horwich's argument is why the inference within our speculation regarding translation business also takes part in our informant's mind? When Quine claims that the reference of a term '*n*' of a remote language is inscrutable in translation, it is the translator who finds there to be no evidence to support her translation of the word '*n*' as referring to *m* or *m**. It is we who are doing the translating. We want to know which translation manual, *T* or *T**, is correct or adequate in the light of this sort of speculation. It is we who are puzzled, not our informant. There is no such a parallel of inferential behaviours between our informant and us while doing translation. This confusion, it seems to me, comes from Horwich's taking the disquotational schema of reference in an absolute sense. Without recognising this, his criticism is at best misleading, if not mistaken.

The second problem is that Horwich's argument shows that he gets the whole idea of the indeterminacy thesis wrong. In the case of radical translation, Quine shows that in translating a remote language, the field linguist tries to establish semantic facts of the translated language. However, Quine's thesis is that, even if we get the totality of assertibility conditions of our informants, the translation manual is not fixed. What we are not doing in translating a remote language is to translate its semantic theory. In terms of the linguistic practice of our informants, if there are semantic facts to be established, then we can form a general semantic theory. If the field linguist already has difficulties to fix the reference scheme of the translated language, the inference given by Horwich is to make the determination of translation more impossible. True, if the proxy function argument is valid, then reference is inscrutable and the ontology of a given language or a

given theory has to be relativised to be comprehended. Hence translation of another language is indeterminate. However, the indeterminacy thesis does not hinge its significance on the validity of the proxy function argument. The notion of proxy function is used to make a point *regarding* semantical reasoning. Therefore, to actually infer *h* from *h** or to infer *h** from *h* is irrelevant, is incidental.

In addition, it is hard to see, under a proper understanding of the idea of a proxy function, why we cannot do it the other way round and infer *h* from *h**? Indeed, in English 'rabbit' is semantically simpler than 'the cosmic complement of a rabbit', but why on earth do we think that we are authorised to confer this semantic feature to other languages? What sort of evidence do we have to support our assumption that it is an *adequate* characterisation of this semantic feature of a remote language? This is what Quine's qualms are about. In a normal vein, the field linguist might arrive at such a manual. But this is because of the *projection* by her, from something similar to our semantical considerations. In the vein of philosophical theorisation, this projection does not prove anything. It seems to me that what Quine has reservations about—that is, do certain semantic and synthetic features of our language coincide with other languages?—turns out to be a presupposition of Horwich's account.[30]

Moreover, does Quine really accept the claim that the only criterion of the adequacy of translation manuals is assertability preservation? There are two clarifications to make. First, translation of a remote language is an empirical enterprise. Our imagined field linguist has nothing more than normal induction and scientific methods to help her form a

---

[30] Please see the following two sections.

translation manual or manuals. She does not have any evidence other than the observed

occasions, under which indigenous utterances are produced, and the assent/dissent

dispositions of her informants to certain querying indigenous utterances. Horwich's

qualms over the licensing of correct expectations and inferential behaviours from

adequate translation manuals are only a matter of empirical inquiry. If our field linguist is

competent, she will make *right* queries. For our field linguist, the totality of her

informants' linguistic dispositions to assent and dissent, and the corresponding observed

occasions, are the sole depository of empirical evidence. And the translation manual she

forms is not only generated from her study of her informants' linguistic dispositions, but

also from her conjecture of how to segment the indigenous informants' utterances in

terms of her analytic hypotheses, by which the translation business is guided in the very

first place. Second, assertability preservation is not the only criterion for adequate

translation manuals. An adequate manual has to preserve truth. Although Quine does not

explore much of this point in *Word and Object*, he later agrees with Davidson that this is

necessary.[31] In addition, it is possible for some indigenous stimulus analytic sentences to

be translated into ours that are not stimulus analytic. The instance given by Quine is our

translation of a community-wide belief of "All rabbits are men reincarnate", which is not

stimulus analytic in English (ibid. p. 69). This consideration is certainly about truth

preserving. The process of translating consists, no doubt, in confusions and

misunderstanding, but this is the signature of empirical inquiry and is no more than what

we are caught by with respect to normal inductive method. There will always be conflicts,

if the inferences of the indigenous informants are misplayed. It is part of the field

linguist's task to sort these things out. With regard to the totality of her informants'

---

[31] This certainly causes a problem to Quine's theory of truth. I will discuss this in the final section of
Chapter Five.

linguistic dispositions, our field linguist does not have the luxury to hope for any evidence beyond this.

The common disturbance in translation, however, does not reside in the problem of preserving informants' inferences, since the misunderstanding of others' logical reasoning is easily located. The difficulty of translation involves ordinary beliefs, such as in the case of rabbits, men, and reincarnation. Consider folk Chinese religious beliefs as an example. The following are some popular Chinese superstitions: good gods always defend their disciples from attack by evil spirits, disease and disease-like symptoms are indications of demonic possessions, and good gods effect their power through their mortal mediums to repulse demonic attacks. In particular, superstitious Chinese believe that drinking a cup of water mixed with the ashes of Chinese incense sticks dedicated to good gods can cure diseases. It is clear that these beliefs put together explain why superstitious Chinese hold that it is true that ash-water has magical potency. To organise and formulate an inference from these superstitious beliefs in terms of her translations of them is not difficult for our field linguist. What is difficult is to first translate these beliefs themselves, because they are devoid of reasonable scientific justifications and because they are so different from what the field linguist believes. This example demonstrates how inadequate Horwich's recognition of the translation business is. For he thinks that the one and only *adequate* translation manual is always able to generate expectations regarding how the foreigner will respond. He also thinks that such a manual would license us "to have exactly the expectations we would have had if we were dealing with members of our own linguistic community—modulo the substitution of the corresponding foreign sentence for our own" (Horwich 1998*a*, p. 205). What will members of the English speaking community

respond to with respect to the translating sentences of the beliefs above characterised, if they do not hold them true? The expectations between how our informants will respond to a query and how our fellow speakers will respond to the translating sentence of the queried one are not always the same. An adequate translation manual, according to Horwich's elaboration, probably just amounts to a guidebook of daily conversations, and not even a very good one. How to keep the balance between holding our analytic hypotheses, conforming to the totality of linguistic dispositions of indigenous informants, and preserving all truths held by them is the major task of the field linguist. And it is for this reason that the translation business is pragmatic.

Horwich thinks that Quine's argument at best shows the indeterminacy of word meaning, but not sentence meaning (Horwich 1998*a*, p. 200). But it is apparent that this is not the case. The concept of stimulus synonymy is only an *ersatz* of its intuitive counterpart. Quine's thesis of indeterminacy of translation shows us that searching an empirical account of intuitive meaning of words and sentences is in vain. Horwich's criticism of Quine's qualms does not fail because of its inconsistency but its inaccuracy. Horwich did not quell Quine's qualms, but his account of meaning is very likely to be quelled by Quine's qualms.

3. On Horwich's Use-Theoretic Account of Word Meaning

Horwich's theory of meaning contains many aspects, such as the relations between his account of meaning and the concepts of truth, reference, and normativity, which will not be the main focus in the following two sections. In order for a full account of proposition to be given, Horwich provides his use-theoretic account of word meaning and then

develops his account of compositionality, which explains how the constituents of a sentence combine together. My major task in the two consecutive sections is to analyse them respectively.

Horwich's most important formulation of a given word's meaning is given by the following:

(M) The meaning of a word, w, is constituted by a use property, $u(x)$, according to which the overall use of w is governed/best explained by an explanatorily basic regularity, where an acceptance property $A(x)$ possessed by w gives the circumstances in which certain specified sentences containing it are accepted (cf. Horwich 1998a, p. 3, p. 6, p. 7, and p. 45).[32]

Many things within (M) are in need of clarification, but some will be set aside for the following section, such as the idea of a Fregean proposition. At issue here are, first, the relations between the use properties, regularities, and the acceptance properties. Second, what does it mean to say that one property is constituted by the other? And finally, how can we make sense of (M)? I will start with the second one.

From *Truth* to *Meaning*, Horwich rejects the idea of a reductive analysis of semantic notions such as, in the present case, "x means F" is reductively analysable in terms of a relation T such that $T(x, f)$. The reason for adopting a form of reductive analysis normally, is that it preserves logical form shared by both the reduced property and the one to which the former reduces (Horwich 1998a, p.22). In the above case of 'x means F', the common logical form is the relational schema Rxy. Horwich calls this the "Constitution Fallacy"

---

[32] For Horwich, the meaning of a predicate is a concept—that is, the meaning of 'red' is the concept RED. The illustration of this reasoning is left out because it is not relevant to my discussion below.

(ibid. p. 25). According to Horwich, it is always possible to *reduce* a superficial relational property to a monadic underlying property.[33] It is clear then, that what Horwich rejects is not a reduction from one property to another, but that this reduction has to preserve the logical form of the reduced property. Further, in the philosophical vein, to say that one property is reduced to the other implies that they are identical. But this is not what Horwich has in mind. Rather, that a property *P* constitutes anther property *Q* is what makes *Q reduce* to *P*. The instance given by Horwich is that being water *reduces* to being made of H_2O. That is, first, they apply to the same thing and, second, the facts about being water are best explained by that being water and being made of H_2O apply to the same thing.

As Schiffer indicates in "Horwich on Meaning" (Schiffer 2000), this is not quite right for the constitution relation between meaning properties and use properties.[34] For what has to be explained in such a constitution relation are not facts about a word's meaning property, but facts about the *use* of words that have that meaning property (p. 531). In other words, what has to be explained are facts about the acceptance of those sentences containing the word. Here I come to explicate the relations between the use properties, regularities, and the acceptance properties. Normally one will go for the property that is indispensable for the explanation of these facts to be the meaning-constituting property. That is, normally, the acceptance property $A(x)$ possessed by a word *w* should be what the meaning property of *w* reduces to. This is, however, not the case according to Horwich.

---

[33] The three *italic* words of 'reduce' within this paragraph indicate what Horwich has in mind when he talks of a reduction. However, as Schiffer argues, Horwich does not provide us an acceptable explanation of why the reduction should not happen the other way round.

[34] In discussing the instance raised in the last paragraph, Schiffer says that, for properties ϕ and ψ, ϕ constitutes ψ if (a) ϕ and ψ apply to the same things, and (b) facts about ϕ is explained by (a) (p. 530). Probably because Horwich uses 'reduce' and he uses 'constitute', he misplaces 'ϕ' for 'ψ' in (b).

He says:

> [F]or each word, *w*, there is a regularity of the form [:] All uses of *w* stem from its
> possession of acceptance property *A(x)*, where *A(x)* gives the circumstances in which
> certain specified sentences containing *w* are accepted. (Horwich 1998*a*, p.45)

It is clear that what constitutes meaning properties is not the acceptance properties but the use properties, according to which certain regularities are given in the form depicted in the quoted words. There is good reason for such a detour. It is possible for a word, *v*, to have the conjunctive (acceptance) property '*A(x)* and *B(x)*' so that we are compelled to wrongly admit the synonymous relation between *w* and *v* (ibid. p. 58, footnote 5). According to Schiffer, this is not exactly correct. It is not that *w* and *v* would be synonymous, but *v* would mean both *A(x)* and *B(x)* (Schiffer 2000, p.531). And this is not what we want to have. Moreover, Horwich claims that it follows from his theory that it is not possible to explain why a given meaning property reduces to the acceptance property to which it happens to reduce. Schiffer agrees with him on this point, because one cannot deduce that a word has the meaning property it does, from premises including that the word has the particular acceptance property that it does (ibid. p. 534). Otherwise, a word, *v*, with a conjunctive acceptance property is confused with another word, *w*, which has one of the conjuncts as its acceptance property.

After we have sorted these things out, we can move to the third issue, namely, how to make sense of (M). The first entangled part is the idea that a word's *overall* use can be governed by a regularity, explanatorily basic or not. Schiffer's example is that a soprano might warm up her vocal chords by singing 'osseous' ten times in succession (Schiffer 2000, p. 531). This use of 'osseous' is, and should, without doubt not be governed by a regularity according to which, say, the rest of the word's use is governed. As Schiffer

indicates, this is not a genuine difficulty for Horwich. When he talks about the overall use

of a word, Horwich is talking about the overall use of a word that is explained by the

acceptance of sentences containing the word. What, then, will count as a person's

acceptance of sentences that contain the word? According to Horwich, it is a person's

uttering the sentences containing a given word assertively to herself that counts as her

private acceptances of those sentences (Horwich 1998*a*, p. 94).

But this account does not really solve the problem. In a metaphysical sense, what

Horwich elaborates is that it is impossible for two words, *w* and *v*, to be exactly alike

except that they do not have the same meaning property. That the word *w* or *v* has a

meaning property, $m(x)$, is a consequence of its having a use property, $u(x)$. On the other

hand, it is also metaphysically impossible for *w* and *v* to be exactly alike except they have

different use properties. A word having its use property is the consequence of its having a

certain meaning property. In the light of this consideration, the meaning property of a

word is metaphysically indistinguishable to its use property. That is to say, a word *w* has

$m(x)$ if and only if it has $u(x)$.[35] However, in an epistemological sense, the regularity

governing the overall use of a given word is determined by its use in empirical inquiries,

not the other way round. For, if it is the other way round, then we should rather reduce

use properties to meaning properties. This amounts to saying that the overall use of a

word is *guided* by its having certain meaning property. In other words, the meaning

property of a given word *w*, acts like a norm that constrains the use of it. And this is what

---

[35] Based on this feature, Horwich reduces the meaning property of a given word to its use property. This is
with a shred of doubt though. As Schiffer indicates, metaphysical equivalence is a two-way street. But to
reduce one to the other presupposes that there is an asymmetry between them. Horwich does not offer us a
reasonable account with regard to this (Schiffer 2000, p. 535).

Horwich does not accept.[36] Therefore, the problem is how to account for alien uses? It will not do if they are dismissed as incorrect use by the regularity. For an alien use might as well be explained by some other regularity. If the use of a word is incorrect according to *a* regularity of its use, then alien uses of a word are correct in the sense that they are also governed by *a* regularity according to which most uses of that word are also explained. It is trivial to say that, given a regularity of the overall use a word, some uses are correct if they are governed by the regularity and some uses are not if they are not governed by it. The question we asked is rather to request a justification of the distinction between what is granted to go into the use fact of a word and what is not. This leads us to the elaboration of two further questions of (M) regarding the notions of explanatory basicness and best explanation. And my suspicion is that these two do not contribute to the full determination of use properties and thus meaning properties of words, as Horwich wishes them to do.

Horwich sometimes varies the content of (M) by saying that what is governed by the regularity is the *basic* use of a given word. In distinguishing his account from MEANING HOLISM, which embraces every aspect of a word's overall use, he claims that *not every use* of a given word is included in the construction of explanatorily basic use regularity. Some use facts of a word constitute *the use* of a given word and some other facts do not, and what makes the distinction is the difference between *explanatorily basic* use properties and the rest (Horwich 1998*a*, p. 60). He so explains,

> In other words, the way to pick out the particular use property of a word that comprises
> what we call "the use" is to find the use property that provides *the best explanation of all
> the others*. (ibid., my *italics*)

---

[36] Cf. the first, the second, the third, the eighth, and the tenth chapters of *Meaning*.

For the sake of clarity, 'the use' from now on is specified as 'the basic use'.[37] What

concerns his account of word meaning then, is to first construct *the explanatorily basic*

regularity of a given word's *basic* use. However, it is implied by Horwich's illustration

above that there are other regularities capable of accounting for the basic use of a given

word. The only difference between these regularities is the capacity of delivering the best

explanation of the not basic use of the word. His example of this use distinction invokes

the word 'planet'. Our disposition to accept "There are nine planets," so Horwich explains,

does not constitute *the use*—that is, the basic use—of 'planet', whereas our disposition to

accept "Planets orbit stars" is plausibly the constituent of *the use* (ibid. pp. 59-60). He

continues to elaborate this point by saying that, within the range of use of a given word,

there is indeed some sort of distinction, one that is similar to the analytic/synthetic

distinction. The fact that we sometimes acknowledge that there is a change of a word's

meaning is due to the fact that the word's explanatorily basic use property, by which the

basic use of it is explained, changes (ibid.).

This, however, strikes me as very implausible. The distinction between the basic use of

the word 'planet' and the other part of its use is a consequence of the scientific theory

held true at a particular time. What goes into the meaning of 'planet' is purely a

contingent fact of scientific inquiry. In terms of Ptolemaic astronomy, all stars and planets

orbit earth in one way or the other. It certainly follows that a different explanatorily basic

regularity governs the basic use of the word 'planet'. There is no unified way to

distinguish the basic use of each word from what is not basic. This is probably all right.

---

[37] Horwich's distinction between *the use*, or in other words, the basic use, and not basic use of a word is not
analogous to the intuitive distinction between 'literal use' and 'extended use'. Much extended use is either
based on the shape or pronunciations of a word, or is a deliberate misuse of it. In either case, the extended
use is not *explained* by the literal use.

The genuine problem is that this is really impractical.[38] The only three examples Horwich

offers to elaborate his account are not helpful—that is, 'and', 'red', *and* 'true'. The

acceptance property of the word 'and' is as follows:

(AND) The acceptance property that governs a speaker's overall use of "and" is (roughly)

his tendency to accept "*p* and *q*" if and only if he accepts both "*p*" and "*q*".

(Horwich 1998*a*, p. 45)

But can (AND) account for uses similar to the *italic* 'and' two sentences back? Well, it

can account for the *italic* 'and' three sentences back, but not always for other sentences

containing this word. (AND) is about the acceptance property of the logical connective

'and' in ordinary English. This connective, however, is also used as a connective to

connect two or more grammatical subjects or objects in ordinary English. I do not think

that every sentence containing the word 'and' used in this sense, such as "Romeo and

Juliet are lovers," or the last 'and' in (AND), can be *best explained* in terms of (AND). In

most cases, it just transfers a two or more places primitive predicate into a complex

compound of sentences.

This brings us to ponder what it means to say that the explanatorily basic use property is

about the basic use of a given word. What makes certain use the *basic* use of a word, *w*, is

certainly not in terms of whether or not the use property, $u_b(x)$, governing that certain use

is explanatorily basic.[39] This is utterly circular. For what makes $u_b(x)$ explanatorily basic

is its capacity to best explain the rest of *w*'s use. So $u_b(x)$ depends for its feature of

explanatory basicness, on a previous distinction between the basic and not basic use of *w*.

---

[38] For instance, who can give the distinction between what sorts of dispositions of accepting sentences containing the word 'table' are about the basic use of it and which are not? Can furniture designers or manufacturers?

[39] The tiny '*b*' in '$u_b(x)$' indicates that it is explanatorily basic with regard to the basic use of *w*. The same applies to the tiny '*b*' in the next paragraph.

Following the discussion so far, (M) is derived from three more fundamental principles, (Ub), (Mb), and (BP). They are:

(Ub)   The range of use of a word, w, is divided into two categories, basic and not basic, which in turn are decided by a non-intentional way, $sk(x)$[40];

(Mb)   The basic use of a word, w, is governed by an explanatorily basic use property, $u_b(x)$, according to which the basic use of w is governed by an explanatorily basic regularity, $r_b(x)$, where an acceptance property $A_b(x)$ possessed by w gives the circumstances in which some specified sentences containing it are accepted;

(BP)   If the not basic use of a word, w, are best explained by the explanatorily basic use property, $u_b(x)$, then (M).

(Mb) is least problematic among the three. (Ub), which is about the distinction between the basic and not basic use of a given word, and (BP), which is about the idea of best explanation, are in need of clarification.

While illustrating his idea of implicit understanding, Horwich borrows Putnam's idea of 'the division of linguistic labour' to explain why the regularities of experts, with different specialised knowledge, are the paradigm samples of the regularities of the use of all kinds of words. But there seems to be another reason for appealing to this idea. The distinction between the basic and not basic use of a given word also has to appeal to an expert who is specialised in the relevant field. For, in many cases, laymen can hardly tell the difference between the basic and not basic use of a word, say, 'post-modern'. Therefore, it seems reasonable to say that the distinction between a word's basic and not basic use is decided

---

[40] The words 'non-intentional way, $sk(x)$' will be further explicated later.

by an expert's specialised knowledge, *sk(x)*. Although Horwich does not consider this, it seems that it is a necessary component of his theory. These two roles played by experts in Horwich's use-theoretic account, will be analysed in the following.

As Schiffer argues, Putnam's idea does not fully explain why the regularities governing the overall use of a word given by different experts, would eventually boil down to the same thing (Schiffer 2000, p. 534). Horwich's response to this most general criticism is that what suits us is not every regularity but the *explanatorily basic* one. So some experts' regularities are either *best explained* in terms of the explanatorily basic one or dismissed in terms of the consideration within the specialised field. This explanation, however, generates two further questions. For Horwich, if there is no *objective* fact regarding which is more explanatorily basic than the other, then this is the case of what he calls MEANING INDETERMINACY. This indicates that Horwich does not intend to explain away this situation in terms of some canons of scientific method, such as simplicity, explanatory power, or conservativeness. This seems right, because these canons are used on a pragmatic basis to evaluate rival *theories* rather than the embodied regularities. The real worry comes from the idea of best explanation. As indicated above, there is no constraint for finding only one regularity being capable of explaining the basic part of a word's use. But only one of them can best explain the not basic use of the word, according to Horwich. And if there are two equally good regularities, it will be a case of MEANING INDETERMINACY. Thus, in what sense does 'best explanation' work as a criterion to somehow *fix* the regularity or, in some cases, regularities to be explanatorily basic in the above characterisation? It seems to me that which explanation is the best, depends on our explanatory purpose(s). One explanation probably has its value as making

the whole explanatory business simple, and the other may have its value as making it easy to comprehend. If what makes a regularity explanatorily basic results from a pragmatic consideration, then there will always be another regularity that is equally explanatorily basic in accordance with a different pragmatic consideration. If this view is correct, then for a word, *w*, there are *many* use properties that can explain, if not best explain, the overall use of *w*.

Even if Horwich's account could successfully pick out the explanatorily basic regularity or regularities by its or their capacity of best explaining the not basic use of a given word, it would have another problem. That is, the distinction between the basic and not basic use of a word is problematic. I have argued at length that this distinction itself is in need of clarification, and the best guess, it seems, is to invoke experts' specialised knowledge, *sk(x)*, to make the distinction. Horwich does not provide us a general account of what is to count as an expert. This is one of Schiffer's important criticisms of Horwich's account of word meaning (Schiffer 2000, p. 532). It is dubious for such an account anyway. It is not problem-proof, though, if such an account is given. For it seems very unlikely for this account to discriminate in favour of some experts but not the others. After all, "the division of linguist labour" is not "the division of linguistic legislator", prohibiting some specialistic use of a word in terms of some other experts' opinion. An astrological expert, for instance, might make a very different distinction between the basic and not basic use of the word 'planet' from an astronomical expert. "There are nine planets" seems to be a significant use fact regarding astrology. It certainly follows from Horwich's account that they have very different explanatorily basic regularities governing the overall use of the word 'planet'. And it is not an advisable line of argument, against the above criticism, to

claim that the astrological regularity governing the basic use of the word "planet" is not

able to best explain, or even explain the rest of the use of it, especially the use fact

regarding astronomy. For the same applies to the astronomical regularity. This is much

worse than the cases of MEANING INDETERMINACY. The claim that there is a set of

regularities, which combined together govern the overall use of a given word, seems to

make sense. It is very unlikely, even in terms of commonsense, that there is a single

regularity to be constructed to account for the *overall* use of a word. People probably

switch the use property in different contexts, such as the use property of the word

"planet" in an astrological context and the other one in an astronomical context.

How to make sense of (M)? The criticism boils down to the following. First, an account

of providing best explanation of something in terms of something else has to be given.

Without this, (M) is not justified because which regularity is capable of providing best

explanation to the not basic use of a word is very likely to depend on our explanatory

purpose(s). Second, a general account of what counts as an expert has to be given.

Without this account, invoking "the division of linguistic labour" is not complete. Third,

even if the second feature is fulfilled, Horwich's account is still not successful. For there

seems to be no way to reconcile different opinions from different specialised fields

regarding what the *basic* use of a given word is. If these questions are not answered, the

whole enterprise of Horwich's explanation of word meaning collapses. The reason is that,

following his reasoning, either there is no way to decide *which* meaning property and use

property are the meaning property and use property of a given word, or there seems to be

no way to decide to *which* use property the meaning property should reduce. Horwich's

account seems to create its own version of indeterminacy of meaning.

There is one last point I want to make regarding how to make sense of (M). Horwich's elaboration of what is an acceptance property is less than clear. According to (M), an acceptance property, $A(x)$, gives the *circumstances* in which certain specified sentences containing the word, w, are accepted. What does 'circumstances' mean? In his instance of the word "planet", "Planets orbit stars" is a use fact that constitutes the basic use of that word. But in what *circumstance* does one accept this sentence? It is highly unlikely that what it means is the necessary theoretical facts for accepting this sentence, for Horwich emphasises that the explanatorily basic regularity along with other necessary apparatus can best explain the overall use of a given word. So *circumstances* should be something other than theoretical facts. It is not clear, then, in what circumstance does one accept "Planets orbit stars." According to Quine's distinction between occasion and standing sentences, which depends on how close the sentences are with their stimulus meanings, most theoretical sentences are standing sentences and thus do not have particular *circumstances* for their acceptance.[41] Therefore, one of the genuine difficulties of Horwich's account is how to account for the acceptance properties of theoretical terms.

There are other more severe difficulties for Horwich's account on a more general basis. According to Horwich's account, the meaning of almost every word is *determinate* at a particular time period of the evolving stages of a given language, except the cases of MEANING INDETERMINACY. This is due to the fact that he appeals to experts' opinions with regard to the explanatorily basic use property of every word, and these experts exist only relative to a particular time period. However, there is always hidden

---

[41] The partial reason for Quine's constructing a technical definition of stimulus meaning is to specify the circumstances in which certain sentences are queried.

indeterminacy at the evolving stages of a given language, according to Ken Akiba (2002), as I indicated in Chapter Three.[42] A hidden indeterminacy by definition is not a manifest indeterminacy of a given language, such as vagueness, relative to a particular time period.[43] The hidden indeterminacy is also objective—that is, it is not decidable in terms of our awareness of it. The instance given by Akiba is the relational predicate, 'heavier than', relative to a pre-Newtonian language. There is a hidden indeterminacy of the use of 'heavier than' in a pre-Newtonian language—that is, it is indeterminate between 'more weighty than' and 'more massive than' from the post-Newtonian viewpoint (p. 225). The best reply I can come up with in terms of Horwich's account is that the use property and hence the meaning property of 'heavier than' in a pre-Newtonian language are determinate according to the expert or experts' opinion at that particular time period. Later in a post-Newtonian language, the use property and the meaning property of it simply changed. So the best way to answer this question is to appeal to a change of meaning as a result of empirical inquiry in general, or scientific inquiry in particular. My intention here is not to evaluate this objection, but to question the envisaged way of replying to it.

In criticising Horwich's account, Schiffer raises a very good objection. That is, if the ambition of Horwich's account is to explain the *public or communicative* linguistic behaviours of speakers of a particular language, then it will flagrantly fail (Schiffer 2000, p. 535). I would like to further elaborate a related point regrading Horwich's account of

---

[42] Akiba does not raise this question for Horwich's theory of meaning but for deflationist theories of truth and reference in general. But it is suitable to make the point here.

[43] Akiba emphasises that this has nothing thing to do with Quine's thesis of indeterminacy of translation because the case he raises is only concerning a single language. But for Quine indeterminacy even happens in the home language. Horwich agrees with Quine that the matter of meaning is at the end of the day about translation. What he does not agree with is that there is any indeterminacy with regard to this subject matter.

what he calls *implicit understanding*.

On the one hand, Horwich claims that, even if a sophisticated scientific inquiry is required in order to reach the conclusion that the word 'dog' has a certain explanatorily basic use property, $u(x)$, people still understand 'dog' perfectly well (Horwich 1998*a*, p. 16). On the other hand, he also claims that, when a member of a linguistic community has only a *minimal understanding* of 'dog', not a full or incomplete knowledge of it, the communal language meaning, namely, $u(x)$, may be correctly attributed to him (ibid., pp. 17-18). For popular words such as 'dog', 'red', and even 'arthritis', this might work. For sophisticated words such as 'supervene' in "*p* supervenes on *q* if and only if *p* metaphysically necessitates *q*," it is dubious. In addition, Horwich's explanation of the distinction between full, incomplete, and minimal understanding of a word is in terms of the degree of similarity between what it means in one's idiolect and what it means in the communal language. And this degree of similarity in turn is constituted by the degree of similarity between the explanatorily basic use property (ibid., p. 18). This is really puzzling. First of all, there does not seem to be any possibility of constructing a *general* technical criterion by which a speaker would be qualified as capable of understanding each word of the communal language. Secondly, it is difficult to see how the degree of similarity between the explanatorily basic use property of a word for a speaker, and the one for a community, has anything to do with the attribution of the communal language meaning of that word to the speaker. Here is an analogy. Suppose that someone bases his understanding of the rules of soccer on his understanding of the rules of Celtic football. How do we make sense of the attribution of the communal language meaning of the rules of playing soccer to him, in terms of the degree of similarity between his understanding of

Celtic football and his understanding of soccer? Further, since Horwich claims that words

of different languages have the same meaning as far as they have the same use property,

the degree of similarity of the use property is also applicable between a foreign word and

a word of mother tongue. Take Quine's famous 'gavagai' as an instance. English speakers

do not have a corresponding word for it, because it is accepted in certain specified

sentences containing it, when a particular kind of insect appears. A new English word, say,

'rabbit-fly', is probably invented for this reason. So English speakers have two use

properties regarding 'rabbit' and 'rabbit-fly'. Let us further suppose that there is no such a

distinction between rabbits and rabbit-flies in the remote language. How do we make

sense of the degree of similarity between an English speaker's use properties of both

words and our indigenous informant's use property of 'gavagai'?[44] Finally, if we want to

make sense of why people can have public communication, then the degree of

dissimilarity between speakers' use properties and the so-called communal use property

of every word should also be taken into account. In many cases, what is difficult to

explain is why and how people can *understand* each other while having different use

properties of many words.

An additional point about the change of meaning in public languages should be addressed.

An instance regarding two four-word Chinese idioms can illustrate this well—namely,

"Morning-three-evening-four" and "Morning-Chin-evening-Chu".[45] These are used

indifferently by most Chinese nowadays, but they actually had a very different meaning in

the old days. Let us use *p* and *q* respectively as shorthand. What '*p*' says concerns trickery.

---

[44] Having the synonymous relation between words of different languages depend on possessing the same use property is highly dubious. The discussion of this is reserved to the following section.
[45] It is hard to find the exact translations of them, so they are written in the fashion of combining the four corresponding English words with dashes.

It comes from a story. A person has three monkeys and he feeds them four chestnuts every morning and evening; due to his poverty he decides to reduce the amount of chestnuts; he tells the three monkeys that, from now on, they will only have three chestnuts for breakfast and four for dinner. The monkeys' protest of this decision forces him to try something different. He tells the three monkeys that, from now on, they will have four for breakfast and three for dinner. The monkeys happily accept this new offer, feeling that their bargain works because the amount of chestnuts for breakfast increases. So this story is saying that the three monkeys are tricked by their master because of their simple-mindedness. What '*q*' says is actually about people being inconstant in all kinds of relationships. It comes from a story with regard to a small country that is sandwiched between two powerful and opposing countries, Chin and Chu. In the morning the messenger of Chin was received in audience by the King of the small country, and the King acceded to Chin's proposal of alliance. Later in the evening the messenger of Chu was received in audience and the King changed his mind to accede to Chu's proposal of alliance. The original meaning of this idiom is about double-mindedness and it is extended to mean the inconstancy occurs in all kinds of relationships. The reason that '*p*' is used indifferently with '*q*' today is mainly because the first and the third words of them are the same—that is, 'morning' and 'evening'—and many people mistake the 'three' and 'four' in '*p*' as "with three or four women or men". It could be argued, however, that the change of meaning might be caused by many reasons. This is just one of the cases. But according to Horwich, we sometimes acknowledge that there is a change of a word's meaning that is due to the fact that *the word's explanatorily basic use property changes* (Horwich 1998*a*, pp.59-60). Contrary to the last point I made, the change of meaning in this particular instance does not result from the change of the similarity between the use

property of the mass majority and the use property of the experts. It rather results from the dissimilarity between them.

I end this section with a discussion regarding Horwich's objection to Kripke's sceptical paradox because Horwich's reply to the paradox seems to suggest the downfall of his own theory of meaning. The sceptical paradox applies to dispositionalism in general, insofar as it is claimed that the reference of linguistic expressions is not determined by our dispositions to linguist behaviours. One major instance given by Kripke is the indifference of the use fact of 'plus' of our linguistic community and 'quus' of an imaginary community. The divergence between the two, by definition, is about their extensions, when numbers are ungraspably huge, so huge that none of us can talk about them, and at this point the triples of numbers to which 'quus' applies are slightly different to that of 'plus'. Horwich's objection to this imaginary case is its question begging nature. He argues in the following way:

> For to insist, without argument, on there being such a possibility is simply to beg the question against the use theory of meaning. What does seem intuitively right is that there are possible *complex* expressions—definable in terms of 'plus'—whose extensions diverge from that of 'plus' in the slight way imagined. But no complex expression—since it will inevitably bear certain use relations to its constituents—can have exactly the same use as a primitive expression. Therefore, in order to construct his counter example to the use theory, Kripke must assume in addition that there could be a *primitive* term (on a par with 'plus') that is coextensive with one of those possible complex expressions, and that, if there were such a primitive, its use (including dispositions concerning the divergent cases) would be identical to our use of 'plus'. (Horwich 1998*a*, p. 223)

Alexander Miller correctly addresses a protest to this argument, that 'quus' is semantically complex but not syntactically complex, and hence has no constituents to which it might have use relations (Miller 2000). He continues to argue, using Goodman's

famous instance of 'bleen/grue', that even if Horwich's point is right, the case of

'blue/green' and 'bleen /grue' still applies (ibid. pp. 167-168). For, in terms of the

formulation of 'bleen' and 'grue', the four predicates are mutually definable in pairs.[46]

The word 'bleen' is not semantically more complex than 'blue' is.

However, Horwich's question begging argument does not seem to be exactly what Miller

thinks it is. After the words cited above, Horwich says:

> But it is plausible to suppose that a foreign word could acquire such an extension only
> either it were defined in terms of some word meaning the same as 'plus', or it were
> applied in some *definite* way to certain triples that, given our limitations, are beyond our
> definite range of application of 'plus'; and in neither of these cases would it be correct to
> say that the use of the foreign word is *exactly* like our use of 'plus'. (ibid.)

The line of Horwich's defence is similar to the one he adopts to argue against Quine's

proxy function in the case of indeterminacy of translation. The basic idea there is that we

infer 'the cosmic complement of a rabbit' from 'a rabbit' because the former is a complex

term bearing some use relations to the latter. The exact reason is of course unknown, but

Horwich somehow seems to confuse these two very different argumentations as being

one and the same. In his argument against proxy function, the line is clear that 'the

cosmic complement of a rabbit' is indeed semantically and synthetically complex. In the

case of 'plus' and 'quus', there is no reason whatsoever to think that 'quus' is a

syntactically complex expression. But except this slip, Horwich's point is different from

that made by Miller. Horwich seems to argue that 'quus' and 'bleen', be they syntactically

complex or not, are constructed from the *actual* predicates 'plus' and 'blue'. Although in

the case of 'blue/green' and 'bleen/grue' they are mutually definable, the pair of

---

[46] That is, X is grue iff x is green at or before a certain time t or blue after t.
X is bleen iff x is blue at or before a certain time t or green after t.
X is blue iff x is bleen at or before a certain time t or grue after t.
X is green iff x is grue at or before a certain time t or bleen after t.

'bleen/grue' is constructed from the pair of 'blue/green', not the other way round. In other words, unless we first construct 'bleen/grue' from the actual predicates 'blue/green', there is no way and no need to define the latter in terms of the former. Creating some eccentric predicates in the light of the actual ones does not confer the former the same use relations between the latter and the rest of our linguistic enterprise. The idea behind this defence, it seems, is that the eccentric predicates lack the interconnections of use facts of the complicated reticular structure of our linguistic enterprise within which the actual predicates, names, singular and general terms are nodes of the structure. However, by the same token, why should the same or very similar reticular structure of use relations of one language apply to another very different language? In other words, Horwich's entire use-theoretic account of meaning depends for its plausibility on some contingent use facts of a language, in the present case, English. Different languages share the total contingent use facts of English. This is hardly acceptable. It will be shown in the following section that this is more apparent in Horwich's account of compositionality.

4. Inflating Compositionality

According to Horwich, to understand one's own complex expressions, one only needs one's understanding of their parts and how they are combined. More specifically, as far as one has figured out *how* a sentence is constructed from primitive constituents and providing that one knows the meanings of those constituents, one *automatically* qualifies as understanding the sentence (Horwich 1998*a*, p. 155). Nothing else is needed for this sort of *implicit* understanding. This is the general idea of Horwich's view on compositionality. So in providing an account of word meaning and the knowledge of how words composed in complex expressions, the meanings of complex expressions are

revealed. Two things are noteworthy. First, since this understanding is implicit, *how* constituents of complex expressions combine, in terms of his principle of compositionality, does not involve process of any kind. Second, recollections of the meanings of words (complex expressions being themselves an inference or within one), the pragmatic contexts of complex expressions, and what the world is like, are not related in one way or another to one's automatic understanding of complex expressions. They are entirely irrelevant. Furthermore, Horwich claims that his principle of compositionality deserves to be called *deflationary*. By this he is saying that the principle of compositionality is compatible with *all* kinds of theories of word meaning. So the subject of this section boils down to the following several issues. First, what is Horwich's principle of compositionality and what is it deflationary about? Second, is this principle able to explain the use fact regarding complex expressions? Finally, does it make sense to say different languages somehow share the same set of synthetic features?

Horwich's principle of compositionality can be put as follows:

(PC) A synthetically complex expression has its meaning as determined by the

      meanings of its constituents and, together with, the manner they combine.

The first task is to explain what the mode of combination is. To use Horwich's favourite example "Dogs bark" and his capitalising way of specifying the meaning-constituting property to explain that the meaning of "Dogs bark", *x*, results from putting terms whose meanings are **DOG** and **BARK**, in that order, into a schema whose meaning is *NS V* (Horwich 1998*a*, p. 156).[47] Since the meaning-constituting properties of words reduce to

---

[47] For notational reasons, the meanings of schemata are not just capitalised but is also italicised, and the '**S**' on the other hand is not italicised in order to indicate the plural ending. This somehow foreshows the difficulty of his account.

their use properties, all capitalised words can also be replaced by use properties. The most important part of this formulation is certainly about the schema "*n*s *v*" and its meaning, or rather its *having* meaning.[48] The general idea behind this specification of the meaning of a complex expression is what Horwich calls the Fregean Principle. That is,

(FP) The meaning of the result of applying combinatorial procedure *P* to a sequence

of primitives = The result of applying combinatorial procedure *P* to a sequence

of primitives. (ibid. p. 180)

Horwich continues to explain (FP) in a footnote:

> Note that a combinatorial procedure, *P*, acts on any sequence of functions and objects…Therefore *P* can perfectly well act both on a sequence of words (including schemata) and on the sequence of their meanings. (ibid.)

This is really a peculiar explanatory strategy. And it is not Fregean. John Collins in "Horwich's Schemata Meet *Syntactic Structures*" (Collins 2003) objects that, normally, one does not assign semantic values to schemata because schemata are abstract descriptions of classes of phrases made from schematic symbols describing classes of words.[49] It is the phrases or words they describe have meanings. "[I]t is not the schemata that composes word meanings," Collins continues, "[but] it is the meanings themselves, as it were, that compose, axiom schemata merely describe this." (Collins 2003, p. 427) This is quite right. However, let me bypass this problem and continue my analysis.

Even if the above-mentioned problem is put aside, in what sense can one say that (PC) is *deflationary*? The most important claim regarding the deflationary nature of (PC) is of course its compatibility with all kinds of theories of word meaning. Given that a

---

[48] Cf. the last note. The same notation applies to "*n*s *v*".
[49] For instance, one does not say that the equivalence schema "The proposition *that p* is true if and only if *p*" is true but that its instances are.

Davidsonian truth conditional account is the major rival of Horwich's accounts of word meaning and the principle of compositionality, and a Davidsonian account falls short of an inflation of the theory of truth, according to Horwich, an account like Horwich's with no such inflation is hence deflationary. This seems to be the implicit reasoning of Horwich. However, this is only correct when a Davidsonian account defines compositionality in terms of satisfaction. If this is the case, then Horwich's account is in this sense deflationary. But it is not. On the contrary, that such a Davidsonian account does not employ schemata, much less conferring meanings to them, is a virtue of it.

Jerry Fodor and Ernie Lepore are two strident critics of the claim of the compatibility between the deflationary nature of (PC) and all kinds of theories of word meaning. Among various arguments they made together or separately, two are discussed in the following. The first one is about principle of uniformity:

(PU) The semantic values of phrasal expressions of syntactic category c are

    constituted in the same way as the semantic values of those of its lexical parts of

    category c.[50]

(PU) and (PC) jointly imply that the stereotype theory of word meaning is false. (PC) says that the meanings of complex expressions are constructed from their constituents. (PU) says that the meaning of any phrasal expression is constituted in the same way as its constituents are. Take a stereotype theory as an instance, the meaning of 'pet fish', say, according to (PU) is also a stereotype. The stereotype of 'pet fish', however, is not constituted by the stereotypes of 'pet' and 'fish'.[51] So if (PC) and (PU) both hold, any

---

[50] This formulation is Philip Robbins's in "What Compositionality Still Can Do" (p. 330).
[51] This is because the stereotype associated with the phrase 'pet fish' has some features that are not features of the stereotype associated with the word 'pet' or the stereotype associated with the word 'fish'. For instance, living in a home aquarium is not a feature of the stereotype associated with 'fish' or 'pet'.

stereotype theory is false. Therefore, (PC) is not compatible with all kinds of theories of word meaning. This is also Philip Robbins's strong deflationism position, according to which (PC) together with other widely-accepted semantic principles such as (PU) *do not* set constrains on the nature of word meanings.

Horwich's reply is easy to anticipate. He rejects (PU) outright. (PU) and (PC) are simply incompatible with each other in Horwich's account. In "Deflating Compositionality" (Horwich 2001), this is pretty much his response. Horwich hardly explains why (PU) is wrong, except to say that we should resist the impulse to accept (PU) (ibid., p.380). But this is not an answer at all, especially when (PU) seems to be intuitively correct. If Horwich is right, then it is difficult to explain why there is a concept of 'horse' but no concept of 'white horse'.

Horwich provides another reply: the meaning of the *phrase* 'pet fish' should be the result of substituting words meaning what 'pet' and 'fish' do—in this case, the stereotypes associated with them respectively—into a schema meaning what '*a n*' does; the phrase itself does not possess anything similar to what its constituents possess (Horwich 2001, p. 380). Horwich's argument is that, if one accepts that the meaning of a word is engendered by its associated stereotype (or inferential role, or recognitional capacity), one does not have to accept that the meaning of a complex is composed by the properties possessed by its constituents. He then continues to say that this would be an instance of the deflationary view of compositionality (ibid.). According to this, given either the stereotype theory or indeed any theory of word meaning and (PC), (PU) does not hold. I think Horwich is right on this one, though (PC) alone, it seems, trivially dismisses the

stereotype theory as unqualified. For the stereotype theorists simply do not accept the

idea that the stereotype associated with, say, 'pet fish' is the fusion of the stereotypes

associated with 'pet' and 'fish'.[52] And this is one of the important reasons for them to

adopt (PU). I seriously doubt that this can be explained the other way round, for the

incompatibility of (PC) and (PU) is straightforward. To form an argument in terms of the

assumption that one who holds (PC) also holds (PU) does not seem to be feasible. In

other words, the argument from Fodor and Lepore, and the one from Robbins's strong

deflationism too, seem to deviate from the right line of argumentation at the very

beginning. The difficulty, it seems, is about (FP). It might be less problematic for the

stereotype theorists to accept both (FP) and (PC) to be applicable to sentences, but with

respect to phrases this is entirely different. Accepting (FP) is equal to accepting that

schemata have meanings. And the meanings of schemata force the stereotypes of words

in a sequence of a phrase to fuse. So the issue here is not that (PC) together with (PU)

imply the falsity of the stereotype theories, but that the stereotype theories are not

compatible with (PC) in the cases of phrases.[53] In terms of this fact, the stereotype

theories probably should be rejected outright, for they cannot explain why one can

understand the composite *meaning* of the phrase 'pet fish' without any involvement of the

stereotype of pet fish. It follows, however, that (PC) has little to do with being

deflationary or not, because not every theory of word meanings is compatible with (PC).

The second argument is quite different from the first one. Fodor and Lepore continue to

dance to Horwich's tune. But this time the argument is based on the less controversial

---

[52] The problem of this line of argument is implicit in this very sentence. If one agrees with (PC), the meaning of 'pet fish' will not be a stereotype but a construction property resulting from applying the stereotypes associated with 'pet' and 'fish' into the schema '*a n*' whose meaning is "*A N*".
[53] Strangely enough, the *italic* 'phrase' excerpted from Horwich's words at the beginning of this paragraph is actually replaced by the word 'sentence' in Horwich's article.

part of (PC)'s applying to sentences. I will follow Horwich's formulation of this argument:

(1) To understand certain complex expressions indicates the fact that some other complex expressions built from the same elements are understood. For instance, "John snores"/"Flounders swim" and "John swims"/"Flounders snore". ('systematicity')

(2) Therefore, the meaning of a word does not depend on the complex expression in which it appears. ('context-independence')

(3) But the stereotype associated with the word 'swim' in "Flounders swim" is not the same as the stereotype associated with that word in "John swims."

(4) Therefore the meaning of 'swim' is not engendered by an associated stereotype.

(Horwich 2001, p. 381, footnote 15)

Horwich summarises his objection to this argument in three points. First, premise (3) has to go. The stereotype associated with the word 'swim' is not varied in every place it occurs. Second, even if this argument were persuasive it could tell against only the stereotype theory but not every theory of word meaning in general. Third, this argument does not depend on compositionality but on context-independence. If the consequence is to prevent any theory of word meaning from being compatible with (PC), that is because of context-independence rather than compositionality. Let me discuss these issues in a reverse order.

To the third point, it is hard to see how Horwich's reply can work. This is because context-independence seems to be the consequence of (PC). If a word's meaning depends on the contexts in which it appears, and if (PC) also holds, then the meaning of a syntactical complex expression and the meanings of its constituents are mutually

determined. This is circular and contrary to Horwich's own view. Therefore, context-independence is the consequence of (PC). The second point is a bit complicated. On the one hand, Fodor and Lepore argue that (PC) sets constraints on all kinds of theories of word meaning. On the other hand, this argument is concentrated on the incompatibility between (PC) and the stereotype theory in particular. It might be all right for Horwich to object to this argument in terms of its failure of proving that (PC) sets constraints on all kinds of theories of word meaning. But it is devastating for him to accept this argument even in a counterfactual sense, for this suggests that he concedes that (PC) is not deflationary. To the first point, this is the typical reply we expect. But it is only natural for Horwich to make the claim that the stereotype associated with the word 'swim' is not varied in every place it occurs, it is not for the stereotype theorists to make such a claim. As a matter of fact, Fodor and Lepore specifically explain what they mean by saying that the stereotype associated with the word 'swim' in "Flounders swim" is not the same as the stereotype associated with that word in "John swims". They say:

> [T]he stereotype of people-swimming is much different from the stereotype of flounder-swimming since the latter, but not the former, adverts to the exercise of fins. (Fodor and Lepore 2001, p. 365)

I cannot see how Horwich's reply is actually responding to this clarification.

This is probably not important anyway, if we consider Robbins's criticism of Horwich's claim that (PC) is deflationary as correct. According to Robbins, (PC) is not deflationary even if it is seen as claiming something weaker, something similar to what Robbins calls weak deflationism—that is, (PC) itself, not together with other semantic principles, is compatible with all kinds of theories of word meaning. According to Robbins, a *superstrong* holistic theory of word meaning, for instance, is not compatible with (PC).

By definition, superstrong holism is the view that the meaning of a word in a language is at least partly constituted by the meanings of other words and phrases in the language, including phrases of which the word is a constituent (Robbins 2001, p. 333). It is clear that superstrong holism is not compatible with Fodor and Lepore's 'context-independence'. Of course superstrong holism is not plausible, especially to Horwich. But this is beside the point. Can we weaken Horwich's claim by saying that (PC) is only compatible with *plausible* theories of word meaning? I am afraid not. This means that (PC) indeed sets constraints on theories of word meaning. So Horwich's (PC) is not deflationary.

It was already argued in the discussion of Fodor and Lepore's first argument that the stereotype theories of word meaning are perhaps too implausible themselves to be compatible with (PC). Superstrong holism, on the other hand, provides us an incontestable case that it is itself incompatible with (PC) according to Horwich's own view. Horwich is probably so enthusiastic to establish a *deflationary* semantic enterprise that he falls short of addressing Schiffer's comment that the words 'deflationary' and 'inflationary' are more emotive rhetoric than philosophically significant notions.

In addition, the two sample sentences in the argument two paragraphs back are with different schemata according to Horwich's account. It is '*n v*s' for "John swims" and it is '*n*s *v*' for "Flounders swim". It seems that both sides take this as a minor difference. But is it?

The sketchiness of Horwich's account of synthetic schemata makes it very difficult to

understand (PC). It is so sketchy that every construction property, resulting from applying

the meanings of the constituents of a sentence to a syntactic schema in a specific order, is

unique. It even cannot tell the difference between well-formed schemata, such as '*n v*s'

for "John swims," and ill-formed ones, such as '*v*s *n*' for "Swims John." To begin with,

this account fails to explain why there are structurally synonymous relations between

sentences. Horwich disposes of this sort of question in terms of the *fine-grained*

characteristic of his account. Structural synonymy, it seems, only makes sense when

considering a coarser-grained account of sentential meanings, according to which the

synonymous relations are explained in terms of their having the same semantic structure

(Horwich 1998*a*, pp. 163-164). The instance considered by him only concerns the active

and passive, but there are many varieties. For example, how are we able to explain the

synonymous relation between "Horwich's theory is not easy to prove," "To prove

Horwich's theory is not easy," and "It is not easy to prove Horwich's theory"? Each of

them is with a unique construction property. Except for this, Horwich's account is also

not able to explain structural ambiguity. According to Horwich's account, every sentence

has a unique construction property. If this is not the case, it is because either one or some

words it contains is vague or is ambiguous or has more than one explanatorily basic use

property. Structural ambiguity is left out. Apparently certain syntactic schemata such as '*a*

*n*s *and n*s *v*', where '*a*' means an adjective and '*and*' means 'and', are ambiguous.[54] We

do not know whether or not the second noun is bounded by the adjective. Further, this

account is too sketchy to explain complex syntactic features, such as the plural and

singular, tense, and many others. If the point is to provide an account of compositionality

---

[54] I do not even know how to fill in the connective 'and' except by adding '*and*' to this schema. This is a
general difficulty for understanding Horwich's account. Collins correctly observes that the symbol '*be*' in a
schema, say, '*n be a*' is not itself meaningless.

to explain ordinary use facts regarding complex expressions, then that every sentence has a unique construction property does not help.

Further, Collins criticises Horwich's account as not being able to reconcile with the general view of syntactic theories developed lately. It seems that certain syntactic features do not associate with syntactic schemata but with words they contain. One instance Collins gives is the different syntactic features between the word 'probable', which only takes finite clauses, and the word 'likely', which takes both finite and infinite clauses. Moreover, these features are not determined by any a priori reason, gleamed from reflection upon what language *must* be like or what it is *for* (Collins 2003, p. 418). Syntactic schemata do not explain this sort of syntactic feature. If any account of meaning or linguistic structure is inconsistent with such restrictions, it should be rejected outright on the grounds that it misdescribes the facts (ibid.). As Collins characterises, Horwich's syntactic schemata are *flat*. It lacks the necessary hierarchical structure imposed on a collection of words to accommodate their various semantic features. If Horwich's account aims to explain how people can understand complex expressions of a language, such as English, in terms of their implicit understanding of nothing more than a class of words and a class of syntactic schemata of that language, then this account fails miserably. On the one hand, this account fails because of the reasons characterised above. On the other hand, if the syntactic features of words are absorbed into the range of implicit understanding of the use properties of words, which is very unlikely to be accepted by Horwich, then all schemata are redundant.

If the acceptance property of a given word also gives circumstances in which the

syntactic features of the word are accepted, then the use property does not just govern the use of the word but also its syntactic connections to other words. If this is the case, then it is not that easy to distance the use-theoretic account from MEANING HOLISM. Further, apparently not every syntactic schema is applicable to every word. Horwich owes us an explanation of this. In addition, some words of languages such as Chinese, which do not have a syntactic item as clauses, are not translatable.

Finally, even if Horwich's account is capable of coping with all these difficulties, the attempt to construct a general account of proposition fails. In terms of my personal painful experience, translating more than ten books from English to Chinese, the syntactic structures of these two languages are very different. The following three points themselves are only examples. First of all, Quine is right on the point that there is no place-to-place corresponding between sentences of different languages. There is not even a guarantee that sentences have similar length. For instance, the Chinese translation of the sentence "Planets orbit stars" contains at least ten Chinese characters.[55] Second, Chinese has a very different grammar from English. The following are several examples.

(1) There are no plural endings in Chinese.

(2) There is no such a concept of clauses in Chinese grammar. If someone writes a

    sentence containing clauses exactly in an English fashion, it will be corrected.

(3) In English, there is a particular grammatical device to distinguish countable nouns

    from uncountable ones, such as one says 'a glass of water' instead of 'a water'. But

---

[55] Even Quine's consideration is very English. For instance, you cannot tell whether or not "Gavagai" is a sentence. This leads to the question of how to segment utterances of a remote language. The reasoning behind this is that pronunciations of words of English, or Latin phyla in general, contain syllable or syllables. Words with many syllables consist in many alphabets accordingly. But Chinese words do not have these features. According to the pronunciation of "Gavagai", it is about three Chinese characters for sure if it is Chinese.

this does not hold in Chinese. Almost every noun has a particular kind of terms like 'glass' in 'a glass of water'. Instead of saying 'a tree', for instance, Chinese say 'a __ of tree'.

(4) More importantly, verbs do not have tense variations.[56]

And this can go on, and on, and on.

Third, similar to English, Chinese relies both on words and phrases to form sentences. But there is no corresponding relation between a word and the complexes containing it that can be preserved in an English-Chinese or a Chinese-English direction. The sentence "Planets orbit stars" is almost a perfect example for the characterisation above. For instance, Chinese do not say 'a planet' or 'many planets' but say something like 'a __ of planet' and 'many __ of planet'. (Remember, no plural endings.) Most importantly, the reason for using at least ten Chinese characters to translate it is simply because the corresponding word of 'orbit' in Chinese is only a noun. Chinese does not have a corresponding verb to the English verb 'orbit'. So we have to translate this three word English sentence into something whose English translation is similar to the sentence "Planets in their orbits move around stars." Since there is no corresponding verb of 'orbit' in Chinese, inevitably we must use an alternative Chinese verb. This is not enough to translate "Planets orbit stars" because the substituted Chinese verb does not make plain that planets move around stars in their *orbits*. The translation of the noun 'orbit' has to fill in to ensure the translation to be correct. In order to explain this deviant case, what sort of syntactic schemata can Horwich invoke? The word 'orbit' being both a noun and a verb in English or being only a noun in Chinese is entirely contingent. I do not see how

---

[56] Therefore, Chinese does not have a corresponding grammatical device as subjunctive conditionals. This is probably the reason why many Chinese-speaking people have difficulty with understanding counterfactuals properly.

Horwich can explain away the above problem of translating "Planets orbit stars" into Chinese.

The fundamental difficulty of Horwich's use-theoretic account of word meaning and compositionality is its fine-grained character. According to this account, words of different languages are synonymous if and only if they have the same use property. The word 'celibataire', for instance, is translated as 'bachelor' rather than 'unmarried man' because in terms of the fine-grained character of Horwich's account 'celibataire' and 'bachelor' are both primitives but 'unmarried man' is not (Horwich 1998*a*, p. 163). We already saw this line of reasoning in many places of Horwich's account. However, in Chinese the translating words of both 'bachelor' and 'unmarried man' are complex expressions. The corresponding Chinese phrase of 'bachelor' is a combination of a two-word phrase whose meaning is similar to the word 'single' in 'a single life' and a word whose meaning is similar to 'man'. Sentences of different languages are synonymous, according to Horwich, if and only if they result from applying words with the same use properties in an exact order in a schema shared by those languages. But we already saw the flaw of this way of thinking in the instance of "Planets orbit stars" and its Chinese translation. They do not even fit in the same syntactic schema. It is inexplicable to believe that all contingent facts of English words and sentences are preserved in foreign languages. The particular semantic and synthetic features of foreign words are finely ground into meaningless pieces by Horwich's mill. Quine's qualms, characterised properly, explain why this failure happens. To look for a word-word corresponding manual for translation, one has to be clear first what *a* word is. Unless this is solved, the so-called fine-grained semantics can hardly begin. But in solving what a word is for a

remote language, it turns out that translation is indeterminate. Horwich's account does not quell Quine's qualms but is quelled by them.

Final Remarks

In the second section I considered how Quine's scepticism toward meaning is misunderstood by Horwich. As I have argued in sections three and four, neither Horwich's use-theoretic account of word meaning nor his account of compositionality can work. This raises doubt regarding the sensibleness of constructing an account of proposition without any complication caused by the concept of truth. This further weakens the rationale of adopting the minimalistic rather than the disquotational approach toward the deflationary concept of truth.

Chapter V

Between Quine's Disquotationalism

and Horwich's Minimalism

Almost every adherent of different deflationary theories of truth, holds that there is some sort of equivalence relation stronger than material equivalence between the two sides of all biconditional instances of the disquotational schema.[1] This feature of the relation is essential to the explanation of three major factors regarding the success of a deflationary theory of truth. The first of these is the functional role of the truth predicate to do the semantic ascent required to express infinite conjunctions and disjunctions. The second is the adequacy of the explanatory role in terms of which all facts involving truth are explained along with other suitable explanatory resources, and the third is the proper ascription of the truth predicate. Many philosophers locate their criticisms of disquotationalism in general, including minimalism, in one factor or another. In my discussions of these criticisms I will also provide my evaluation of Quine's disquotationalism and Horwich's minimalism. This chapter thus consists of four sections. The first three discuss the three factors for the success of a deflationary theory of truth mentioned above. The fourth questions Quine's attitude towards disquotationalism.

1. The Equivalence Relation and Semantic Ascent

In order correctly to set the explanatory order regarding the equivalence relation and semantic ascent, I will begin my discussion with the sense equivalence relation.

---

[1] As said in Chapter One, the deflationary theories of truth considered in this dissertation are disquotational and minimalist theories only.

1.1 The sense equivalence relation

Many opponents of disquotationalism claim that the sense equivalence relation holds between truth ascription sentences, such as "'Snow is white' is true", and the sentences to which the truth predicate is ascribed, in this case "Snow is white". Although this seems to be problematic for disquotationalism, it is in fact equally problematic for minimalism as well. But let us look at disquotationalism first.

Richard Schantz in "How Deflationists Diverge from Tarski" (Schantz 2000) makes the following claim:

> The general idea on which [the disquotational account] is based is that a sentence of the form '*s* is true' has the same meaning or the same cognitive content as the sentence *s*. (p. 318)

In the case of "'Snow is white' is true if and only if snow is white," he cites Quine's words of "calling the sentence true, we call snow white" in *Philosophy of Logic*, to support this reading of disquotationalism. However, Schantz continues to discuss the functional role of the truth predicate that cancels the effect of semantic ascent, and then to the expressibility of infinite conjunctions and disjunctions. No explanation is provided as to why the cited words are taken to support the claim that the sense equivalence relation holds between instances such as "'Snow is white' is true" and "Snow is white". In "The Metaphysics of Truth", Devitt also claims that when the truth term is attached to the quotation name of a statement, it yields a statement that is equivalent to that statement, such as the classic equivalence relation between "'Snow is white' is true" and "Snow is white (p. 583). If this sort of criticisms concentrates only on the truth ascription of singly given sentences without further elaboration, then they are based on the *redundantist*

readings of disquotationalism. In fact, some of opponents make this plain. For instance, Wolfgang Künne in "Disquotationalist Conceptions of Truth" (Künne 2000) distinguishes the disquotationalist redundancy thesis from the disquotationalist abbreviation thesis of infinite conjunctions and disjunctions.[2] And he criticises them *separately*.[3] This line of criticising disquotationalism, and minimalism as well, is not very promising, for none of the major supporters of disquotationalism go for the sense equivalence explanation explicitly.[4] For instance, following in Quine's footsteps, Hartry Field rejects the interpersonal and intrapersonal synonymous relations between linguistic expressions in general.

However, if there is no sensible connection between the redundantist reading and the abbreviation reading of the disquotational theory of truth, then it is hard to make sense of either. Of course, the abbreviation reading is partially the consequence of reading disquotationalism in the cases of singly given sentences in a redundantist way. A more

---

[2] Künne raises an argument inspired by Quine, Argument From Entailment, against the redundantist reading. Based on Davidson's observation in "What is Quine's View of Truth" that the instances of the disquotational schema are guaranteed to be true only in the very special case, namely, for a given person's idiolect, Künne reformulates the disquotational schema as
(D*) '*p*' is true *as used by me now* if and only if *p*.
"'Snow is white' as used by me now is true" hence entails
(a) "Snow is white" as used by me now is meaningful,
and
(b) Something is sometimes used by someone.
However, "Snow is white" entails neither of them.
In a footnote Künne indicates that Quine does not have very much sympathy for this argument. But I do not think that Quine objects to either (a) or (b) as a consequence of (D*), at least not (a). In "Meaning and Inference" the argument introduced by Quine is about the inference from 'Fa' or '$\sim Fa$' to $\exists(x)$ ($Fx \lor \sim Fx$). What Quine objects to is the assumption that the meaning of 'Fa' exists if 'Fa' is meaningful, but he has no objection to something like (a). What is rejected is the inference from the meaningfulness of 'Fa' to the existence of the meaning of 'Fa'. Further, (b) is ambiguous because the very *something* that is used could be the sentence itself or the meaning of the sentence. Quine only disagrees with the assumption that the latter suggests the existence of the meaning of 'Fa' taken as an entity. In either case it is not right to think that Quine disagrees with Künne's argument. For (a) is enough to discern the difference in meaning between "'Snow is white' as used by me now is true" and "Snow is white".
[3] Künne leaves out Quine as one of those that go for the sense equivalence explanation.
[4] In the following, however, it will be shown that some deflationists are committed to this explanation implicitly.

suited and hence weaker manner is to regard the disquotational truth predicate as

providing the necessary mechanism to *express* infinite conjunctions and disjunctions,

arguably in various degrees. Even in this weaker sense, Anil Gupta ("A Critique of

Deflationism", Gupta 1993*a*) urges us to acknowledge that, in order for the deflationary

truth predicate to express infinite conjunctions and disjunctions, the sense equivalence

relation in the instances of the disquotational schema, and minimalist equivalence schema

as well, is essential. Gupta's criticism is applicable to both the disquotational schema and

Horwich's equivalence schema. For the sake of simplicity, I will only consider the

disquotational schema in the following discussion.

Gupta begins his criticism with an example according to which the expressive power of

the truth predicate is manifested. Suppose we wish to affirm all sentences of the form

      ____ & snow is white.

Our wish to affirm all the instances of this form amounts to affirming the following

infinite conjunction by filling the blank of the form with all sentences:

(1) [Sky is blue & snow is white] & [Chicago is blue & snow is white] & ….

The disquotational function of the truth predicate makes (1) equivalent to

(2) ['Sky is blue' is true & snow is white] & ['Chicago is blue' is true & snow is white]

& ….

As far as the sentence names are what a universal objectual quantification ranges over,

using the pronominal variable 'x', (2) can be formulated as

(3) For all sentences x: [x is true & snow is white].

Based on this setting, Gupta's argument goes as follows. For (3) to express (1), (1) and (2)

need to be *equivalent*. In order for the truth predicate to play its role in forming

generalisations such as (3), (3) and (1) have to be *equivalent* too. This is accepted by

almost everybody, as the claim that the truth predicate has the function of *expressing*

some infinite conjunctions.[5] The difference between disquotationalism and

non-disquotationalism lies in how strong the equivalence relation is and whether or not

this is the only function played by the truth predicate. The disquotation function of the

truth predicate makes certain that (1) and (2) come out equivalent. As for the claim made

by the disquotationalist that the truth predicate is *only* a device of disquotation, the

equivalence relation between (1) and (2) has to be strong enough to justify it. For (1) to be

expressible in terms of (3), the relation between (1) and (2) can be varied from material,

necessary, to sense equivalence. The material and necessary equivalence relations only

yield that T-sentences are true or necessarily true respectively. But substantive truth

theorists have no problem agreeing with the necessary truth of T-sentences. Indeed,

according to Gupta, they have to.[6] Only when the relation between (1) and (2) is sense

equivalence, is the claim that the truth predicate is only a device of disquotation and

nothing else justified. Under this strong reading of the equivalence relation, however, the

claim that (3) expresses (1) is plainly false. (3), as a universal statement, and the

conjunction of its instances, namely, (2) do not have the same sense.[7] For Gupta, this is

the consequence of confusing what Quine calls the difference between affirming the

universal and affirming all its instances (Gupta, 1993*a*, p. 63).

Gupta's argument is twofold. On the one hand, it shows that both the disquotational and

---

[5] Infinite disjunctions are also expressible in terms of the same claim. Since the main issue here is about the expressibility of infinite conjunctions only, the part regarding infinite disjunctions is dropped from now on.
[6] Argument for this, please cf. Gupta 1993*a*, p. 69.
[7] Künne raises a similar argument, Argument From Conceptual Overloading. Please cf. Künne 2000, p. 187-188.

the minimalist theories of truth can at most explain all the instances of generalisations regarding truth, but not those generalisations themselves. On the other hand, the meaning of 'true' is not fixed by what Horwich calls the inclination to accept, without evidence, all the instances of the equivalence schema, for the acceptance of these instances cannot sufficiently explain the particular use of the truth predicate in truth generalisations.

The main point made by Gupta is the complaint that the disquotational and the equivalence schemata at most can explain every instance of a given truth generalisation but not the generalisation itself. Or, in other words, all the instances of a given generalisation can be derived from these schemata and other auxiliary resources, as long as these resources are truth-free, but not the generalisation itself. This complaint results from the fact that, without an infinitary inference rule, how can one explain the use of the truth predicate within a truth generalisation, such as "Every sentence of the form '*p* or not *p*' is true", in terms of the explanation of the uses of the truth predicate within all of its instances? If accepting the sense equivalence reading is not an admissible alternative, then disquotationalism and minimalism are not capable of explaining all the uses of the truth predicate. This criticism, on the other hand, does indicate that, as far as the disquotationalist and the minimalist can find a way to explicate the reasoning of infinitary inference, the sense equivalence reading and the related problems are dispensable. However, discussions with regard to the reasoning of infinitary inference are highly related to the disquotational and the minimal accounts of the adequacy of the explanatory role of the truth predicate in terms of which all facts involving truth are explained. Therefore, Gupta's criticism is insignificant, if in the following section the reasoning of infinitary inference can be explained.

However, Gupta does not only raise one argument only. But the discussion of his further arguments is postponed. Gupta does not contest the claim that instances of the disquotational schema are necessarily true. According to him this is an inspirational insight. However, some opponents of disquotationalism and minimalism even disagree with this claim. For them, disquotationalism and minimalism cannot even make sense of the necessary equivalence relation of the instances of the disquotational schema or the equivalence schema. So I turn to this issue first.

1.2 The necessary equivalence relation

We have discussed, in Chapter Two, the criticism made by Marian David that the instances of the disquotational schema, such as "'Snow is white' is true if and only if snow is white," are not necessary truths. He is not alone in this line of criticism. The scope of this criticism is different from one opponent to another. For instance, James Van Cleve in "Minimal Truth Is Realist Truth" (Cleve 1996, p. 869) and Pascal Engel in *Truth* (Engel 2002, p. 47) only ascribe this problem to disquotationalism whose primary truth bearers are sentences. McGrath, on the other hand, includes minimalism as well (McGrath 1997*b*, pp. 79-84; McGrath 2000, pp. 35-39).

The basic idea behind this line is that truth depends on meaning. Whether or not an instance of the disquotational schema holds, depends on the contingent fact of what the quoted sentence in the instances means. Künne presents an argument according to this reasoning (Künne 2000, p. 180). Let the sentences

(S) Snow is white

and

(T) "Snow is white" is true

embed in a modal context, as consequents in a counterfactual conditional:

(MS) If we all came to use the sentence "Snow is white" for saying that grass is red, it
would not be the case that *snow is white*.

(MT) If we all came to use the sentence "Snow is white" for saying that grass is red, it
would not be the case that *"snow is white" is true*.

(MT) is true, but (MS) is false. Künne then argues that, if the ascription of truth to "Snow

is white" in (T) were nothing more than the ascription of whiteness to snow in (S), the

application of the modal operator "it would not be the case that" to (S) and (T) could not

impose the assignment of different truth values to (MS) and (MT) (ibid.).[8] Therefore, the

instances of the disquotational schema are contingent. And the contingency of the

disquotational schema is very likely to force people always to confuse (MS) with (MT) or

(MT) with (MS).

McGrath thinks that strong deflationary theories of truth, such as Horwich's minimalism,

cannot get away from this attack. Although Horwich employs propositional forms to

construct his minimalism, according to McGrath the instances of the equivalence schema

are embedded in the same muddy field.[9] This is correct. When formulating the

equivalence schema, Horwich claims that for every English declarative *sentence*, '*p*',

there is a corresponding equivalent *sentence*, "the proposition that *p* is true". It is easy to

see that, with a little adjustment, Künne's argument also applies to minimalism.[10] In

---

[8] Künne raises this argument against the redundantist reading of disquotationalism. But the abbreviation
reading is very likely to be doomed if this is established.
[9] David in "Minimalism and the Facts About Truth" also notes this.
[10] Precisely because of this reason, McGrath proposes his *weak* deflationism according to which only the

terms of the observation that minimalism eventually has to gravitate to the mud, regarding

propositions as the primary truth bearers does not seem to provide a significant difference

from disquotationalism in dealing with truth ascriptions in the cases of vague, ambiguous,

and indexical contexts. If there is any way to dispose of these problematic cases, it will

not only be available to minimalism.

This line of criticism affects almost all the truth theories of the major deflationary

theorists, perhaps except Quine's. They respond in one way or another in order to

maintain the important feature of disquotationalism, that all instances of the

disquotational schema are necessary. Horwich on the other hand does not seem to care too

much about this. Perhaps he takes this as a minor point under which the core of

minimalism is intact. He does consider the instances of the equivalence schema as

necessary though. In a footnote he says that the alternative strategy—the one that he

prefers—is to keep the theory of truth un-modal and simple, and instead to derive the

necessity of the axioms of the minimal theory from a separate theory of necessity. And the

theory proposed by him is that necessary truths are specified by their being *explanatorily*

fundamental. Establishing this theory, and providing that the minimal theory is

explanatorily basic, it follows that the axioms of it are necessary.[11] As David observes in

"Minimalism and the Facts about Truth" (David 2000), this theory of necessity is plainly

false. He says, "[N]atural science contains explanatorily fundamental laws that are not

necessary truths" (p. 167). Or consider Gupta's example in "A Critique of Deflationism",

the statement that no chemical reaction will produce caustic soda from saltpetre and

---

theory of propositional truth is deflationary and other theories of sentence, belief, and utterance truth are
inflationary.
[11] Please cf. Horwich 1998*b*, p. 21 footnote.

sulfuric acid is considered as necessary but is still in need of a deeper explanation by chemistry of why it holds. If this statement is necessary, then according to Horwich's theory of necessity it is explanatorily fundamental. But this contradicts the fact that chemistry provides a deeper explanation for it. Now Horwich has to reject the necessity of this statement, which might cause a cleavage between the intuitive set of necessary truths and the set of necessary truths according to his theory, and insist that chemistry is explanatorily fundamental. Even putting aside David's observation that this insistence is ill-founded, there is another more serious problem. What is explanatorily fundamental normally depends on the explanatory purpose. It is hard to believe that there is an absolute order of explanatory fundamentality relative to practical purposes of explanation. It follows that Horwich fails to give us any reason to accept that the instances of the equivalence schema are necessary.

For solving the problem of the modal status, Vann McGee provides a modified disquotational schema (DSI):

(DSI) "Φ" is true in my idiolect if and only if Φ.

He then claims that all instances of (DSI) are necessary. As long as we keep the hidden indexicals straight, he so explains, we always arrive at (MT) rather than (MS).[12] If this is the end of his explanation, this account is less appealing. For this account leaves everybody in the dark when accounting for other people's use of the same group of sentences. McGee extends his idiolect disquotationalism to a global one by quelling Quine's qualm of translation. For him,

> [T]he actual field linguist never finds herself in the position of Quine's radical translator, because the field linguist is willing to posit [psychological] hypotheses that Quine wishes to

---

[12] Please cf. McGee 1993, pp. 90-91.

disallow to the effect that the native's conceptual system is broadly similar to her own.
(McGee 1993, p. 99)

McGee claims that the problem of the inscrutability of reference under this consideration is untenable. There are some differences one cannot mistake between the conceptual roles for the word 'rabbit' and for the phrase 'undetached rabbit parts'. Consequently, for an indigenous informant whose conceptual frame is at all similar to my own, the word 'gavagai' used by her more closely resembles my use of 'rabbit' rather than 'undetached rabbit parts'. I will not consider this explanation more fully. However, let me point out two things. First, there is no dispute that the field linguist is very likely to translate 'gavagai' as 'rabbit' rather than 'undetached rabbit parts', as I indicated in the previous chapter. There is even no dispute over bringing some psychological hypotheses into account in the practical vein of translation. What is in dispute is whether or not this translation is determinate. So it is better not to include psychological hypotheses in considering radical translation in the theoretical vein. McGee's account seems to miss the point. Second, it is unclear how the relation of resemblance accounts for translation. More importantly, resemblance is not a transitive relation. Hence, the chance of a tenable account of the global disquotationalism is slim. The price of accounting for the necessary status of all instances of the disquotational schema is not low.

Hartry Field gives a somewhat similar analysis in "Deflationist Views of Meaning and Content" (Field 1994). That is:

[A] person can meaningfully apply "true" in the pure disquotational sense only to utterances that he has some understanding of; and for such an utterance u, the claim that u is true (true-as-she-understands-it) is cognitively equivalent (for the person) to u itself (as she understands it).[13] (p. 105)

---

[13] There is a qualification according to Field that keeps the two from being fully cognitively equivalent. This is because the claim that u is true makes an ontological commitment to the utterance u, while the

This is what Field calls *pure disquotational concept of truth*. Further, Field explicates the notion of cognitive equivalence in a footnote:

> [T]o call two sentences that a person understands 'cognitively equivalent' for that person is to say that the person's inferential procedures license a fairly direct inference from any sentence containing an occurrence of one to the corresponding sentence with an occurrence of the other substituted for it.[14] (p. 106)

According to Field this is certainly not an instance of substitutional quantification. And he claims that this cognitive equivalence implies that this kind of inferences is empirically indefeasible and even conceptually so, if semantic paradoxes can be ignored. This cognitive equivalence between 'u is true (as I understand it)' and the utterance u (as I understand it) guarantees that "u is true iff u" is conceptually necessary, (if not logically necessary). However, this account owes us a complete explication of the sort of inference that is appealed to regarding cognitively equivalent sentences, according to which the notions of truth and truth preserving cannot play any role.

The next issue is how to apply the disquotational truth predicate to sentences that the speaker does not understand.[15] This application is definitely not in a pure disquotational sense, because pure disquotational truth is by definition not about sentences or utterances that a given speaker does not understand. The account that allows the truth predicate to apply to utterances or sentences that one does not understand is therefore called by Field the *extended disquotational theory of truth* (1986, p. 59; 1994, p.128). The first way to explain extended disquotational truth is to take a theory of content into account. A foreign sentence is true in the extended disquotational sense as far as it is (intrapersonally or even

---

utterance itself does not. But this is a relatively minor point and will be kept tacit in the following discussion.

[14] There are certainly some qualifications of this explication. But for the sake of simplicity, they will be ignored.

[15] This will be further clarified in the third section.

interpersonally) synonymous with a sentence that I understand. The general description of it is (ED):

(ED) for any foreign sentence *s*, *s* is true iff *s* is synonymous with a sentence 'p' (that I

   understand) and p.

The notions of *intrapersonal synonymy* and *interpersonal synonymy*, according to Field, have to be explained in terms of something in which the notion of truth conditions does not play any important role. Ayer's verificationism is among those theories that can provide the desired help. In fact, two sentences are interpersonally synonymous according to verificationism if and only if they have the same set of verification conditions. However, verificationism is undesirable because of its overly simplified character. This extended version of disquotational truth is still less preferable, according to Field, even if the theory of content is more refined. This is because it is probably better not to mix some semantic elements into the logical aspects of truth (Field 1994, p. 130).

Field then proposes the notion of *quasi*-disquotational truth to replace extended disquotational truth in counterfactual circumstances. (For simplicity, I will use subscripts '$_{pd}$' and '$_{qd}$' to distinguish the pure disquotational truth predicate from the quasi-disquotational truth predicate.) The definition of quasi-disquotational truth is as follows:

(QD) $^\uparrow$ (S is true$_{qd}$ iff $\sum$p ($\exists m$ (*m* is the meaning of S and @ (*m* is the meaning of 'p')) and

   p)). (Field 1994, p. 131)

The idea is to invoke not just the notion of synonymy, but the notion of meaning as well, and to find a un-truth-theoretic way to explain it. It also appeals to an *actually operator* '@' that temporarily undoes the effects of the modal operator (ibid.). One of the main

reasons why this quasi version is important, is that it is more likely to be our ordinary use of the truth predicate as maintained by Field. This is because we seem to incline to accept certain counterfactuals such as "If we had used 'Snow is white' differently it would have had the truth conditions that grass is red." And to avoid the accusation that the ordinary way of using 'true' is not the purely disquotational sense, we have to appeal to something similar to quasi-disquotational truth.

There is another option for explaining the truth ascription to sentences that one does not understand though, and that is to define the truth of foreign sentences as being *true relative to a correlation* of it to one of our sentences. A sentence is true relative to the correlation if our sentence that is in correlation with it is true in the purely disquotational sense (Field 1994, p. 128). Of course, Quine's translation theory is among those that can explain the correlation between foreign sentences and our own. The feature of this option is that it is lacking an objective synonymy relation. The correlation, such as a translation between sentences, is highly context-sensitive so that it does not make sense to talk about correct or incorrect translations, but good or bad ones. Consequently, the application of the truth predicate to sentences that I do not understand does not have objectivity either. This option allows us to give a *quasi*-disquotational explanation of the counterfactual (MT). That is,

> In considering counterfactual circumstances under which we used "Snow is white" in certain very different ways, it is reasonable to translate it in such a way that its disquotational truth conditions relative to the translation are that grass is red. (Field 1994, p. 133)

Since Field follows in Quine's footsteps in holding a sceptical view toward the notion of

synonymy, intrapersonal and interpersonal, in the Postscript (2001) of "Deflationist Views on Meaning and Content", he goes for the second option and makes some further refinements to the definition of quasi-disquotational truth:

(Qd) S is true$_{qd}$ (at a possible world v) iff S is to be translated by a sentence of mine (in the actual world) that is purely disquotationally true (at v). (2001, p. 152)

Apparently, Field's account uses the conceptions of pure disquotational truth and extended disquotational truth to explain the necessary status of all instances of the disquotational schema, and further uses the conception of quasi-disquotational truth to make sure that we will always reach something similar to (MT) in counterfactual circumstances.

According to McGrath, this is not satisfactory. He points out that Field's strategy is to conceive of the relevant modal intuitions as special cases of intuitions about truth in other languages (McGrath 1997*b*, p. 82). The language used by our counterfactual selves in the relevant world is treated as another language, truth in which can thus be explained in the extended disquotational way as truth in actual other languages (pp. 82-83). The question McGrath raises is this: the issue is not just whether or not the counterfactual intuitions can be accounted for by deflationist, but whether she can *explain* the facts these intuitions register (p. 83). We need to know why in the counterfactual intuition, (MT), "Snow is white" would have the disquotational truth conditions relative to a translation that grass is red. Why do these objective counterfactual intuitions relate to practical translations? There is still another feature of Field's account to consider. The extended- and quasi-disquotational truths do not overstride the given speaker's own pure disquotational truth, unless we are given some practical hypotheses such as that the

translation/interpretation manuals of people speaking the same language tend to be convergent. Most importantly, McGee and Field, especially Field, seem to be committed to the sense equivalence relation implicitly.

The damage from arguments similar to Künne's is not only to deflationary theories of truth. If the instances of the disquotational schema and the equivalence schema are contingent, then the disquotational schemata of reference and satisfaction are also doomed. If we all came to use the sentence "Snow is white" for saying that grass is red, then "Snow is white" would not have had the disquotational truth conditions that snow is white. In Chapter Two I mentioned Halbach's reply to this sort of argument, which, according to his terminology, he refers to as 'Levy's argument'. The main point is that, under this sort of counterfactual reasoning, there are no necessary truths. For instance, "The prime number between four and seven is smaller than six" would have not been necessarily true if the phrase 'prime number' were used differently, as, for example if it were used to mean what 'even number' means. Therefore, it is not just that we should follow Field (1994) in emphasising that there are logical uses of the truth predicate that are immune from Levy's argument, but that there *must* be uses of the truth predicate to which this sort of arguments does not apply (Halbach 1999, p. 173).[16] This leads us to the final discussion of the equivalence relation.

1.3 The material equivalence

The origin of disquotationalism comes from Quine's ideas of semantic ascent and descent.

---

[16] Halbach also gives us a quite different answer to the modal problem in Halbach 2000. Since his main concern is to provide a solution to the generalisation problem, his solution will be discussed in the following section.

The disquotational schema of truth therefore is supposed to grasp these ideas. The peculiar feature of semantic ascent and descent, as I discussed in the second section of Chapter Two, is that the mentioned sentence and the used sentence are exactly the same—the same sentence with the same empirical content or the same interpretation. It simply does not make sense to substitute talk of whether "Snow is white" is true for talk of whether snow is white, when the sentence "Snow is white" in "'Snow is white' is true" is used for saying that grass is red. Therefore, counterfactual reasoning does not directly apply to the instances of the disquotational schema. Though disquotationalism takes sentences as the primary truth bearers, this does not amount to its being the case that it is a degenerate case by which the fundamental feature of semantic ascent and descent cannot be correctly grasped. Furthermore, this does not suggest that counterfactual reasoning does not apply to sentences with the truth predicate, but that counterfactual reasoning about sentences with the truth predicate is derivative. For instance, Field's quasi-disquotationalism is derived from pure and extended disquotationalism.[17]

More precisely, we should understand the disquotational schema in terms of the same sentence being mentioned and used relative to a language with a fixed interpretation scheme. This understanding of the disquotational schema is difficult to form in a schematic formulation. McGee's idiolect disquotationalism, that "Φ" is true in my idiolect if and only if Φ, and Field's pure disquotationalism, that "u is true (true-as-she-understands-it)" is cognitively equivalent (for the person) to "u itself (as she understands it)", cannot fully illuminate this point. However, based on the proper understanding of the disquotational schema, the instances of this schema only need to be

---

[17] Whether or not Field's quasi-disquotationalism is tenable is not at issue here.

held materially equivalent. So we may follow my characterisation in Chapter Two to

formulate the disquotational schema as an axiom schema in a truth theory, within which

all disquotation sentences of the base language sentences are axioms, as follows:

(D$_Q$) '*p*' is true iff *p*,

> where the first '*p*' is replaced by any sentence of the base language and the
>
> second '*p*' by the same sentence with the same interpretation.

Using Gupta's example, (3) expresses (1) because applying the disquotational schema to

each instance of (3) in terms of the material equivalence reading, cashes out exactly each

corresponding conjunct of (1).[18] In *Truth*, Horwich says the following:

> [T]he explanatory basic fact about our use of the truth predicate is our tendency to infer
>
> instances of 'The proposition *that p* is true' from corresponding instances of '*p*', and vice
>
> versa, whenever (a) each '*p*' is replaced with tokens of an English sentence, (b) these
>
> tokens are given *the same interpretation* as one another, (c) under that interpretation they
>
> express the content of a statement (a proposition), and (d) the terms 'that' and
>
> 'proposition' are given their English meanings. (p. 126, the *italics* in (b) are mine)

Therefore, my reasoning regarding how to interpret the disquotational schema of truth is

not incompatible with Horwich's explanation of the equivalence schema of truth.

The necessary status, in addition, is not essential to the explication of the expressibility of

infinite conjunctions in terms of the truth predicate. Granted, the material equivalence

relation of the disquotational schema, (D$_Q$), is a bit stronger than the usual one, for the

sentences mentioned and used are restricted to the same sentence with the same

interpretation. This is precisely why Quine takes truth as *immanent* to home languages,

for Tarski's T-sentences are not automatically instances of (D$_Q$).[19] The expressibility of (1)

---

[18] I am just borrowing Gupta's own case to show his misconception. The criticism that it is incoherent to use (1) to (3) in constructing the argument still holds.

[19] This is why I made the claim in the second section of Chapter Two, that the stronger disquotational property under Quine's account is not preserved under the recursive definition of satisfaction.

in terms of (3) in the above instance, depends on both the material equivalence relation

and the additional restriction about the sentence mentioned and used in every instance of

(3). However, this restriction does not yield any modal relation to the schema and its

instances. If it should be held that the disquotational schema yields instances with a

necessary status, then it has to be because of other reasons.[20]

In terms of all the deflationists I discussed so far, only McGrath and Horwich propose the

material equivalence relation for explicating the expressibility of infinite conjunctions

using the truth predicate. In the last chapter we saw that McGrath proposes his finite

theory of propositional truth as that, for all propositions p, p is materially equivalent to

Tr(P). In the postscript of the second edition of *Truth*, in answering the criticism from

Gupta mainly discussed in 1.1, Horwich replies:

> The generalizing function of truth is perfectly well fulfilled as long as instances of the
> truth schemata, understood as merely *material* biconditionals, are accepted.
> But this point [That is, Gupta's criticism of the sense equivalence relation between 'The
> statement *that p* is true' and '*p*'] does not undermine the minimalist story about the
> function of truth; for, as just mentioned, that function requires merely that the
> generalizations permit us to *derive* the statements to be generalized—which requires
> merely that the truth schemata provide material equivalences.
> (Horwich 1998*b*, p. 124)

By applying the propositional form or, in other words, propositional structure of '$<p>$ is

true iff p' to the proposition, say, "$<$Snow is white$>$", they can ensure that the

propositions on both sides of 'if and only if' are exactly the same proposition that snow is

white.[21] The functions of semantic ascent and descent in turn make sure that the material

---

[20] However, considering for what reason one should consider disquotation sentences as necessary, will not
be discussed..

[21] Horwich clearly makes this point in "A Defense of Minimalism", p. 153.

equivalence relation is enough to explicate the expressive function of the truth predicate.[22]

In considering the case of disquotationalism, requiring that the sentences mentioned and used in every instance of the disquotational schema are the same sentences relative to a language with a fixed interpretation scheme is vital to a proper understanding of the disquotational schema, (D$_Q$). This, on the other hand, also shows that although the disquotational schema and the equivalence schema are not sense equivalent, they are still capable of explaining the expressing of infinite conjunctions.[23] Whether or not disquotationalism and minimalism can explain general facts involving truth—that is, the affirmation of infinite conjunctions—will be discussed in the following section. But it is clear that so far I cannot find anything giving minimalism the advantage over disquotationalism.

The second aspect of Gupta's criticism rests on Horwich's account of the meaning of 'is true'. In *Meaning*, Horwich's explication of the meaning of the word 'true' is as follows:

(TRUE) The acceptance property governing our total use of the word "true" is the inclination to accept instances of the schema 'the proposition *that p* is true if and only if *p*'. (p. 45)

In a footnote (23) of "A Defense of Minimalism" (Horwich 2001*a*), Horwich makes it clear that he distinguishes *inclinations* from *dispositions* of accepting instances of the equivalence schema because of consideration regarding the liar paradox. This, in turn, is one of the important issues of the coming section. Now, Gupta claims that, for

---

[22] Providing that the reasoning of infinitary inference is well explained by disquotationalism and minimalism.

[23] This is probably one of the reasons why Horwich would like to keep his theory of truth un-modal and simple.

deflationism to dethrone the truth-conditional account of meaning, the instances of the

equivalence schema have to explain the meaning or sense of 'true'.[24] By this, he means

that our understanding of the meaning of 'true' depends entirely on our prior

understanding of the meaning of *all* (declarative) sentences of a given language. Again,

this cannot be correct, according to Gupta, because the conceptual resources of

understanding 'true' are enormously massive. It is not difficult to find the basic

assumption of this criticism—that is, the sense equivalence relation. However, in terms of

my clarification of the understanding of the disquotational schema, it is without doubt that

this assumption is totally wrong, and hence so is Gupta's argument.

Quine certainly would not agree with Horwich's (TRUE) as explicating the *meaning* of

'true'. Nevertheless, the general idea behind it is probably acceptable in that in Quine's

terms, all instances of the disquotational schema are stimulus analytic.[25] Following my

clarification of the understanding of the disquotational schema three paragraphs back, it is

quite clear that a proper understanding of the disquotational schema can be nicely

separated from any account of meaning, provided that these accounts of meaning are

non-truth-conditional.[26] What is essential to our understanding of 'true' is the peculiar

feature of the disquotational schema, that the sentences mentioned and used in every

instance are exactly the same sentence with the same interpretation. What is not essential

is what the sentence mentioned and used means. The truth predicate is really transparent

---

[24] Gupta's argument is based on Dummett's argument in his celebrated paper "Truth", according to which the sense equivalence relation is also regarded as essential to deflationism.
[25] Stimulus analytic sentences are by definition those sentences that speakers would assent to under any circumstance.
[26] There are, of course, some necessary conditions for the applicability of the truth predicate; for instance, with the syntactic feature of declarative sentences and with assertoric content. But this is not unique to disquotationalism, but is true of deflationary truth theories in general, especially the minimal theory of propositional truth. Please see section 3.

in this sense. It also nicely accommodates the fact that what is true depends doubly on meaning and reality. Further, it is unnecessary to restrict the language, in the present case, English, to be relative to a particular space-time frame—that is, the language can be seen as evolving over time. There is no need to adopt McGee's idiolect disquotationalism or Field's pure disquotationalism, because the uses of the truth predicate of people speaking the same language incline to coincide with one another. For, unlike McGee and Field, the language we are talking about is not anyone's idiolect. The impression that the extension of the disquotational truth predicate varies each time new terms are introduced into the language, is caused by the definitional formulation, such as David's (D), (DM), and (D$_M$) discussed in Chapter Two, according to which the range of truth ascription depends on the substitution class. It is now clear that this is nothing but a misconception. If any new term introduced into a language causes problems to the disquotational theory, then it equally causes problems to substantive theories too. English speakers would accept instances of the disquotational schema, such as "'Snow is white' is true if and only if snow is white," no matter what. Another reason not to adopt McGee's or Field's accounts, is the prominent role played in them by the notion of content.

Before we move on to the following section, one thing should be noted. In "Analyticity, Carnap, Quine, and Truth" (David 1996), David observes that Quine's argument in "Two Dogmas of Empiricism" against Carnap's definition of analyticity—that is, the definition in terms of the semantical rules or, in other words, meaning postulates of a given language—has a similar impact on other semantic concepts, including the concept of truth.

Carnap's idea is that, once you choose a convenient language form, there is a set of

(analytic) sentences, each of which is true according to the semantical rules of this

language. In other words, when the language is fixed, a certain set of analytic sentences is

also fixed. The drawback of this account, Quine argues, is that it only labels some

sentences as 'analytic' or 'semantical rule' and leaves them unexplained. Further, this

account transfers the unclearness of the concept of analyticity to the concept of semantical

rules (cf. Quine 1953*b* section 4).

A closer investigation and criticism occurs in "Carnap and Logical Truth" (Quine 1954),

written three years later.[27] One of the main points is about the distinction between

legislative and *discursive* postulation. Legislative postulation institutes truth by

convention and discursive postulation is merely the selection of certain true expressions

(from a pre-existing body of truths) as a basis from which to derive others (p. 111). There

is also a similar distinction between legislative and discursive definitions. Following these

distinctions, most semantical rules and postulates in Carnap's *rule book* should be

classified as discursive ones, if Carnap is attempting to explicate the intuitive concept of

analyticity. These semantical rules and postulates simply state pre-existing truths, whose

selection is arbitrary, or relative to our demonstrating purposes or interpretation.

Consequently, there are no statements that are *true by* these rules or postulates. Otherwise,

all true sentences should be considered as semantical rules or postulates since they are

true *by* themselves. If Carnap's attempt is to construct a new (artificial) linguistic system

to explicate the intuitive concept of analyticity, then those rules and postulates are

legislative even though this legislative procedure is still arbitrary. If the reason for

---

[27] "Two Dogmas of Empiricism" was written in 1951.

constructing a formalised language L is to capture the intuitive concept of analyticity, then this legislative procedure can never be successful. These rules, or postulations, are constructed in language L and hence do not belong to ordinary languages. The analytic statements in language L, then, are not coincident with the analytic statements in ordinary languages. It follows that the attempt at capturing the intuitive concept of analyticity fails[28]. In a nutshell, if it is impossible to define the concept of analyticity as analytic-in-L for a variable L ranging over all languages in terms of a restricted, indexed concept of analyticity, then to define true-in-L for a variable L ranging over all languages in a similar fashion is also not possible.

Surprisingly, David so complains, Quine later conceded that the debate between him and Carnap merely concerns the explanandum (David 1996, p. 284).[29] David's surprise, I think, results from his misconception of Quine's idea about truth. It is rather dubious that analyticity can sensibly be defined as S is analytic for L (for both variable S and L). For instance, the claim that "Bachelors are unmarried men" is analytic for any language L is problematic. This claim implies that it makes sense to talk about the analyticity of this sentence in Mandrin, say. But for Mandrin speaking people who do not understand English, this sentence is nothing more than gibberish. On the other hand, Carnap's account consists of two parts. First is about the definition of analyticity relative to a given language. Then utilising the notion of translation to see whether or not the above-mentioned instance can make sense in another language. Although Quine has a strong view against the absoluteness of meaning and translation, this is irrelevant to the

---

[28] M. G. White in "The Analytic and the Synthetic: An Untenable Dualism" (1950) made a similar criticism as follows: "What some philosophers do is to pretend that natural languages are really quite like these artificial languages; and that even though there is no rule-book for them, people do behave *as if* there were such a rule-book" (ibid.: 82).

[29] Quine admits this in *Dear Carnap, Dear Van*, Quine to Carnap, 1951-3-29.

explanatory *procedure* of Carnap's account. Similarly, 'true in *L*' (variable *L*) should be explicated in a similar fashion.

In "Notes on the Theory of Reference", Quine says:

> In Tarski's constructions, moreover, we have an explicit general routine for defining truth-in-*L* for individual languages *L* which conform to a certain standard pattern and are well specified in point of vocabulary. We gave indeed no similar single definition of 'true-in-*L*' for variable '*L*'; but what we do have suffices to endow 'true-in-*L*', even for variable '*L*', with a high enough degree of intelligibility so that we are not likely to be averse to using this idiom. (p. 138)

In his paper, David twists 'true-in-*L* for individual languages *L*' and "'true-in-*L*' for variable *L*" into 'indexed truth predicates restricted to individual fragments of English' and 'true-in-English'. This is the consequence of his failing to appreciate Quine's reasoning as characterised above. Considering David's misinterpretation of disquotationalism, this is not surprising. The *general routine* is not available in defining analyticity, purely because 'is analytic' is not transparent. The words 'general routine' can of course be regarded as about the disquotational schema. The corresponding truth predicates in different languages are similarly accounted for in terms of the corresponding disquotational schemata of them. Therefore, whether or not we can have a general definition of 'true' depends on how strong our account of meaning is. This will be discussed in the third section.

2. Explaining General Facts Involving Truth

For disquotationalism and minimalism to be feasible theories of truth, the most important task is to explain the overall use of 'true' in terms of the corresponding schemata of truth. For instance, one might incline to accept the sentence "'Water is transparent' is true if and

only if water is H_2O". The disquotational schema and the equivalence schema, along the line of its material equivalence feature, can well explain this. I only give an explanation here in terms of the disquotational schema. This sentence is a derivation from the scientific explanation of the transparency of water, according to which water is transparent if and only if it is H_2O and the disquotation sentence that "Water is transparent" is true if and only if water is transparent.[30] Among these explanations, the most important part is about general facts involving truth. Or, in other words, does affirming all instances have a relation to affirming the corresponding generalisations close enough to establish the relation that, by affirming a generalisation or an infinite conjunction, one can affirm the other? In the following, we will first see some answers provided by two disquotational theorists, Field and Halbach. Then I will discuss one of Horwich's instances of *modus ponens*—that is, if one sentence/proposition implies another, and the first one is true, then so is the second. By doing this, I seek to discover whether or not disquotationalism and minimalism could satisfactorily account for general facts involving truth. Moreover, I would like to show that in what reasoning one can depend the explanation of all general facts involving truth on the material equivalence relation occurring within each instance of the disquotational and the equivalence schemata. And finally, I aim to show a deficiency that is unique to minimalism, that is, the problem of the liar paradox.

But let me briefly discuss some alternatives advanced by some defenders of disquotationalism. Field (1994) proposes to begin from the idea of building schematic

---

[30] A more complicated case is generated by substituting the word 'because' for 'if and only if' in the sentence explained and the scientific explanation. Horwich seems to think that this is problem free (Horwich 1998*b*, pp. 49-50 and pp. 105-106). Although I doubt that this will work, I ignore this issue.

variables into the language and subjecting them to a rule of substitution.[31] Unlike

Horwich, Field accepts the truth schema itself, and not just its instances, as part of a

language. It follows that this device naturally gives the disquotational theory of truth the

desired deductive power to justify generalisations. For instance, to prove that, whenever

modus ponens is applied to true premises, the conclusion is true, we can do as follows:

1. If p, then q   premise

2. 'p' is true   premise

3. 'p' is true iff p   assumption, disquotationalism

4. p   from 2, 3

5. q   from 1, 4

6. 'q' is true iff q   assumption, disquotationalism

7. 'q' is true   from 5, 6

So 'If p, then q' and 'p' is true, then 'q' is true

All general facts involving truth are supposed to be derived in a similar fashion.

According to Field, there are at least three advantages of this account. First, the theory of

truth does not need to be revised as the language is expanded. Second, the use of the

schematic formula "'p' is true if and only if p" in the axiomatisation seems to more

directly capture the core of the notion of truth. Finally, and most significantly, according

to Field, the theory does not depend on our having a compositional account of the

functioning of the other devices in the language. That is, the accounts of reference and

satisfaction are expendable in explicating compositionality (cf. Postscript of Field 1994, p.

142).

---

[31] Field supplements his view, in the Postscript of Field 1994, with the claim that the syntactic theory also needs to be done with schematic variables. For the sake of simplicity, the discussion of this issue is left out.

Presumably, the second advantage is just about my *proper understanding of the*

disquotational schema in the discussion of the material equivalence relation. The third is

important. It is about how truth *commutes*, borrowing Halbach's terms, with logical

connectives. In terms of a similar procedure, for instance, it is not hard to reach the

conclusion:

'p or q' is true if and only if 'p' is true or 'q' is true.

Shapiro raises one problem of Field's proposal. He questions the satisfactoriness of this

proposal in accounting for sentences that one does not understand, even in the extended

disquotational sense (Shapiro 2003, p. 129). Extended disquotationalism is not helpful

because the substitution range of schematic variables still boils down to the same class of

sentences that one understands. This will be discussed in the next section. Horwich, on

the other hand, questions the reasonableness of attributing schematic reasoning to the

great majority of people. This seems to be too sophisticated an activity to be attributed to

ordinary people (Horwich 2001*a*, note 20).

Halbach also gives us a quite different answer in Halbach 2000. According to his

modalised disquotationalism (MD), the necessary status of the disquotation sentences,

that is, the instances of the disquotational schema, can be accounted for by the notion of

truth-analyticity.[32] He uses the following as an axiom of truth:

> "'*A*' is true" and "*A*" are truth-analytically equivalent for every sentence "*A*"
>
> of the base language. (p. 163)

'Truth-analytically equivalent' in turn can be defined in terms of truth-analyticity and

---

[32] For the preliminary formulation, please cf. the second section of Chapter Two, under the title 'A
preliminary set-up of a proper formulation', and the third section of the same chapter, under the title
'Expressibility'.

material equivalence:

"*A*" and "*B*" are truth-analytically equivalent if and only if (the proposition

expressed by) the sentence "*A* if and only if *B*" is truth-analytic (ibid.).[33]

In addition to all the disquotation sentences in the truth theory, two further principles are

added in order to define truth-analyticity. That is,

(L) Truth-analyticity is closed under predicate logic;

and a reflection principle

(R) For every formula "*A*(*x*)" with one free variable *x*, the following is derivable:

For all "*B*": if "*A*(*x*)" is truth-analytic for "*B*", then *A*("*B*") (p. 164).

Then it can be claimed in the truth theory, that all disquotation sentences (of the base

language) are truth-analytic. The virtue of this account, according to Halbach, is its

capacity to allow truth to commute with the connectives (p. 165). Unlike Field, the

reasoning is not schematic, but is about universally quantified sentences. For instance, it

can be proved that truth commutes with the connective of negation:

1. All disquotation sentences are truth-analytic.   Assumption, MD

2. For all "*A*": "'*A*' is not true iff not *A*" is truth-analytic.   from 1, (L)

3. For all "*A*": "'not *A*' is true iff not *A*" is truth-analytic.   from 1

4. For all "*A*": "'not *A*' is true iff '*A*' is not true" is truth-analytic.   from 2, 3

5. For all "*A*": "not *A*" is true iff "*A*" is not true.   from 4, (R)

All the general facts involving truth are supposed to be proved under similar reasoning.

However, there is one thing not very clear in this account. Halbach does not appeal to a

notion of what Boghossian calls impure analyticity to define truth-analyticity. If the

---

[33] Halbach is well aware of the *de re* character of his formulation and offers a solution to it. Please cf. Halbach 2000, pp. 163-164.

notion of 'truth-analyticity' is defined in virtue of the meaning of the truth predicate, then truth-analyticity cannot be used as an *auxiliary* notion to help axiomatise truth. Similar to Horwich, (except for the differences between the primary truth bearers), Halbach holds that the meaning of the truth predicate is fixed and exhausted by nothing beyond the disquotation feature (Halbach 2000, p. 168). Thus, a sentence is truth-analytic if and only if it is either a disquotation sentence or if it follows logically (not by virtue of the meanings of non-logical vocabulary) from the disquotation sentences (ibid.). Halbach also claims the following:

> According to the view of modalized disquotationalism, being logically provable from the
> disquotation sentences and being [*truth-analytic*] are the same. (p. 168)

So truth-analyticity is conceived in terms of provability. The problem for this account is Halbach's claim that the predicate 'is logically provable from the disquotation sentences', in the sense of classical predicate logic, can be defined within the *base* theory (ibid.). And according to Halbach, a base theory is formulated by the base language consisting of only sentences without the truth predicate (p. 155). Then, how can 'logically provable from the disquotation sentences' be defined in a theory within which disquotation sentences cannot even be formed? I think what Halbach has in mind is that 'logically provable' but not 'logically provable *from disquotation sentences*' can be defined within the base theory. And the predicate of 'logically provable from disquotation sentences' can hence be defined in the *truth* theory.

As I indicated in the last subsection, the original idea of disquotationalism comes from semantic ascent and descent. The truth predicate has its use in helping us express, as well as affirm, infinite lots of sentences, and infinite conjunctions.[34] The issue regarding the

---

[34] Again, the discussion of infinite disjunctions is left out hereafter.

deductive power of the disquotational theory is raised in considering the relation between affirming an infinite conjunction and affirming a corresponding (universal) generalisation. It is easy to reach the conclusion that affirming a universal generalisation implies affirming all of its instances, and through the disquotational schema, all conjuncts of the original infinite conjunction are gained. The other way round is problematic, due to lack of an infinitary rule of inference.

When some infinite conjunctions have their conjuncts with syntactic or proof-theoretic features, it is easiest to see how adding some axioms and/or rules of inference can help *explain* the affirmation of a generalisation from the affirmation of all its instances. The truth theory accompanied with additional axioms and/or rules of inference provides *derivations* of the affirmation of a given generalisation from the affirmation of all its instances. Both Field and Halbach hold that the disquotational theory of truth is not simply an axiomatisation of truth, in terms of the disquotational schema as the axiom schema, but a theory *about* an axiomatic system of truth. Since all the axioms resulting from the instances of the disquotational schema are deductively too weak to account for certain general facts involving truth, more axioms and/or rules of inferences are needed to enhance the deductive power of the truth theory. If the additional axioms and/or rules of inference are truth-free, they are added to other explanatory resources (Field's). If not, they are added to the truth theory (Halbach's). Although both Field and Halbach emphasise that these additional axioms and/or rules of inference are not substantial, it is doubtful that they do not inflate the truth theory in one way or another.

Field's account introduces schematic letters and the corresponding substitution rule into

our natural language, and first order propositional logic is thus absorbed into the language we speak as a part of it.[35] In addition to attributing schematic reasoning to ordinary people, one striking consequence is that infinite conjunctions with certain synthetic features are *all* expressible. The function of the disquotational truth predicate of expressing some infinite conjunctions vanishes. The use of the truth predicate is to properly form disquotational truth conditions, and a side interest is to distinguish true generalisations from false ones. Another striking consequence is that, it seems, some model-theoretic conceptions are built-in the language spoken by ordinary people. One can derive that, for all sentences, 'If S, then S' is true, where the schematic letter 'S' ranges substitutionally over all sentences of the language.[36] In other words, that the schema 'If p, then p' is logically valid can be proved in principle by ordinary people who speak the natural language they speak. For Halbach, on the other hand, the disquotational truth predicate cannot play the most important role of expressing and affirming infinite conjunctions, so the notion of truth-analyticity is indispensable for this purpose. It seems that the notion of disquotational truth more or less fades out at this stage. More importantly, in order for the truth predicate to commute with objectual quantifiers, some intensional predicates of other semantic concepts are needed. The more we need, the less important the truth predicate is. Both accounts complicate the truth theory in order to establish the soundness of first order logic. But the original task is only to establish the claim that, by justifying that one can affirm an infinite lot of sentences by affirming the corresponding generalisation, infinite conjunctions can be expressed in terms of the truth predicate. Whether or not the price is so high, that the name, if not the spirit, of

---

[35] Field proposes to develop a new kind of logic to solve the liar paradox, so he does not go for a hierarchical system.

[36] In view of Field's quasi-disquotationalism, it seems that consideration regarding possible world semantics is comprised in his account.

deflationism is put in danger, is worth considering.[37] Since neither Quine nor Horwich go for the strategies suggested by Field's account and Halbach's, the discussion will be left here.

Now, let us return to the discussion of Horwich's minimalist explanation, based on the minimal theory of truth (MT), of the general fact mentioned at the beginning of this section. It runs as follows:

1. Logic provides us with facts like [dogs bark & (dogs bark → pigs fly)] → pigs fly, that is, with every fact of the form

   $[p \ \& \ (p \rightarrow q)] \rightarrow q$.

2. Therefore, given MT, we can go on to explain every fact of the form

   $[<p> \text{ is true } \& \ (p \rightarrow q)] \rightarrow <q> \text{ is true}$.

3. But from the nature of implication, we have all instances of

   $(<p> \text{ implies } <q>) \rightarrow (p \rightarrow q)$.

4. Therefore we can explain each fact of the form

   $[<p> \text{ is true } \& <p> \text{ implies } <q>] \rightarrow <q> \text{ is true}$.

5. And therefore, given MT, we get each fact of the form

   $<[<p> \text{ is true } \& <p> \text{ implies } <q>] \rightarrow <q> \text{ is true}> \text{ is true}$.

6. But it is a peculiar property of propositions that any general claim about them—any characterization of *all* propositions—is made true by the infinite set of particular facts associating that characteristic with each individual proposition.

7. Therefore, in light of 5 and 6, we can explain the general fact:

   Every proposition of the form, $<[<p> \text{ is true } \& <p> \text{ implies } <q>] \rightarrow <q> \text{ is true}>$, is

---

[37] Whether or not they make some extra metaphysical and/or epistemological commitments is not able to be discussed here.

true. (Horwich 1998*b*, pp. 21-22)

This explanation is clearly unavailable to the disquotationalist, because of premise 6. This turns out to be insignificant, for Horwich later abandons this line of reasoning. Earlier, the success of this explanation was crucial for Horwich, to demonstrate the deductive power of the minimal theory, its capacity for deriving a universal generalisation from all its instances, and also the fulfilment of the necessary condition for the minimal theory to be an adequate theory of truth.[38]

It appears that because of premise 3, Horwich intends to explain the notion of logical implication. But this makes the whole explanation unsound, for the strong modal relation between $<p>$ and $<q>$ in '$<p>$ implies $<q>$' is missing from premises 2, 3 and premise 4. So it is probably more charitable to take Horwich's argument as being about the notion of material implication—about *modus ponens*. This slip does not affect my criticism of his argument anyway. I shall follow David's observation that this must be a typographical error and hereafter change 3 into the following:

3' ($<p>$ implies $<q>$) $\leftrightarrow$ ($p \rightarrow q$).[39]

The idea behind this explanation is apparently to take the general fact as an infinite conjunction, "'<Snow is white> implies <Grass is green> and <Snow is white> is true, then <Grass is green> is true,' '<Sky is blue> implies <Dogs bark> and <Sky is blue> is true, then <Dogs bark> is true, …'", and then to explain how we can arrive at the conclusion that this conjunction can be expressed in terms of the minimalist truth predicate. And hence *modus ponens*—if one sentence/proposition implies another, and the

---

[38] Premise 6 is new to the second edition of *Truth*, so is 7. It is not difficult to see why Gupta complains in both "A Critique of Deflationism" (1993) and "Minimalism" (1993) that Horwich (in the first edition of *Truth*, 1990) only explains all the instances of a generalisation but not the generalisation itself. I will skip the discussion of the old version.

[39] Please cf. David 2000, note 19.

first one is true, then so is the second—can be explained in terms of the minimal theory along with some truth-free explanatory resources. David's interpretation (2000) is that Horwich's idea of explaining general facts involving truth is to *deduce* these facts from the minimal theory, along with some truth-free resources. This does not seem to be misleading, for it is Horwich himself who advances such an idea in some places.[40] What David wants to show is that the deducibility of these general facts eventually boils down to an extended notion of loose deducibility.

There are several problems for Horwich's argument. The most important one is that, as I discussed in the last chapter, premise 3 will inevitably go infinite at the level of schematic argument. Since we have the problem of being short of an infinitary rule of inference, this is probably the reason for Horwich to provide a meta-argument. Apparently it is the sixth premise that makes Horwich think he successfully proves the deducibility. In the Postscript of Horwich 1998*b*, Horwich says:

> However, it seems to me that in the present case, where the topic is *propositions*, we can find a solution to this problem [That is, there is no logically valid infinitary rule]. For it is plausible to suppose that there is a truth-preserving rule of inference that will take us from a set of premises attributing to each proposition some property, *F*, to the conclusion that all propositions have *F*. No doubt this rule is not *logically* valid, for its reliability hinges not merely on the meanings of the logical constants, but also on the *nature* of propositions. But it is a principle we do find plausible. We commit ourselves to it, implicitly, in moving from the disposition to accept any proposition of the form '*x* is *F*' (where *x* is a proposition) to the conclusion 'All propositions are *F*'. (p. 137, the last *italic* word is mine)

Supposedly, the *disposition* to accept any proposition of the form '*x* is *F*' is something to do with premise 5. The inference rule, which is characterised in 6, is thus to licence us

---

[40] In at least two places he specifically uses 'derive' (Horwich 1998*b*, pp. 11-12) or 'deduce' (pp. 24-25) to account for what he means to say about *explaining* general facts involving truth.

from premise 5, to derive 7. What Horwich does not tell us is what sort of *nature*

possessed by propositions makes us *implicitly* commit to this rule of inference.[41] David,

on the other hand, offers an interpretation of this alleged nature of propositions. He says:

> It must be because [Horwich] thinks that propositions…obey the following principle: *If a*
> *proposition exists in any possible world, then it exists in every possible world*. But this
> principle holds only on some (e.g., Fregean) conceptions of the nature of propositions. (David
> 2000, p. 172)

Russellian propositions are precluded from the minimal theory then, if this is the case.

This contravenes Horwich's assumption that minimalism is compatible with both Fregean

and Russellian propositions, and even compatible with a hybrid kind of these two

(Horwich 1998*b*, p. 17). This also damages his account regarding the relation between

beliefs and propositions (ibid., Question 31). For David, the more serious deficiency of

Horwich's assumption that premise 6 is a rule of inference, to which we are implicitly

commit, is its failure of explaining *how* this envisaged rule might work in the absence of

the assembling of infinitely many premises, or an infinite premise. Furthermore, this

envisaged rule is only supposed to work for some specific objects, namely, propositions.

This, according to David, shows that Horwich implicitly slackens the deducibility in need

of accounting for the deductive power of the minimal theory, to a much looser one.

Consequently, Horwich has to admit that substantive theories of truth can always find

some equally loose deducibility to fit their explanation of their adequacy in accounting

for our acceptance of all instances of the equivalence schema and all general facts

involving truth.

I do not think that Horwich is guilty of *implicitly* loosening the required deducibility.

Because in *Truth*, second edition, Horwich states the following:

---

[41] Armour-Garb expresses the same view in Armour-Garb 2004, note 13.

My inclination is to think that *the principles of deductive logic are neither susceptible of, nor in need of, demonstrative justification*. Instead, I would follow Quine (1953) in supposing that the logical principles relied upon in scientific theorizing are justified in so far as they are members of the simplest way of accommodating experience. I would say, moreover, that the logical principles deployed *outside* science in ordinary life—which may or may not differ from those used in science—are not subject to revision in light of experience. Hence these commitments are a priori; their justification is pragmatic; purely practical goals are furthered by adopting them.[42] (p. 76, the first string of *italics* is mine)

If there is no need to demonstratively justify the principles of deductive logic, within or without science, then what does it mean to say that the minimal theory, along with some truth-free explanatory resources, can *derive* or *deduce* general facts involving truth, including the principles of deductive logic such as the one I discussed?[43] There seems to be an apparent incoherence. The more sympathetic interpretation is that the required deducibility is loosened in the very first place. What Horwich is guilty of is not making this *explicit*.[44]

Why Horwich eventually intends to justify those general facts by appealing to the unique nature of propositions is a separate question. In the cited words two paragraphs back, we can see that Horwich recognises the logical invalidity of the rule of inference characterised as such. He probably thinks, separately from the thought regarding explanation cum derivation/deduction of general facts involving truth, that the principles of deductive logic can be explained in an even looser sense. This shows yet another

---

[42] This is partially the reason why Horwich does not go for the strategies advanced by Field. Moreover, the claim that "the logical principles deployed *outside* science in ordinary life—which may or may not differ from those used in science—are not subject to revision in light of experience" results from the need for squaring this with his use theory of meaning. That is, the demarcation between common-sense logic and various kinds of logical theories.

[43] Apparently it is not just about the principles of deductive logic but also about other general facts, such as that true beliefs about how to attain our goals tend to facilitate success in achieving them. To account for this general fact and not only its instances, it is inevitable that Horwich will appeal to the unique nature of propositions.

[44] A sympathetic guess is that in revising the first edition of *Truth* these things did not coherently come to Horwich's mind.

incoherence in his account. Horwich holds that the minimal theories of utterance truth and belief truth can be constructed, with suitable auxiliary assumptions, from the minimal theory of propositional truth.[45] He says:

> Ordinary language suggests that truth is a property of propositions, and that utterances, beliefs, assertions, etc., inherit their truth-like character from their relationship to propositions. However, the above derivations show that this way of seeing things has no particular explanatory merit. The truth-like conception for each type of entity is equally minimalistic. And by assuming any one of them we can easily derive the others.
> (Horwich 1998*b*, p. 102)

But the minimal theories of utterance truth and belief truth certainly cannot provide the explanation of the general fact of *modus ponens* in the way characterised above, for that explanation depends on the unique nature of propositions. Thus Horwich's minimalist account is incoherent.

There is another notable fact, indicated by Armour-Grab (Armour-Grab 2004, note 13), that makes Horwich's position entirely incoherent. At the very beginning of *Truth*, Question 1, that is, Horwich specifically makes the following claim:

> However this commitment [to the existence of propositions], though controversial and in need of some defence (…), is much less substantial than it might seem at first. For it presupposes very little about the *nature* of propositions. (p. 16)

The reason given by Horwich, to relieve the worry of introducing substantial elements into the minimal theory by committing ourselves to the existence of propositions, is that this commitment presupposes very *little* about the nature of propositions. But premise 6 is important for Horwich to account for the loose deducibility of general facts involving truth. The adequacy of the minimal theory turns out to be absolutely hinging on not just the commitment to the existence of propositions but also on their very *nature*. The

---

[45] Please cf. note 15 of the previous chapter or Horwich 1998*b*, pp. 98-102.

minimal theory is thus not very *minimal*, much less deflationary.

In latter times, Horwich seems to modify his argument, abandoning the idea of appealing to the unique nature of propositions. In "A Defense of Minimalism" (2001), Horwich replies to Gupta and Soames' criticism that minimalism only explains how to reach all instances of a generalisation but not the generalisation itself.[46] Gupta in Gupta 1993*a* and 1993*b* made the criticism that Horwich refers to and Soames in Soames 1997. In Horwich 1998*b*, the argument about *modus ponens* characterised above is supposed to be one of Horwich's most important replies to this criticism. Nevertheless, in Horwich 2001, in accounting for general facts involving truth he substitutes another premise for premise 6 mentioned above. It seems that he withdraws the original strategy in Horwich 1998*b*, though he does not make this explicit, at least not in this article.

The new premise added to truth-free explanatory resources is as follows:

> Whenever someone can establish, for every *F*, that is *G*, *and recognizes that he can do this*, then he will conclude that every *F* is *G*. (Horwich 2001, p. 157)

In explicating this premise, Horwich continues:

> We can establish, for every proposition of the form, $(p \rightarrow p)$, that it is true. Moreover, we can become aware of this ability. For we can see that any of the propositions in question may be designated by some expression of a certain form. And we can also come to see that there is a general strategy which, given any such expression, will establish the sentence that results from writing the expression before "is true". (ibid.)

This mathematical inductive reasoning squares with Horwich's attitude toward the principles of deductive logic, and thus makes his account more coherent. If adding this

---

[46] This is actually part of the criticism. In terms of this part, Gupta and Soames go on to conclude that the equivalence schema does not really suffice to account for *all* uses of the truth predicate. It follows that the meaning of 'true' is not fixed by our inclination to accept all instances of the equivalence schema. However, it is clear that the key issue here is about the explanation of how we move from affirming all instances to affirming the corresponding generalisation.

new premise validates the argument, and accordingly all generalisations involving truth are explained in terms of this line of reasoning, it does not seem that this line of reasoning is not available to disquotationalism.[47] For the significance of this premise does not rely on the *nature* of propositions. And this reasoning is especially well-suited to Quine's thought that truth, even logical truth, hinges on reality; that in each instance of the generalisation, sentences with the form '$p \rightarrow p$' are true, our eyes are on reality—that is, for instance, if snow is white then snow is white. Hence I again fail to find any reason to take minimalism as more desirable than disquotationalism.

There is a possible objection to my claim that this strategy is also available to the disquotationalist. One might cast some doubt on the claim that disquotationalism is entitled to do so, by arguing that what confers the ascription of the truth predicate to an expression is the proposition designated by it. This line of reasoning is not applicable to utterances or sentences, for, without the propositions designated by them, their interpretations can vary from one case to another. Contrast this criticism to my interpretation of disquotationalism, however, and it is easy to see that it is not justified. For in every instance of the disquotational schema, the sentence mentioned and used is with a fixed interpretation.

According to Armour-Grab (2004), however, Horwich's new strategy is in tension with his treatment of the liar paradox. Remember the basic claim of the minimal theory, that by applying the propositional form or, in other words, propositional structure of '$<p>$ is true iff p' to *every* proposition, we yield an axiom. In dealing with the liar paradox, using

---

[47] The soundness of this new premise will not be discussed hereafter, for it is not the main issue.

'#' to abbreviate a liar-paradoxical instance "THE PROPOSITION FORMULATED IN CAPTITAL LETTERS IS NOT TRUE," Horwich says:

> But [to deny classical logic] cuts too deep; [to deny (à la Tarski) that the concept of truth can be coherently applied to propositions, such as <#>,] also smacks of overkill; and [to deny that the sentence in capital letters succeeds in formulating a proposition] goes against the fact that, for any condition *C*, one might happen to believe that the proposition meeting that condition is not true—which…would imply that 'The proposition meeting condition *C* is not true' expresses a proposition…Therefore the only acceptable solution is [to reject certain instances of the equivalence schema]: only certain instances of the equivalence schema are correct.[48] (Horwich 1998*b*, p. 41)

Horwich does not provide any solution in order to bar liar-paradoxical propositions in Horwich 1998*b*. In a footnote (21), he acknowledges what Gupta pointed out to him that the need to restrict the instantiation of the equivalence schema is in tension with the ability of the minimal theory to capture schematic generalisations (p. 42). For not every instance of a given generalisation is entailed by the generalisation. However, Horwich says,

> [S]uch problem cases are few and far between; so the utility of truth as a device of generalization is not substantially impaired by their existence. (Horwich 1998*b*, p. 42)

This only justifies why we should maintain the minimal theory, because of its utility, but does not justify the theory itself. Similarly, the strategy of appealing to the nature of propositions certainly is not helpful either, and the same is true for the new mathematical inductive reasoning.

Normally there are two options to deal with liar-paradoxical expressions. We can follow Tarski's insight and go for a hierarchical structure of languages. In each stage we restrict the target language to consist only of sentences without the truth predicate or without the higher order truth predicate. Or alternatively, we can adopt deviant logic to allow

---

[48] Words in square brackets are Horwich's, in the previous paragraph of my quote.

truth-value gaps into which the liar-paradoxical expressions fall. Both options have their deficiencies. Going hierarchical is probably too artificial. Adopting deviant logic is, as Horwich argues, to cut too deep. Nevertheless, the first option is not available to Horwich, for he has no way to go hierarchical.[49] And there is good reason why he should not, because the meaning of 'true' is then not fixed through our inclination to accept the instances of the equivalence schema. Horwich apparently turns down the second option for a similar reason that this jeopardises his use theory of meaning. Consider his instance of (AND) discussed in the previous chapter. Presumably, the counterparts of all logical connectives can be defined use-theoretically in a similar way. Adopting deviant logic creates problems for this. Notice though, that these two options are open to the disquotationalist. In my formulation of Quine's disquotationalism, I go all the way down in terms of the first option. The second option is also available to Quine, though I think that it is not very Quinean.

We have not discussed Horwich's response yet. The tension is that, in dealing with the generalisation problem, liar-paradoxical propositions cause trouble, and in dealing with the liar paradoxes, the explanation of truth generalisations is incomplete. Presumably, any principle or additional axiom added to the minimal theory has to satisfy both sides that are in tension. That is, by restricting the instantiation of the equivalence schema, the explanation of truth generalisations does not exclude any instance. We can agree with Horwich that the utility of truth as a device of generalisation *is not* rejected by the few and far between cases of the liar paradox, but the minimal theory *is*. One way to do it is to

---

[49] In a footnote (McGrath 1997*b*, p. 74, 10), McGrath suggests that we can distinguishes between all propositions, those with the truth predicate in various ways and those without the truth predicate, and that we can allocate them to different levels. Relative to each level, then, we can develop a theory of propositional truth for that level. It is doubtful that this is doable. But I will ignore this approach.

convert all truth generalisations into axioms of the minimal theory. But this is clearly not acceptable for Horwich for the same reason that he does not go for the two options mentioned above. Another way is to add the term 'non-paradoxical' to all truth generalisations to make them restricted. For instance, for *modus ponens*, if one non-paradoxical sentence/proposition implies another non-paradoxical one, and the first one is true, then so is the second. But according to Armour-Garb (2004, pp. 499-500), not all truth generalisations can be treated in the same way. For instance, true non-paradoxical disjunctions have at least one true non-paradoxical disjunct. Liar-paradoxical propositions can still be one of the disjuncts.[50] As McGrath indicates, that a proposition might generate such a paradox is in some cases a contingent matter (McGrath 1997*b*, p. 74, note 10). It follows that there is no a priori method for identifying all and only those problematic propositions.

We can conclude for this section that the minimal theory has serious problems in dealing with the liar paradox. The consequence is that its explanatory capacity for affirming truth generalisations is compromised. On the other hand, disquotationalism is in a better shape. Unless Horwich's minimalism has a drastic and consistent change in both of the minimal theory of truth and the use theory of meaning, it seems that disquotationalism is a better choice.

3. The Proper Ascription of the Truth Predicate

Many opponents incline to regard deflationary theories of truth as having genuine difficulty in accounting for truth ascription. The issue is about truth ascription to

---

[50] Armour-Garb also gives a good counterexample in the case of blind ascription. Cf. Armour-Garb 2004, pp. 500-501.

non-understood sentences. The difficulty of truth ascription under this condition is divided into four types. They are, non-understood sentences of a given language of a given speaker (NUSL), non-understood sentences of different languages (NUSDL), sentences of the possible extension of a given language that are non-understood now (NUSPEL), and sentences that are in principle non-understandable to human beings (NUSH). Disquotationalism seems to be in more trouble with this difficulty. So I shall begin with a discussion of disquotationalist accounts. However, we will soon discover that this criticism focuses on disquotationalism because of some particular interpretation of this sort of theory.

The criticism that disquotationalism has difficulties in accounting for the truth ascriptions to non-understood sentences, stems from various readings of disquotationalism, mainly based on readings of Field's account, perhaps McGee's too, and something similar. Among others, there are Gupta's sense equivalence reading, Künne's redundantist reading, and David's definitional reading. I already showed in the first section that Gupta's reading results from an incoherent assumption and a misconception of disquotationalism, so in the following only Künne's reading and David's reading will be briefly discussed.

Since these readings are mainly about Field's disquotational theory, I should first elaborate his account a bit. Field claims that the pure disquotational truth predicate should be explained as true-as-one-understands-it in every instance of the disquotational schema, and each sentence on the right hand side should also be explained as what one understands. Hence the truth predicate is only applicable to sentences that a given speaker understands in the pure disquotational sense. Truth ascription to other sentences is

understood in the extended disquotational sense according to which the proper translation of non-understood sentences of a given speaker is needed. Two problems emerge. First, different speakers are inclined to have very different understandings of the truth predicate, and hence their ascriptions of it are different in principle. Second, there seems to be a conspicuous difference between cases of NUSL and those of NUSDL, for it is dubious whether or not the notion of translation can play any role in the cases of NUSL.[51]

Künne raises the following argument against Field's pure disquotationalism. Suppose two people, Brad and Rachel, are talking about a newspaper report, and Brad comments on it by saying, "That is true" while in disagreement with him, Rachel says, "No, it is not true" (Künne 2000, p. 189). The startling consequence of pure disquotationalism is that there is no disagreement between them at all. For the truth predicates used by them are different in principle.[52] Or Brad may say "If the report is true, then the government is in trouble," and Rachel agrees with him by saying "The report is true". Brad then concludes, "The government is in trouble". It seems that they build up an argument based on *modus ponens* (ibid.). But this is hardly the case, for again they use 'true' differently. This criticism is similar to Gupta's:

> [O]ne first acquires the concept 'true-in-my-present-idiolect' and then using it acquires the full-fledged 'true.' The problem of explaining how one gets from 'true-in-my-present-idiolect' to 'true' seems…much harder than that of explaining 'true' using a limited ideology". (Gupta 1993, p. 81, note 48)

Field owes us an explanation of how different speakers of the same language eventually come to use 'true' indifferently.[53]

---

[51] Even if the speaker uses a hybrid language, the distinction between the languages she speaks and those that she does not speak still makes sense. To simplify my discussion, only unilingual cases are in considering.

[52] I modify Künne's argument a bit to suit my argumentation.

[53] I also made a similar objection to Field's account in 1.2. That is, the extended- and quasi-disquotational

Now turn to the second problem. If the truth predicate is primarily applied to sentences that a given speaker understands, then there is a difference between non-understood sentences that belong to the language spoken by her and those that do not. For the latter, the truth ascription to them is explained by Field's extended disquotationalism, according to which a non-understood sentence is disquotationally true only if there is a proper translation between it and a corresponding sentence that the speaker understands. For the former, the notion of translation has no place. It seems that Field has to use the idea of NUSPEL to explain the cases of NUSL. But there is supposed to be a difference between NUSPEL and NUSL. The consequence of pure disquotationalism, then, is the confusion between truth ascription to sentences of a given language that one *does not* understand now, and to sentences of the possible extension of it that one *could not* understand now.

David strongly disagrees with Field's pure disquotationalism. The austere disquotationalism, David so argues, is its lack of reliance on any notion of meaning. However, the notion of understanding is within the same family of notions as meaning, and hence cannot legitimately to be used by a physicalist account of disquotationalism (David 1989, p. 608). For Field, and McGee as well, to reserve the necessary status of instances of the disquotational schema, the price to pay is to introduce the notion of understanding into the account. It might be claimed that 'true-as-one-understands-it' and 'as one understands it' at both sides of pure disquotation sentences somehow cancel each other out—that is, it does not specify *the* understanding of the target sentence. In other

---

truths do not outpace the given speaker's own pure disquotational truth, unless we are given some practical hypotheses such as that the translation/interpretation manuals of people speaking the same language tend to be convergent.

words, the two 'understand' appearing strings of words are only used to restrict the range of substitutional variable '*u*' in "'*u*' is true (as one understands it) if and only if *u* (as one understands it)". According to David, this is only a cheat. Although David's example is his disquotational definition of truth, (D), which I discussed in Chapter Two, his point is also applicable to the above claim.[54] The point David made is that:

> Of course we can keep a definition [or a full explanation of a schema] clean of uncomfortable notions if we only use those notions in an "informal commentary" on the definition [or the schema]; but in this way we only keep the *definition* [or the *schema*] clean and not the *theory* in which the definition [or the schema] is embedded. (David 1989, p. 608)

As I indicated earlier in 1.2, Field's pure disquotationalism is inadequate in being implicitly committed to the sense equivalence relation. Conversely, this criticism is not applicable to (D$_Q$), for the notion of interpretation involved is not about the sentences appearing at the two sides of the biconditionals. If one has any problems with it, one should equally object to Tarski's formulation of Convention T.

Furthermore, Shapiro raises an argument against Field's disquotational account in Shapiro 2003. Here is the background case on which the argument depends. A guru makes a pronouncement about set theory; a disciple thinks that everything the guru says is true; the disciple does not understand set theory; a logician who distrusts the guru's set-theoretic pronouncements would like to draw their number-theoretic consequences, which the disciple does understand.[55] Shapiro argues that if in addition the disciple trusts the logician's acumen about logic (though not about set theory), then the disciple ought to be able to reason in terms of the notion of logical consequence, that the number-theoretic

---

[54] The definition is as follows:
  (D) *x* is a true sentence $=_{DF} (\sum p)(x$ is 'p', & p) (David 1989, p. 604; David 1994, p. 89) .
[55] This is actually Field's characterisation of Shapiro's case in Postscript (2001) of Field 1994. When Field made this characterisation, Shapiro's paper was not yet in publication.

consequence of what the guru says is true. The upshot of Shapiro's argument is that, though the disciple understands the conclusion, the reasoning is frustrated by the pure disquotational fact that the disciple cannot meaningfully apply 'true' to the guru's utterances. In considering how to extend the truth schema as we expand the language, Field concedes that "once we come to regard an expression S as an acceptable declarative sentence, then even if we have no understanding of it (or virtually none) still we ought to accept the corresponding instance of the truth schema" (Postscript (2001) of Field 1994, p. 147).[56] Field also acknowledges that, according to the cognitive equivalence thesis of pure disquotationalism, truth attribution to an expression not yet understood is out of the question. In order to settle this conflict, he proposes to think of this matter in terms of indeterminacy—that is,

> [A] sentence I don't understand is maximally indeterminate in content for me, and so is
> an attribution of truth to it; but the indeterminacies are 'tied together'. (ibid.)

According to Field, this modified view handles Shapiro's case neatly. The disciple regards the guru's sentences as "in a potential expansion of the disciple's language, so that there is no difficulty in his carrying out the reasoning" (ibid., p. 148).[57] In addition, Field argues that the guru problem is also a prima-facie problem for minimalism (ibid.). For the synonymous relation between the disciple's expressions and the guru's is unavailable to a minimalist. Therefore, Horwich also has to provide an explanation of how the disciple's reasoning is possible.

It does seem to me that Field gets the wrong idea of Shapiro's criticism. From my

---

[56] It seems to me that, in the case of language expansion, the schematic reasoning introduced by Field to natural language is capable of accommodating the acceptance of the corresponding instances of the truth schema. However, as a full-blooded physicalist, Field still relies on computational psychology to explicate his pure disquotationalism.

[57] Field indicates that his reply still holds if the guru's language diverges from the disciple's. In this case, the expansion of the disciple's language is to translate the guru's utterances.

formulation of disquotationalism, mainly with respect to the schema (D$_Q$), it is not hard to

find an explanation based on translation or possible expansion of the language to

accommodate this problem. My theory does not hinge its interpretation of the

corresponding schema on the notion of understanding. It is, then, a practical question

whether or not the disciple in Shapiro's case can reason in terms of his assumption that

everything the guru says is true, in terms of the logician's acumen in logic, and the notion

of logical consequence. It depends on the justification of the disciple's *faith* in the guru's

utterances, as well as in the logician's acumen in logic. Let us suppose that the logician is

trustworthy. If the part of everything the guru says that the disciple understands comes out

true, then the disciple at least is at a departure point for some inductive reasoning. If the

remaining part also comes out true after the disciple expands his language or translates

those utterances into his language, then the disciple's faith is fully justified. However,

Field is never in a position to regard Shapiro's case as a practical question, for the

disciple's whole reasoning is blocked in the first place. Because of this, Field has to

appeal to indeterminacy to respond, cumbersomely perhaps, to Shapiro's criticism,

whereas my disquotational theory does not have to.

Regarding Field's criticism to Horwich, I think that he makes a good point. Similar to

Field, the disciple's reasoning is blocked at the very beginning according to the minimal

theory of propositional truth. Shapiro criticises Horwich's minimalism similarly. The

disciple is unable to understand the corresponding instances, about what the guru says, of

the equivalence schema. He can only express it vacuously in terms of his faith in the guru

and the logician who tells him that the set-theoretic utterance of the guru's is a legitimate

declarative sentence. Furthermore, it is dubious why the disciple should believe in the

notion of logical consequence, for he only has instances of what he understands. In other words, Horwich's mathematical inductive reasoning has its effectiveness only for sentences that one understands. For sentences that one does not understand, there are no corresponding propositions to be grasped.[58]

It is the very *nature* of propositions, it seems to me, that makes the disciple's reasoning fail. That is, one does not know what proposition is expressed if one does not understand the expression (does not know the meaning of it). The fact that the disciple does not understand the guru's utterances implies that the disciple does not know what propositions are expressed by those utterances. Consequently, the disciple cannot apply the minimalist truth predicate to the guru's words. In other words, there is a practical problem of truth ascription for the minimal theory of propositional truth in the cases of NUSL. This also affects the explanation of general facts involving truth. The infinite conjunctions or disjunctions are not fully captured by attaching the truth predicate to generalisations, for the propositions not known by a speaker should not appear in the instances of the generalisations. The possible expansion of language does not much help here. Based on the assumption, or perhaps, as Horwich would say, common sense view, that for any declarative *sentence* there is a corresponding proposition, Horwich formulates his equivalence schema. For sentences that a given speaker does not understand, the function of the minimalist truth predicate that plays the role of forming generalisations, is only fulfilled by accepting the aforementioned assumption that those non-understood sentences express propositions because of their declarative feature.[59] In other words, the

---

[58] Please see the following criticism to Horwich.
[59] This assumption is not that unimportant. It is one of the main reasons why Horwich rejects moral expressivism. With respect to this, I am with Michael Smith (1994*a* and 1994*b*)in holding that deflationary theories of truth should not take any stand on issues like this.

minimalist truth ascription to many instances of a given generalisation is founded on our implicit, if not explicit, acceptance of (D$_Q$). The function of the minimalist truth predicate only makes sense when (D$_Q$) is presupposed. Therefore, the very reasoning of Horwich's, which gives the primary position to the minimal theory of propositional truth, is in doubt.

In dealing with propositions not yet expressible, Horwich refers to a possible extension of English to explain. He says:

> [W]e can characterize the 'equivalence axioms' for unformulatable propositions by considering what would result if we *could* formulate them and *could* instantiate those formulations in our equivalence schema. Thus we may specify the axioms of the theory of truth as what are expressed when the schema
>
> (E) '<*p*> is true iff *p*'
>
> is instantiated by sentences in any possible extension of English.
> Alternatively, instead of identifying the axioms indirectly in terms of how they would be *expressed*, we can solve the problem by directly specifying the *propositional* structure which all and only the axioms have in common. (Horwich 1998*b*, p. 19, note 3)

What we do not know, according to this passage, is whether or not the minimalist truth predicate is applicable to propositions such as in the cases of NUSH. The option for the disquotationalist is clear. There is no intelligent way to apply the disquotational truth predicate to expressions (of some languages) that are in principle not understandable to us. Horwich, on the other hand, has no choice but to accept the very controversial assumption that we human beings are capable *in principle* of understanding all possible propositions. Again, there is no reason, as far as I can see, to give the minimal theory of propositional truth a privileged position.

In terms of the first three sections, there is little difference between my formulation of disquotationalism and Horwich's minimalism, except, of course, minimalism is a theory

212

of propositional truth. If we take the equivalence schema not to apply to propositions but

sentences, and providing that those conditions set up by Horwich for a proper

understanding of the equivalence schema are satisfied, then we can hardly find any

difference between (D_Q) and (E). However, according to the analysis in this and the

second section, there are some very serious deficiencies in the minimal theory of

propositional truth. So we may say that disquotationalism seems to be preferable.

4. Is Quine a Disquotationalist?

Whether or not Quine is an advocate of disquotationalism can be discussed in two parts.

One is about some inconsistent passages about which theory of truth he is in favour of.

For instance, in Quine 1995 he says:

> Science is seen as pursuing and discovering truth rather than as decreeing it. Such is the
> idiom of *realism*, and it is integral to the semantics of the predicate 'true'.[60] (p. 67,
> emphasis mine)

The other part is about his theory on meaning. For instance, in Quine 1981*b* he states the

following:

> "It was left to Davidson to recognize Tarski's theory of truth as the very structure of a
> theory of meaning" (p. 38).

We will first spend some space to settle the former.

According to Davidson, Quine is among those who are in favour of deflationism. This

seems to be correct, for Quine does urge that disquotationalism is all there is to say about

truth. For instance, in Quine 1990*b* he says the following:

> [T]here is surely on impugning the disquotation account; no disputing that 'Snow is

---

[60] Davidson suggests that this passage seems to suggest a non-deflationary position. This is probably not
the case. Considering Grover's viewpoint (Grover 2000) that the importance of scientific pursuit is to
search for what-is-true, according to his terms, the *italic* word two sentences back can be perfectly
explained as the collection of all what-is-trues in terms of disquotationalism.

white' is true if and only if snow is white. Moreover, it is a full account: it explicates clearly the truth or falsity of every clear sentence. It is even a more than full account: it imposes a requirement on the truth predicate that is too strong for any predicate within the language concerned—on pain of contradiction. (p. 93)

Of course, there is also an almost indisputable passage: "Truth is disquotation" (p.80).

However, there are also unmistakeable passages suggesting that we should not interpret Quine's disquotationalism as deflationary. There is a strong flavour of the coherence theory of truth in the following passage:

> It is rather when we turn back into the midst of an actually present theory, at least hypothetically accepted, that we can and do speak sensibly of this and that sentence as true. (Quine 1960, p. 24)
> Where it makes sense to apply 'true' is to a sentence couched in the terms of a given theory and seen from *within* the theory, complete with its posited reality." (ibid., emphasis mine).

Then again, Quine seems to be in support of the correspondence theory as well. In Quine1990*b* he says:

> Yet there is some underlying validity to the correspondence theory of truth, as Tarski has taught us. Instead of saying that 'Snow is white' is true if and only if it is a fact that snow is white, we can simply delete 'it is a fact that' as vacuous, and therewith *facts* themselves: 'Snow is white' is true if and only if snow is white. To ascribe truth to the sentence is to ascribe whiteness to snow; such is the *correspondence*, in this example. (p. 80, emphasis mine)

So far as it is holophrastic, for Quine, to say that 'Snow is white' is true is in fact to ascribe whiteness to snow. This ascription conversely does not rely on the analysis of sentential structure. The connection between our understanding of the world and the world itself is through our sense experience. The empirical slack between the meagre sensory input and the torrential output allows us the latitude of ontological relativity to vary the reference schemes of a given theory while preserving the distribution of truths

(or, in other words, the sentence-to-sentence structure).

Quine seems to vacillate between the coherence theory and the correspondence theory. The inclination to a coherence theory of truth, according to Davidson, is a mistake. This inclination is shown in Quine's indecision between his *sectarian* and *ecumenical* positions towards choosing a true theory among empirically equivalent ones (Quine 1986). Some citations might help. For sectarian position, Quine states:

> One of the two systems of the world must be deemed false even if we know them to be empirically equivalent. My reason for it was naturalism. (Quine 1986, p. 156)

For ecumenical position, he states:

> I opted for truth of both systems of the world, finding it offensive to my empiricist sensibilities to declare otherwise. (ibid.)

The difficulty with the ecumenical position is as follows:

> Two empirically equivalent systems of the world may be *logically incompatible*, and hence incapable of being simultaneously viewed as true. (ibid.)

After performing Davidson's expedient, that "the incompatibility is resoluble simply by re-construing some theoretical term in that sentence as a pair of distinct homonyms" (ibid.), the ecumenical position does not have this difficulty since empirically equivalent systems of the world might be logically non-equivalent but never in conflict.

Davidson does not accept Quine's indecision. For him, the sectarian position is completely flawed. It is hard to see how a sentence, with interpretation unchanged, can be true for one person but not another or can be true for a person at one time but not at another. "This difficulty," Davidson argues, "seems due to the attempt to import epistemological consideration into the concept of truth" (Davidson 1990, p. 306). And for Davidson the import of epistemological consideration means to follow the approach of a

coherence theory to account for the concept of truth.[61]

We also have good reason to reject Quine's inclination to a correspondence theory. Even if Quine considers his theory of truth as some sort of correspondence theory, his theory is too different from the mainstream theories considered in Chapter Two to be included as one of them. The concept of truth does have a close relation to reality. As I said before, truth is about reality, about what there is. This is what semantic ascent is about. It is by semantic ascent that we can substitute the talk of sentences for the talk of reality. Thus to say that "Snow is white" is true is to say that snow is in reality white. This seems to be the reason Quine claims that his theory of truth is a kind of correspondence theory (Quine 1990*b*, p. 80). As long as we do not possess some sort of telepathic communicative ability that does not involve any linguistic forms whatsoever, we need linguistic expressions to help us articulate facts. Whether or not we can specify a fact without any linguistic form involved is less important here. What is important is that we have to rely on linguistic forms to specify facts in all kinds of communications. Hence Quine's point can be illustrated as the claim that any given sentence of a target language is true if and only if this sentence articulates a fact.

According to most realists, truths as well as facts should be independent from scientific theories. However, Quine's ideas of truth and fact do not respect this view. Our account of what reality should be conceived of, according to Quine's holism and thesis of under-determination, makes no absolute sense.[62] It depends on how we construct

---

[61] According to Colin McGinn, disquotationalism can be used to reject the classic coherence theory of truth. Please cf. McGinn 2000, pp. 194-195.
[62] The thesis of under-determination is that theories or theoretical formulations can be empirically equivalent—that is, they imply the same set of empirical consequences—while being logically incompatible

hypotheses to interpret pieces of empirical evidence. According to holism, we have the

latitude to modify our theory, to redistribute truth-values over some sentences of our

theory, even logical truths, when a conflict with experience happens.[63] And the

modification or redistribution of truth-values can occur in more than one way. It seems

that we have different ways to capture reality. The thesis of under-determination pushes

this to an extreme degree, for it implies that all possible observations fail to determine the

unique nature of reality.[64] Therefore, what sorts of facts we have are relative to the theory

we hold. Otherwise, logically incompatible theories that account for all possible

observations are unintelligibly in support of the same set of facts. So, different

comprehensive scientific theories claim different sets of facts. It follows that talking about

the correspondence relation in an absolute sense is pointless. Hence Quine's

correspondence theory is not the normal kind in which most realists are interested. Even

---

with each other in a non-trivial sense. The logical incompatibility of two theoretical formulations might simply be that they have a different usage for two theoretical terms, say, 'electron' and 'molecule', which do not appear in the linguistic vehicles of their empirical consequences. What are said in theoretical formulation T about 'electron' is about 'molecule' in another formulation U, and similarly about the word 'molecule'. The seemingly logical incompatibility coming from the divergent usage of terms, for Quine, is superficial and can be dispensed with by *reconstrual of predicates*. What 'reconstrual of predicates' means is 'any mapping of our lexicon of predicates into open sentences' (Quine 1975, p. 320).

[63] One might argue that logical truths are different from other sentences in the sense that they are devoid of content, or that their truth has nothing to do with their content. The disputable part is whether logic is a compendium of the broadest traits of reality or it is just an effect of linguistic convention (Quine 1970, p. 96). "Must all right-minded men agree on logic, or is it every language for itself?" (ibid.) In translating a remote language, however, the logic of the translated language is at worst incommensurable with ours but never in conflict, for the conflict only discredits our translation. Following this line, it seems that logic is naturally an effect of linguistic convention. But the truth predicate plus some other Quinean theses—that is, holism, indeterminacy of translation, and inscrutability of reference—separate logic from language. "The truth predicate serves the crucial purpose, in oblique generalisation, of disquotation. Logical theory, despite its heavy dependence on talk of language, is already world-oriented rather than language-oriented; and the truth predicate makes it so." (ibid., p. 97)

[64] Gilbert Harman in "Meaning and Theory" suggests that there are two notions of empirical equivalence. The first one is that two theories are empirically equivalent if and only if they are *equally supported by* all observational evidence. The second is that two theories are empirically equivalent if and only if they are *compatible with* all observational evidence. In "On Popper's Negative Methodology"(Quine 1974*a*), Quine argues that there is no confirmational support of natural laws or the like but only non-disconfirmational support of them. (cf. ibid., pp. 218-20) There is no observational evidence that can confirm a hypothesis or a law or the like — pieces of evidence only support them in the sense that they do not refute them. In other words, what the pieces of observational evidence support is the compatibility between a theory and observations. Therefore, there is no genuine distinction between the concepts of 'equally supported by' and 'compatible with' for Quine.

though Quine seems to express his preference for realism in Quine 1995, his account of the correspondence theory has no attraction for other realists. The resultant account of realism that Quine supports is hardly the realism that we are familiar with.

Quine's inclination should not be about the correspondence theory of truth, but rather about a *correspondence intuition* regarding truth. In considering the passage cited from Quine 1990*b* (p. 93), this is apparent in the following words quoted from "Comment on Lauener" (Quine 1990*a*):

> I keep to the correspondence theory of truth, but only holophrastically: it resolves out into Tarski's disquotational version of truth rather than a correspondence of words to objects.
> (p. 229)

The correspondence intuition can be phrased neatly as follows: the sentence mentioned and used in all instances of (D$_Q$) is exactly the same sentence with the same interpretation. In other words, when one uses a sentence, for instance, "snow is white" to make an assertion about the reality, one can equally assert that "Snow is white" is true.[65]

Now let us move onto the second part of this section. My formulation of disquotationalism, based on (D$_Q$), as I indicated above, separates the theory of meaning nicely from the theory of truth. If we have a theory of meaning as strong as Horwich's,

---

[65] McGinn in McGinn 2000 adheres to what he calls *thick* disquotationalism, according to which the concept of truth is simple, without empirical essence or nature, unanalysable, but *substantial* (p. 202). Truth is a robust property and is disquotationally definable because it is the one and only *self-effacing* property whose application conditions can be stated without making reference whatsoever to itself (p, 198). Moreover, truth is essentially a device of inference; by learning that a certain proposition is true one can infer a fact of the world in terms of making a transition from the truth of the proposition to that very fact. Disquotation sentences accordingly are read as follows: from left to right, they are read as reality-implying—that is, by knowing that truth applies to a proposition one comes to know *facts about the world* (p. 199); from right to left, the truths are supervenient on the facts but do not collapse into facts and propositions, since it is an irreducible property (pp. 205-206). Except for seeing truth as a robust property, almost all the features above can be explicated in terms of my interpretation of (D$_Q$). What thick disquotationalism does not tell us, is how to explain the feature of semantic ascent. How can such a substantial truth predicate form generalisations by which those corresponding infinite conjunctions and disjunctions are expressed?

then in the cases of NUSDL the disquotational truth predicate is equally applicable to sentences of other languages. Even under this consideration, disquotationalism is still more preferable than minimalism according to the analyses in the last two sections. If such a strong theory of meaning is not available, then truth ascription varies from one translation/interpretation manual to another. Disquotationalism under this condition is still in a better shape, for minimalism inevitably collapses. The genuine problem of my formulation of Quine's disquotationalism emerges from his theory of meaning, from his inclination to accept truth conditional semantics.

For instance, in Quine 1974*b*, he says:

> [I]n learning language, we are learning how to distribute truth values. I am with Davidson here; we are learning truth conditions. (p. 65)

It is certainly tempting to go after Michael Williams' view in "Meaning and Deflationary Truth" (Williams 1999) that, providing Field's disquotational theory as the sample, the concept of truth need not play an *explanatory* but only an *expressive* role in the major parts of a Davidsonian theory of meaning. For example, the imputation of logical form to alien utterances, ascribing semantic relations of reference and satisfaction to nonlogical, subsentential components of utterances, and most importantly the principle of charity.[66] Establishing this view, a Davidsonian theory of meaning can join the deflationary club. However, according to Jeffrey Hershfield and Deborah Hansen Soles in "Reinflating Truth as an Explanatory Concept" (2003), this view is totally a misconception of a Davidsonian theory of meaning.

---

[66] Also cf. "On Some Critics of Deflationism" (2002), especially p. 156 in which Williams argues that the role of 'true' played in the truth conditions of observation sentences is just generalising and expressing.

I merely indicate two points made by Hershfield and Soles, which mean that any attempt

to deflationise Davidsonian theory of meaning is doomed. The first point regards

Williams's claim that reading first order logic into alien utterances is *neutral* with respect

to the issue of whether truth is deflationary, and the point runs as follows:

> Williams seemingly overlooks the ontological ramifications that come along with
> Davidson's application of a Tarski-style theory.
> In particular, [Williams] seems to ignore the fact that the assignment of first order logical
> form, with a standard objectual interpretation of the quantifiers, carries certain
> ontological import: using such a theory to generate truth conditions for utterances of a
> speaker is, among other things, to impute to that speaker beliefs about the existence (or
> non-existence) of various (sorts of) things. (Hershfield & Soles 2003, p. 36)

So the imputation of logic into alien utterances is to commit an alien informant to

something ontologically to which a deflationary concept of truth cannot do. The second

point is about the principle of charity, which Williams characterises as the maximisation

of truth and observational adequacy:

> Davidson's principle of charity is a precondition of any interpretation at all. As such it is
> part of a rationalizing enterprise which involves attributions of belief (and meaning and
> desire). The attribution of contentful belief can and will involve attributions of what the
> interpretee counts as evidence; it will also involve occasional attributions of erroneous
> belief, that is, the attribution of some belief to the interpretee which the interpreter does
> not hold true but for which the interpreter can find an explanation for the fact that the
> interpretee does hold it true. But error presupposes an explanatory concept of truth, for it
> represents the case where what is held true comes apart from what is actually true.[67]
> (ibid., p. 37)

It does not seem to make any sense, in both of the maximisation of truth and

observational adequacy, to base Field's extended disquotationalism to use the mapping

relation between one's own utterances and alien ones, to explain Davidsonian principle of

---

[67] In Horwich 1998*a*, Horwich claims that, except the divergence in the explanatory order, "Davidson's principle of charity arguably boils down to [his] use theory of meaning" (p. 72). Considering the possibility of the ascription of error belief, either Horwich's interpretation of this important principle of Davidsonian theory fails, or he implicitly employs truth substantively as an explanatory concept. In either way, his objections to Davidson's theory are significantly weakened.

charity.

Final remarks

At the end of the day, I do not look at any possible reconciliation with sanguinity. There are only two options left. One is to go all along the line of disquotationalism and to abandon any idea of accepting the truth conditional theory of meaning. Or we can follow Davidson in seeing truth as a substantive primitive, and hence depart from disquotational reasoning. So we may come to the conclusion that either one should adopt truth conditional semantics or, if one tends toward deflationism, one should choose the Quinean disquotational theory based on (D_Q) over Horwich's minimalism.

Conclusion

At the beginning of the Introduction, I listed some issues regarding deflationism that had yet to be clarified. These were:

1. What is a proper formulation of a disquotational schema of truth?

2. What is the modal status of disquotation sentences or the axioms of the minimal theory of proposition truth?

3. How sensible are the presuppositions involved in asking these questions about the modal status of sentences or axioms?

4. How should we give the deflationary theory of truth the deductive power it needs so as so as to adequately explain all general facts involving truth?

5. What is the deflationary conception of reference? Is it feasible to explicate the deflationary conception of truth in terms of the disquotational concepts of reference and satisfaction?

6. What is an appropriate account of propositions for use in a minimalist account of truth?

7. What sort of equivalence relation occurs within disquotational sentences or minimalist axioms of truth?

8. Between disquotationalism, taking sentences as the primary truth bearers, and minimalism, taking propositions as the primary truth bears, which is the more promising approach to a deflationary theory of truth?

Hopefully this dissertation has made some contribution to a more appropriate

understanding of disquotationalism and minimalism. The first question is of course a very general one, which is answered throughout the work. The second question has been answered by the way in which I address many criticisms of disquotationalism, and some of minimalism as well, which stem from a misinterpretation of the disquotational schema of truth. For instance, David, Künne, and many others, criticise disquotationalism for falling short of accounting for the modal status of disquotation sentences. According to McGrath, this is also a problem for Horwich's minimalism. In the second section of Chapter Two I showed that this is the consequence of misinterpreting the disquotational schema of truth as ascribing the truth predicate to sentence types disregarding what contents they have. This sort of misinterpretation turns the whole reasoning of disquotationalism and minimalism into a grotesquery. Further, in Chapter Five I argued that the necessary status of disquotation sentences is not indispensable when explaining all general facts regarding truth—this of course also addresses the fourth question above. If there is any reason for disquotation sentences as well as the axioms of the minimal theory to be necessary, it must be because of something else.

The seventh question is dealt with in part by my discussion of Gupta's criticism of disquotationalism and minimalism that the equivalence relation within any instance of the disquotational schema, as well as any instance of the equivalence schema, has to be sense equivalence in order to maintain the expressing feature of the truth predicate. In Chapter Five, I showed that the sense equivalence reading does not concern anyone as long as a suitable explanation of the reasoning of infinitary inference is given. And this reasoning in turn was discussed in the second section of Chapter Five.

We can see that these criticisms can be answered by a proper understanding and a proper formulation of the disquotational schema. Following the clarification of Quine's reasoning on the disquotational feature of the truth predicate, namely, semantic ascent and descent, my formulation of the disquotational schema within the truth language relative to a target language is given in Chapter Two. This formulation addresses the first highly general question in the list. The schema is:

(D$_Q$) 'p' is true iff p,

    where the first 'p' is replaced by any sentence of the target language and the second 'p' by the same sentence with the same interpretation.

Further in Chapter Five, I argued that, in terms of this formulation, the sense equivalence as well as the necessary equivalence relations are not necessary in order for the truth predicate to express infinite conjunctions. The material equivalence relation is enough for the truth predicate to do this, providing that, of course, a suitable account of the reasoning of infinitary inference is available. Thus the second, third, and seventh questions from the list are addressed by my reply to objections raised by David, Künne, and many others, which do not arise according to the proper understanding of (D$_Q$). Granted, the material equivalence relation of the disquotational schema, (D$_Q$), is a bit stronger than the usual one, for the sentences mentioned and used in all disquotation sentences are restricted to the same sentence with the same interpretation. This is precisely why Quine takes truth as *immanent* to home languages, for Tarski's T-sentences are not automatically instances of (D$_Q$). The expressibility of infinite conjunctions in terms of the corresponding generalisations depends on both the material equivalence relation and the additional restriction about the sentence mentioned and used in every instance of the generalisations. However, this restriction does not yield any modal relation to the schema and its

instances.

I go some way towards dealing with question five in my discussion of the disquotational conception of reference. Along with the disquotational feature of the truth predicate, the disquotationalist and the minimalist normally adopt a disquotational theory of reference, considering the disquotational schema of reference and all its instances to provide a genuine reference scheme for the target language. Some disquotationalists, such as Halbach and McGee, even claim that the disquotational concept of truth can be explicated in terms of the disquotational concepts of reference and satisfaction. I showed in Chapter Five that McGee's account is based on a misunderstanding of Quine's indeterminacy thesis. Halbach's account, on the other hand, is different. In order to enhance the deductive power of the disquotational theory of truth, Halbach proposes to explicate the concept of truth in terms of the inductive axioms of reference and satisfaction. In Chapter Three, I argued that both claims cannot be accepted by Quine.

Question five is also dealt with in much of the second and third chapters—particularly with reference to Halbach's proposal. In those chapters I showed his proposal faces a problem of the explanatory order, as well as a problem of falling short of implying all disquotation sentences of the target language. If the disquotational schema of reference is considered as providing a genuine reference scheme, then Quine's theses of inscrutability of reference and ontological relativity are jeopardised. For Quine's maneuvre, which ends the infinite relativisation of ontology in practice by acquiescing in our own language and taking its words at face value, amounts to the claim that the relativised explication of any given reference scheme of other languages or theories has to rely on unrelativised

semantic notions applied to our own language. In additional to this, I showed that the disquotational schema of reference and all its instances, if taken as providing a genuine reference scheme, cannot account for all the facts involving reference. For instance, consider the hidden indeterminacy of reference in pre-Newtonian formulation of physics. I concluded that a disquotational account of reference, such as Horwich's, is too deflationary to account for referential indeterminacy. A disquotational schema of reference is useful in scientific reasoning, but serious problems emerge when it is seen as providing a genuine reference scheme.

I return now to how I have dealt with the fourth question. One of the most important tasks of disquotationalism and minimalism is to explain all general facts involving truth. Therefore, it is important for disquotationalism and minimalism to show that they have enough deductive power to explain why infinite conjunctions or disjunctions can be expressed in terms of the truth predicate. I found, in Chapter Five, that Horwich's account of explaining general facts involving truth (in Horwich 1998*b*) is incoherent. On the one hand, he claims that the minimal theory of truth does not become substantial because of its commitment to the existence of propositions. For it presupposes very little about the nature of propositions. On the other hand, the crucial arguments for the capacity of the minimal theory of truth to explain all general facts involving truth depend heavily for their validity on the very nature of propositions. Moreover, Horwich claims that he follows Quine in supposing that the logical principles relied upon in scientific theorising are justified in so far as they are members of the simplest way of accommodating experience. Therefore, the principles of deductive logic, according to Horwich, are neither susceptible to, nor in need of, demonstrative justification. This, however, indicates

another incoherence. If there is no need to justify the principles of deductive logic, such as *modus ponens*, why does Horwich appeal to the nature of propositions to prove that the generalisation regarding the principle of *modus ponens* can be derived from all its instances?

Later, Horwich abandons appealing to the nature of propositions, and turns to mathematical inductive reasoning. That is, whenever someone can establish, for every *F*, that is *G*, *and recognises that he can do this*, then he will conclude that every *F* is *G*. This reasoning, however, generates problems regarding the liar paradox for minimalism. In sum, either the truth predicate does not apply to every proposition, or the applications produce paradoxes, but there is no way that Horwich's account can get away from this.

Chapter four contains much of my discussion of the sixth question. In this chapter I consider Horwich's claim, following his illustration of what word meanings and sentential meanings are, that Quine's indeterminacy thesis does not successfully repudiate meaning entities. He argues that this is because reference is not inscrutable. Hence, he argues, Quine's proxy function argument does not show that there are only adequate translation manuals, but not a correct one. In discussing Horwich's QQQ account, I first indicated that his formulation of the alleged argument of Quine is actually not Quine's. The emphasis of Quine's indeterminacy thesis is inaccurately boiled down, by Horwich, to the inscrutability of reference. Probably because of this, the importance of the proxy function argument is overestimated by Horwich in evaluating Quine's indeterminacy thesis. I criticised Horwich's disquotational theory of reference in Chapter Three. What Horwich fails to understand about Quine's indeterminacy thesis, is that in the case of radical

translation, with regard to the totality of her informants' linguistic dispositions, our field linguist does not have the luxury to hope for any evidence beyond this. Even if we get the totality of assertability conditions of our informants, we cannot establish semantic facts of the translated language. True, if the proxy function argument is valid, then reference is inscrutable and the ontology of a given language or a given theory has to be relativised to be comprehended. Hence translation of another language is indeterminate. In contrast to Horwich's overestimation of the importance of the proxy function argument, however, the indeterminacy thesis does not hinge its significance on the validity of this argument. The notion of proxy function is used to make a point *regarding* semantical reasoning. The notion itself is not a semantic notion. To actually infer a reference relation from its proxy or the other way round is irrelevant and incidental.

I also discussed Horwich's use-theoretic account of meaning in Chapter Four. Based on his view according to which the meaning property of a word is metaphysically indistinguishable to its use property, Horwich arrives at the conclusion that each word has its (explanatorily basic) use property, according to which the overall use of a word is best explained. A word of a language is synonymous with a word of another language so long as they have exactly the same use property. In terms of his account of compositionality, applying a schema with a schematic property to the constituents of the sentence in an exact order, one gets a construction property for each sentence. Similarly, two sentences of different languages are synonymous if and only if they have the same construction property. Consequently, a proposition can be explicated as the construction property of the sentence expressing it. Therefore, a successful formulation of a use-theoretic account of propositions makes a minimal theory of propositional truth more materialised.

The first problem of Horwich's use-theoretic account of meaning is that the explanatorily basic use property of a word depends on the distinction between the word's basic use and the rest. Horwich thinks that a use property of the basic use of a word is explanatorily basic if and only if it can best explain the not basic use of the word. However, he does not give us a satisfactory account of how to distinguish the basic use of a word from the rest of its use. It is very likely that this distinction eventually boils down to a contingent fact that can be known only through an empirical or scientific inquiry. This implies that there might be equally explanatorily basic use properties of a word, due to the contingency of the distinction. Secondly and more importantly, Horwich's entire use-theoretic account of meaning depends for its plausibility on some contingent use facts of a language, in his case, English. He seems to assume that different languages all share the total contingent use facts of English. This is hardly acceptable. It is more apparent in his account of compositionality that, according to him, different languages share the same set of synthetic schemata. The fundamental difficulty for Horwich's use-theoretic account of word meaning and compositionality is its fine-grained character. The particular semantic and synthetic features of foreign words are finely ground into meaningless pieces by Horwich's mill.

The fourth and eighth questions are answered by how I show that the deficiencies in Horwich's use-theoretic account of meaning and compositionality suggest a missing link in his enterprise of deflationary semantics. Regardless of whether or not Horwich provides an acceptable account of propositions, the minimal theory of propositional truth is unable to explain all ascriptions of the truth predicate. In Chapter Five, in discussing

the proper applications of the truth predicate in the disquotational or the minimalist sense to sentences or propositions that one does not understand, I found that the minimalist truth predicate cannot be applied to sentences that one does not understand. For one does not know what propositions are expressed by those sentences. As a result, the minimalist truth predicate is not applicable to sentences that one does not understand. This further damages Horwich's mathematical inductive reasoning regarding the minimalist explanation of general facts involving truth. This reasoning has its effectiveness only for sentences that one understands, but not for all propositions. The possible expansion of language does not much help here. Based on the assumption or common sense that for any declarative *sentence* there is a corresponding proposition, Horwich formulates his equivalence schema. For sentences that a given speaker does not understand, the function of the minimalist truth predicate that plays the role of forming generalisations, is only fulfilled by accepting the aforementioned assumption that those non-understood sentences do express propositions because of their declarative feature. In other words, the minimalist truth ascription to many instances of a given generalisation is founded on our implicit, if not explicit, acceptance of (D$_Q$).

In the final section of Chapter Five, I briefly discussed Quine's possibly incoherent attitudes toward the matter of truth. I did not see any reconciliation between his disquotational reasoning with respect to truth and his advocacy of a Davidsonian truth conditional approach. In terms of the comparison between Quine's disquotationalism and Horwich's minimalism, however, I concluded that if there is a strong theory of meaning, then we can have a general account of disquotational truth or minimalist truth. Even given this situation, Quine's disquotationalism is preferable to Horwich's minimalism. If there

is no strong theory of meaning, then disquotationalism has to cope with the following

problem: how do speakers of different languages, or even the same language, use the truth

predicate? But Horwich's minimalism is, on the other hand, destined to fail.

Bibliography:

Akiba, Ken (2002) "Can Deflationism Allow for Hidden Indeterminacy?", in *Pacific Philosophical Quarterly*, Volume 83, pp. 223-234.

Alston, W. (1996) *A Realist Conception of Truth*, Ithaca, NY: Cornell University Press.

Armour-Garb, Bradley (2004) "Minimalism, the Generalization Problem and the Liar", in *Synthese*, Volume 139, pp. 491-512.

Azzouni, Jody (1998) "On "On What There Is"", in *Pacific Philosophical Quarterly*, volume 79, issue 1, pp. 1-18.

Boghossian, Paul (1990*a*) "The Status of Content", in *The Philosophical Review*, April, pp. 247-263.

——(1990*b*) "The Status of Content Revisited", in *Pacific Philosophical Quarterly* 71, pp. 264-278.

Brandom, R. (1994), *Making it Explicit*, Cambridge, Mass.: Harvard University Press.

Brendel, Elke (2000), "Circularity and the Debate between Deflationist and Substantive Theories of Truth", in *Circularity, Definition and Truth*, edited by André Chapuis and Anil Gupta, New Delhi: Indian Council of Philosophical Research, pp.29-47.

Buchanan, R. (2003) "Are Truth and Reference Quasi-Disquotational?", in *Philosophical Studies*, volume 113, no. 1, pp. 43-75.

Christensen, David (1993) "Skeptical Problems, Semantic Solutions", in *Philosophy and Phenomenological Research*, Volume 53, Issue 2, pp. 301-321.

Collins, John (1993) "Horwich's Schemata Meet Syntactic Structures", in *Mind*, Volume 112, Issue 447, pp. 399-431.

David, Marian A. (1994), *Correspondence and Disquotation*, New York, Oxford University press.

——(1996), "Analyticity, Carnap, Quine, and Truth", in *Philosophical Perspectives*, 10, *Metaphysics*, pp. 281-296.

——(2000), "Minimalism and the Facts about Truth", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, pp.161-175.

——(2003) "Truth, Eliminativism, and Disquotationalism", in *Noûs*, Volume 23, Issue 447, pp. 599-614.

Davidson, Donald (1979), "The Inscrutability of Reference" in *Inquiries into Truth and Interpretation* (1984), New York, Oxford University press, pp. 227-241.

——(1990) "The Structure and Content of Truth", in *Journal of Philosophy*, Volume 87, Issue 6, pp. 279-328.

——(1996) "The Folly of Trying to Define Truth", in *Journal of Philosophy*, Volume 93, Issue 6, pp. 263-278.

Devitt, Michael (2001), "The Metaphysics of Truth", in *The Nature of Truth*, edited by Michael Lynch, Cambridge, MA.: MIT Press, pp. 579-611.

Dillard, Peter S. (1996) "Radical Anti-Deflationism", in *Philosophy and Phenomenological Research*, Volume 56, Issue 1, pp. 173-181.

Engel, Pascal (2002), *Truth*, Montreal: McGill-Queen's University Press.

Etchemendy, John (1988), "Tarski on Truth and Logical Consequence", *The Journal of Symbolic Logic* 53, pp. 51-79.

Field, Hartry (1972), "Tarski's Theory of Truth" (including "Postscript" (2001)), in *Truth and the Absent of Fact* (2001). Oxford: Clarendon Press, pp. 3-29.

——(1973), "Theory Change and the Indeterminacy of Reference" (including "Postscript" (2001)), in *Truth and the Absent of Fact* (2001). Oxford: Clarendon Press, pp. 177-198.

——(1974), "Quine and the Correspondence Theory" (including "Postscript" (2001)), in *Truth and the Absent of Fact* (2001). Oxford: Clarendon Press, pp. 199-221.

——(1986), "The Deflationary Conception of Truth", in *Fact, Science, and Morality: Essays on A. J. Ayer's 'Language, Truth and Logic'*, edited by G. Macdonald and C. Wright. Oxford: Basil Blackwell, pp. 55-117.

——(1994), "Deflationist Views of Meaning and Content" (including "Postscript" (2001)), in *Truth and the Absent of Fact* (2001). Oxford: Clarendon Press, pp. 104-156.

——(1999), "Deflating the Conservativeness Argument", *The Journal of Philosophy*, 96, (pp. 533-40).

Fodor, J. (1994) *The Elm and the Expert*, Cambridge, Mass: MIT Press.

Fodor, J.; Lepore, E. (2001) "Why Compositionality Won't Go Away: Reflections on Horwich's 'Deflationary' Theory", in *Ratio*, volume 14, no. 4, pp. 350-368.

Fogelin, Robert J. (1997) "Quine's Limited Naturalism", in *The Journal of Philosophy*, volume 94, no. 11, pp. 543-563.

Forbes, G. (1986) "Truth, Correspondence, and Redundancy", in G. MacDonald and C. Wright eds., *Fact, Science, and Morality*, Oxford: Blackwell, pp. 671-687.

Georgalis, Nicholas (2000) "Reference Remains Inscrutable", in *Pacific Philosophical Quarterly*, volume 81, issue 2, pp. 123-129.

Gibson, Roger, Jr. (1988), *Enlightened Empiricism*, Florida: University Presses of Florida.

Grover, Dorothy (1990) "Truth and Language-World Connections", in *Journal of Philosophy*, Volume 87, Issue 12, pp. 671-687.

—— (2000), "On Locating Our Interest in Truth", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, pp.120-132.

Gupta, A. (1993*a*), "A Critique of Deflationism", in *Philosophical Topics* Vol. 21 no. 2, pp. 57-81.

——(1993*b*), "Minimalism", in *Philosophical Perspectives 7: Philosophy of Language*, edited by J. E. Tomberlin. Atascadero, Calif.: Ridgeview Publishing Company.

——(2002), "An Argument Against Tarski's Convention T", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, 225-237.

Halbach, Volker (1999), "Disquotationalism and Infinite Conjunctions", *Mind* Vol. 108 (429), pp. 1-22.

——(2000), "Disquotationalism Fortified", in *Circularity, Definition and Truth*, edited by André Chapuis and Anil Gupta, New Delhi: Indian Council of Philosophical Research, pp.155-176.

——(2001), "How Innocent Is Deflationism?", *Synthese* 126, pp. 167-194.

Heck, Richard G., Jr. (1997), "Tarski, Truth, and Semantics", in *The Philosophical Review*, Volume 106, pp. 533-554.

Hershfield, J.; Soles, D. H. (2003), "Reinflating Truth as an Explanatory Concept", in *Pacific Philosophical Quaraterly*, Volume 84, pp. 32-42.

Horwich, Paul (1996) "Realism and Truth", in *Philosophical Perspectives*, 10, Metaphysics, pp. 187-197.

——(1997*a*) "Davidson on Deflationism", in U. Zeglen, ed., *Discussions with Donald Davidson: On Truth, Meaning, and Knowledge*, London: Routledge, pp. 20-24.

——(1997*b*) "Implicit Definition, Analytic Truth, and Apriori Knowledge", in *Noûs*, Volume 31, Issue 4, pp. 423-440.

——(1998*a*) *Meaning*, Clarendon Press: Oxford.

——(1998*b*) *Truth*, second edition, Clarendon Press: Oxford.

——(2001*a*) "A Defense of Minimalism", in *Synthese*, 126, pp. 149-165.

——(2001*b*) "Deflating Compositionality", in *Ratio*, volume 14, no. 4, pp. 369-385.

Hylton, Peter (2004), "Quine on Reference and Ontology", in *The Cambridge Companion of Quine*, ed. By Roger Gibson, Jr., pp. 115-150.

Kovach, Adam (1997) "Deflationism and the Derivation Game", in *Mind*, Volume 106, Issue 423, pp. 575-579.

Kraut, Paul (1993) "Robust Deflationism", in *The Philosophical Reviews*, Volume 102, Issue 2, pp. 247-263.

Künne, Wolfgang (2000), "Disquotationalist Conceptions of Truth", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, pp. 176-193.

Leeds, Stephen (1973), "How to Think about Reference" *Journal of Philosophy* 15: pp. 485-503.

McGee, Vann (1993) "A Semantic Conception of Truth", in *Philosophical Topics*, vol. 21, No. 2, spring, pp. 83-111.

McGinn, Colin (2000), "The Truth about Truth", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, pp. 194-204.

McGrath, Matthew (1997*a*), "Reply to Kovach", in *Mind*, Vol. 106, issue 423, pp. 581-586.

——(1997*b*), "Weak Deflationism", in *Mind*, Vol. 106, issue 421, pp. 69-98.

——(2000), *Between Deflationism & Correspondence Theory*, New York: Garland Publishing.

Miller, Alexander (2000), "Horwich, Meaning and Kripke's Wittgenstein", in *Philosophical Quarterly*, Vol. 50, issue 199, pp. 161-174.

——(2001), "On Wright's Argument against Deflationism", in *Philosophical Quarterly*, Vol. 51, issue 205, pp. 527-531.

Nelson, R. J. (1997) "Proxy Functions, Truth and Reference", in *Synthese*, volume 111, issue 1, pp. 73-95.

O'Leary-Hawthorne, John; Oppy, Graham (1997) "Minimalism and Truth", in *Noûs*, Volume 31, Issue 2, pp. 170-196.

Quine, W. V. (1951) "Ontology and Ideology", in *Philosophy Studies* 2, pp. 11-15.

——(1953*a*), "Notes on the Theory of Reference", in *From a Logical Point of View*, Cambridge, Mass.: Harvard university Press, pp. 130-138.

——(1953*b*), "Two Dogmas of Empiricism", in *From a Logical Point of View*, Cambridge, Mass.: Harvard university Press, pp. 20-46.

——(1954), "Carnap and Logical Truth", in *The Ways of Paradox and Other Essays*, revised ed. Cambridge, Mass.: Harvard University Press, pp. 107-132.

——(1960), *Word and Object*, Cambridge, Mass.: M.I.T. Press.

——(1969*a*), "Epistemology Naturalized", in *Ontological Relativity and Other Essays*, New York: Columbia University Press, pp. 69-90.

——(1969*b*), "Ontological Relativity", in *Ontological Relativity and Other Essays*, New York: Columbia University Press, pp. 26-68.

——(1970), *Philosophy of Logic*, Eaglewood Cliffs, N.J.: Prentice-Hall.

——(1974*a*), "On Popper's Negative Methodology", in *The Philosophy of Karl Popper*, ed. by P. A. Schlipp, La Salle, Ill: Open Court.

——(1974*b*), *Roots of Reference*, La Salle, Ill: Open Court.

——(1975), "On Empirically Equivalent Systems of the World", *Erkenntnis* 9, pp.

313-328.

——(1976), "Truth and Disquotation" in *The Ways of Paradox and Other Essays*, revised ed. Cambridge, Mass.: Harvard University Press, pp. 308-321.

——(1977), "Facts of the Matter", in *American Philosophy from Edwards to Quine*, ed. by Shahan and Merrill, Norman, Okla.: University of Oklahoma Press.

—— (1981*a*), "Reply to Eleven Essays" in *Philosophical Topics* 12, pp. 227-243.

——(1981*b*), "Empirical Content" in *Theories and Things*, Cambridge, Mass.: Harvard university Press, pp. 24-30.

——(1981*c*), "Responses" in *Theories and Things*, Cambridge, Mass.: Harvard university Press, pp. 173-186.

——(1986*a*), "Reply to Paul A. Roth" in *The Philosophy of W. V. Quine*, edited by L. Hahn and P. Schilpp, La Salle, Ill: Open Court, pp. 459-461.

——(1986*b*), "Reply to Roger F. Gibson, Jr." in *The Philosophy of W. V. Quine*, edited by L. Hahn and P. Schilpp, La Salle, Ill: Open Court, pp. 155-157.

——(1990*a*) "Comment on Lauener" in *Perspectives on Quine*, ed. by Robert B. Barrett and Roger F. Gibson, Cambridge, Mass.: Basil Blackwell, p.229.

——(1990*b*) *Pursuit of Truth*, Cambridge, Mass.: Harvard University Press.

——(1995) *From Stimulus to Science*, Cambridge, Mass.: Harvard University Press.

Ramsey, F. P. (1927), "Facts and Propositions", *Proceedings of the Aristotelian Society*, suppl. Vol. 7, pp. 153-171.

Robbins, Philip (2001) "What Compositionality Still Can Do", in *Philosophical Quarterly*, volume 51, issue 204, pp. 328-336.

Romanos, George D. (1983), *Quine and Analytic Philosophy*, Cambridge, Mass.: M.I.T. Press.

Roth, Paul A. (1986), "Semantics without Foundations" in *The Philosophy of W. V. Quine*, edited by L. Hahn and P. Schilpp, La Salle, Ill: Open Court, pp. 433-458.

Schantz, Richard (2000), "How Deflationists Diverge from Tarski", in *Circularity, Definition and Truth*, edited by André Chapuis and Anil Gupta, New Delhi: Indian Council of Philosophical Research, pp.317-331.

Schiffer, Stephen (2000) "Horwich on Meaning", in *Philosophical Quarterly*, volume 50, issue 201, pp. 527-536.

Shapiro, S. (1998), "Proof and Truth: Through Thick and Thin", *Journal of Philosophy* 95, pp. 493-521.

——(2003) "The Guru, the Logician and the Deflationist: Truth and Logical Consequence", in *Noûs*, volume 37, no. 1, pp. 113-132.

Sher, Gila (1998) "On the Possibility of a Substantive Theory of Truth", in *Synthese*,

volume 117, issue1, pp. 133-172.

Smith, Michael (1994*a*), "Minimalism, Truth-aptitude and Belief", in *Analysis*, 54.1, pp. 21-26.

——(1994*b*), "Why Expressivists about Value should Love Minimalism about Truth", in *Analysis*, 54.1, pp. 1-12.

Soames, Scott (1997), "The Truth about Deflationism", in *Philosophical Issues* 8, edited by E. Villanueva, Atascadero, Cal.: Ridgeview Publishing Company.

——(1999), *Understanding truth*, Oxford, New York: Oxford University Press.

Sosa, Ernest (1993*a*) "Epistemology, Realism, and Truth", in *Philosophical Perspectives*, Volume 7, Issue Language and Logic, pp. 1-16.

——(1993*b*), "The Truth of Modest Realism", in *Philosophical Issues* 3, *Science and Knowledge*, pp. 77-95.

Tarski, Alfred (1956*a*), "The Establishment of Scientific Semantics" in Tarski, *Logic, Semantics, and Metamathmatics*, tr. J. H. Woodger, London: Oxford university Press.

——(1956*b*), "The Concept of Truth in Formalised Languages" in Tarski, *Logic, Semantics, and Metamathmatics*, tr. J. H. Woodger, London: Oxford university Press.

Van Cleve, James (1996) "Minimal Truth Is Realist Truth", in *Philosophy and Phenomenological Research*, Volume 56, Issue 4, pp. 869-875.

Van Inwagen, Peter (2000), "Generalizations of Homophonic Truth-Sentences", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, pp. 205-222.

Williams, Michael (1999), "Meaning and Deflationary Truth", in *The Journal of Philosophy*, 96, pp.545-564.

——(2002), "On Some Critics of Deflationism", in *What is Truth*, edited by Richard Schantz, Berlin, New York: de Gruyter, pp.146-158.