



REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree *Mphil*

Year *2005*

Name of Author *FITZGERALD, S.*

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS

Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

This copy has been deposited in the Library of *UCL*

This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

DONALD

Davidson's Principle of Charity

Gareth Fitzgerald

Department of Philosophy

University College London

M.Phil Thesis

April 2005

UMI Number: U593886

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U593886

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I hereby declare that this thesis is all my own work.

Candidate's signature:

(Gareth Fitzgerald)

Abstract

The claim of this thesis is that Davidson's Principle of Charity is not a necessary condition on constructing a theory of meaning for a natural language. There are means other than the application of Davidson's Principle of Charity by which an interpreter can achieve the disentanglement of questions of meaning and belief for which Davidson's principle was enlisted. These other means consist in the employment of evidence about speakers, and speaker's interaction with their environment, from the human sciences such as psychology and linguistics. The thesis includes a detailed account of the content, status and development of Davidson's Principle of Charity placed in the wider context of his approach to the philosophy of language. Davidson's views on the evidence available to the interpreter are examined and found to rest on doubtful behaviouristic arguments originating from Quine. The behaviouristic view of the evidence is rejected, and a broader and more open-minded construal of the available evidence is sketched and defended. It is argued, following this reassessment of the available evidence, that Davidson's methodology should be relaxed in the following way. Where the extra evidence is not neutral with regard to the contributions of meaning and belief to a speaker's utterances, the interpreter need not employ a Principle of Charity. The interpreter should enlist hypothetical, rather than aprioristic, maxims subject to constant empirical tests.

Table of Contents

1. Introduction	4
2. Davidson's Approach to the Philosophy of Language	8
2.1. What is a Theory of Meaning?	8
2.2. Davidson and Tarski	10
2.3. Truth theory and Natural Language	12
2.4. Satisfying the Conditions on a Theory of Meaning	14
2.5. Foster's Problem	17
3. Davidson's Principle of Charity	21
3.1. The Origins of Davidson's Principle of Charity	21
3.2. Coherence and Correspondence	26
3.3. The Move from Maximising Agreement to Optimizing Intelligibility	30
3.4. Assertion: The Very General Attitude of Holding True	33
3.5. Holding Belief Fixed to Solve for Meaning	36
3.6. Testing a Theory of Meaning	38
3.7. Making Disagreement Intelligible Against a Background of Agreement	39
4. The Available Evidence and the Status of Charity	42
4.1. Davidson's Principle of Charity as a Necessity of Interpretation	42
4.2. Grandy's Principle of Humanity	50
4.3. McGinn on Davidson's Principle of Charity and the Causal Theory of Belief	58
4.4. Quine and Davidson: A Behaviouristic Legacy	65
4.5. Davidson and Lewis: Lewis's Radical Interpreter	82
5. Some Positive Suggestions and Some Conclusions	89

1. Introduction

Davidson saw the application of a Principle of Charity as unavoidable in interpreting the words and thoughts of others.¹ Davidson's view was that if one does not adopt his Principle of Charity (henceforth, "Davidson's Principle of Charity" reads DPOC) then interpretation becomes impossible. He saw DPOC as the only means by which to disentangle questions of what a speaker means from questions of what a speaker believes. According to Davidson, one cannot know the meaning of an utterance unless one knows what belief the speaker has expressed in making it, and similarly one cannot know what it is a speaker believes without knowing what his utterances mean. DPOC is a device for holding the speaker's beliefs fixed in accord with the interpreter's own responses to the environment and standards of consistency. By holding belief fixed the interpreter can begin solving for the meanings of the speaker's utterances. The meanings of a speaker's utterances are stated in terms of their truth-conditions.

These other means consist in the employment of evidence about speakers and their environments from the natural sciences; evidence relevant to the attribution of meanings and beliefs. Davidson does not afford this evidence a central role in interpretation because he has accepted Quine's behaviouristic construal of the evidence relevant to the linguist. To establish the master claim, two sub-claims must be established: (1) There is admissible, relevant evidence available about speakers which is not utilised in Davidson's Charitable methodology², and (2) Such evidence can serve the disentanglement with which Davidson was concerned. Quine saw the entanglement of meaning and belief as one source of the indeterminacy of translation. This thesis is not claiming to fully resolve that source of indeterminacy; indeed, Davidson did not see his method as resolving it.³

¹ Davidson, 1967, pp.27. Unless otherwise attributed, all quotations are from Davidson.

² A "Charitable Method" for interpretation is one that employs a Principle of Charity.

³ "No single optimum charity emerges; the constraints therefore determine no single theory. In a theory of radical translation (as Quine calls it) there is no completely disentangling questions of what the alien means from questions of what he believes" 1967, pp. 27.

The claim then is that DPOC is not the only available means to disentangling questions of meaning and belief. The employment of knowledge from sciences such as psychology, linguistics and optics is entirely relevant to pinning down a speaker's beliefs. Knowledge from empirical linguistics may be relevant to pinning down a speaker's meanings. If we start out assuming nothing and knowing nothing about the speaker with which we are presented, we might, for example, learn through empirical psychology that he tracks enduring medium-sized objects against contrasting backgrounds much as we do. Or we may learn from studying the workings of the speaker's eyes or ears that they function very differently to our own and are connected to the speaker's brain in some peculiar way. This information should be carefully handled but is entirely relevant to the discernment of the native's beliefs. The availability of such empirical evidence in an impressive quantity would cast doubt on the necessity of DPOC to the disentanglement of belief and meaning that must occur in interpretation. The claim of this thesis is not the absurd one that if, for example, we know that a speaker's eyes work much as ours do then we can conclude that his perceptual beliefs must be the same as ours. This claim relies on a crude view of the relationship between empirical evidence and belief ascriptions. My claim is the weaker one that natural evidence about a speaker is relevant to our discernment of a speaker's meanings and beliefs. To take an extreme case, our knowledge that a man is blind renders it highly unlikely that he has the same beliefs about the visual environment as his sighted interpreter.

Davidson's Charitable methodology should be revised in the context of a more open-minded view of the evidence available to the interpreter. The Principle of Correspondence, a constituent of DPOC⁴, prompts the interpreter to take the speaker to be responding to the same features of the world as he would in those circumstances. This is not something we need assume but a matter that can be illuminated by the sciences. The admission of such evidence does not speak in favour of a reduction of the intentional notions to the natural evidence. Supposing that the evidence one gathered was described in physical terms, then the reductive claim

⁴ 1983, pp.211.

would be that the intentional notions could be defined in scrupulously physical terms. One can hold that natural evidence is relevant to working out someone's beliefs, without being reductive. On several occasions Davidson himself allows that knowledge of the speaker's sensory apparatus, positioning, background capabilities and other facts can be fruitful in interpretation.⁵ Yet no such concessions are to be found in the canonical formulations of DPOC.

Davidson's behaviouristic view of the evidence is the result of two factors. Firstly, there is overwhelming evidence that Davidson accepted Quine's behaviouristic construal of the evidence relevant to the study of language. Not only does Davidson acknowledge his own discussion of Radical Interpretation as resembling Quine's discussion of Radical Translation with respect to the mechanics, he accepts a version of Quine's manifestation argument for linguistic behaviourism (see 4.4). Quine's behaviouristic arguments are highly dubious and Davidson's method suffers for his acceptance of them. Secondly, there was a lingering conviction on Davidson's part that his Charitable method by itself was adequate to good interpretation.⁶ Davidson said that the application of DPOC to the behavioural evidence, together with the satisfaction of the formal constraints⁷ on a theory of meaning, exhaust the constraints on interpretation.⁸ My thesis is a denial of this point: Davidson does not give us good reason to limit the interpreter in this way.

There is a relation between an utterance in observable circumstances and the truth-conditions of that utterance such that the utterance constitutes only one piece of defeasible evidence amongst several sources of defeasible evidence for the truth-conditions. Davidson himself does not instruct the interpreter always to take an utterance to be true to the observable circumstances, he counts utterances on

⁵ 2001b, pp.xix. 1974c, pp.282.

⁶ 2001b, pp.xx. Having accepted Lewis's injunctions over knowledge of sensory apparatus, bad positioning et al, Davidson maintained:

"I do not take this to prove that the evidential base on which I depend is not in theory adequate."

See also 1974c.

⁷ As explained in section 2 of this thesis, the formal constraints involve satisfying Tarski's Convention T on a theory of truth.

⁸ 1974b, pp.196-7.

occasions as defeasible evidence for truth-conditions. The difference between the current suggestion and Davidson's is that he enlists only one source of defeasible evidence, namely behaviour, where one should be prepared to admit many. Davidson agrees that the relation between an utterance on an occasion and its truth-conditions may have no uniform character; there may be factors peculiar to individual cases. But his limited view of the defeasible evidence commits his interpreter to finding a high degree of truth in the speaker. A belief, call it belief*, is identified on Davidson's view only by its place amongst a surrounding network of beliefs. On this view, too much falsity in the surrounding beliefs undermines the purported subject matter of belief*. It is plausible that with the employment of more diverse sources of evidence in interpreting a speaker, one can begin to discern their beliefs in ways that do not require placing those beliefs in a constellation of true belief. In this respect, this thesis is not only a challenge to Davidson but also the offer of some moderation of the strong epistemological consequences he saw his approach to truth and interpretation as having.

The first objective of this thesis is an overview of Davidson's approach to the philosophy of language. The second and most important objective is to give a detailed and coherent account of DPOC. A problem with some of the secondary literature is that it shows only a weak understanding of DPOC and its place within Davidson's method, and so loses its bite as a critique of Davidson. My account of DPOC aims to do justice to the important refinements and additions Davidson made to a formulation he inherited from Quine. The fourth and central section of the thesis presents a challenge to Davidson's pronouncements on the necessity of DPOC to interpretation. Suggestions are made concerning the kind of evidence or knowledge that may be employed in effecting the disentanglement of meaning and belief with which Davidson was concerned. My claim is that such evidence is afforded at most a peripheral role by Davidson and that this marginalisation is unjustified. The thesis concludes with some positive sketches of how the employment of further evidence might go.

2. Davidson's Approach to the Philosophy of Language

This section outlines Davidson's approach to the philosophy of language up to the point at which DPOC enters the picture. Davidson invites us to consider how we should go about constructing a theory of meaning for a language. DPOC becomes pivotal in his answer to this question. To understand this matter we need to equip ourselves with an account of the form that Davidson thought such a theory should take, and a sense of why he thought this question so central to the philosophy of language.

2.1 What is a Theory of Meaning?

The construction of a theory of meaning is intended to provide an answer to the question of how utterances mean what they do. It is, of course, unclear how exactly one should take this question. One reason for this lack of clarity is that it is hard to give an explanation of the term "mean" in "utterances mean what they do". Another reason for the lack of clarity is that the "how" in "how utterances mean what they do" lends itself to a range of approaches. Nonetheless, Davidson proposed that we could answer our original question if we knew how to construct a theory that met three conditions. We can specify these three conditions, and give the rationale for them, before we go on to describe the species of theory that Davidson thought would meet them satisfactorily.

Firstly, the theory must state the meaning of every sentence, actual and potential, of the language in question.

If sentences depend for their meaning on their structure, and we understand the meaning of each item in the structure only as an abstraction from the totality of sentences in which it features, then we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (or word) in the language.⁹

⁹ 1967, pp.22.

If a speaker knows the theory then he should understand the language in question; he should know what all the sentences of the language mean. Davidson notes that the question of 'what we could know' that would enable us to understand a speaker of some language is not the same as the question of 'what we do in fact know' that enables us to understand the utterances of others.

For there may easily be something we could know and don't, knowledge of which would suffice for interpretation, while on the other hand it is not altogether obvious that there is anything we actually know that plays an essential role in interpretation.¹⁰

Furthermore, Davidson's account of how we could come to know a theory of meaning is not intended to be an account of the actual process of first or second language acquisition. A theory of meaning is not an empirical theory of language learning (although it should fit well with one). His question is thus "a doubly relative question: given a theory that would make interpretation possible, what evidence plausibly available to a potential interpreter would support the theory to a reasonable degree?"¹¹

Secondly, the theory must be confirmable without prior knowledge of the detailed propositional attitudes of the speakers of the language:

interpretation¹² cannot hope to take as evidence for the meaning of a sentence an account of the complex and delicately discriminated intentions with which the sentence is typically uttered the central difficulty is that we cannot hope to attach a sense to the attribution of finely discriminated intentions independently of interpreting speech. The reason is not that we cannot ask necessary questions, but that interpreting an agent's intentions, his beliefs and

¹⁰ 1973a, pp.125.

¹¹ *ibid*

¹² "Interpretation" here meaning the construction of a theory of meaning for a speaker. Davidson often talks of a "theory of interpretation" rather than a "theory of meaning". One gets the impression that the phrases "theory of interpretation" and "theory of meaning" are used interchangeably. But Davidson is at pains to point out that interpretation, coming to understand a speaker, involves not only working out the meaning of his words but also constructing a theory of his beliefs and actions.

his words are parts of a single project, no part of which can be assumed to be complete before the rest is.¹³

Thirdly, the theory must identify a finite basic vocabulary of the language from which we can explain the meaning of a potential infinity of sentences. A language with infinite basic vocabulary would not be learnable by creatures of finite powers.

2.2 Davidson and Tarski

With these three conditions in mind, Davidson argued that a theory of truth, in the form of a Tarskian definition of truth for a formal language, could serve as a theory of meaning if adapted so as to apply to natural languages. Tarski sought to give an extensional definition of truth for formal languages, or more specifically a definition schema that a truth theory for any particular formal language must fit. According to Tarski's Convention T, an acceptable theory of truth should yield a theorem of the form:

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

for every sentence "S" of the language in question. "S" in this schema is replaced by a description of a sentence of the object (or target) language and "p" by the described sentence itself, where the metalanguage contains the object language, or by its translation into the metalanguage where it does not. The notion of translation occurs in Convention T: Tarski's condition of adequacy on a truth definition takes meaning (translation) as given when testing a truth definition. The truth definition for each formal language constitutes a partial definition of the concept of truth:

It should be emphasised that neither the expression (T) itself (which is not a sentence but only a schema of a sentence) nor any particular instance of the form (T) can be regarded as a definition of truth. We can only say that every equivalence of the form (T) obtained by replacing 'p' by a particular sentence and 'S' by a name of this sentence, may be considered

¹³ 1973a, p.127.

a partial definition of truth, which explains wherein the truth of this one individual sentence consists. The general definition has to be, in a certain sense, a logical conjunction of all these partial definitions.¹⁴

The formalised languages with which Tarski was concerned could be known by their syntax. Even if the object language and metalanguage were different, the meanings of the terms were still unproblematically available.¹⁵ Davidson intended a reversal of the dependence of the truth definition on the notion of meaning that was to be found in Tarski.

a syntactical test of the truth of a T-sentence would be worthless, since such a test would presuppose the understanding of the object language one hopes to gain.¹⁶

The meanings of natural language sentences and hence translations between natural languages are not unproblematically available; we cannot give the meanings of natural language terms by stipulation and the meaning of natural language sentences is not given in their syntax. So Davidson took our grasp of the truth predicate rather than a notion of translation or meaning to be primitive. Within a theory a primitive term is one which we decide to use without giving an explicit definition in other terms. He assumed that truth was a concept of which we had a grasp prior to constructing a theory of meaning. This is not the place to give a detailed account of Davidson's views on truth but suffice to say, he argued that truth could not be defined in more basic terms.¹⁷ Davidson found it plausible that one good sense to be given to saying someone understands a declarative sentence is that he knows what it is for it to be true.¹⁸ Davidson sought to give an analysis of the meaning of the sentences of a language by detailing the dependence of their truth-conditions on the sentences' composition. What Davidson required of a truth theory for a natural language was

¹⁴ Tarski, 1944, pp.71.

¹⁵ 1974a, pp.150.

¹⁶ *ibid*

¹⁷ 1996, pp.263-278.

¹⁸ See 1967.

that it met Convention T and that it could be confirmed without using the notion of meaning or translation.

2.3 Truth theory and Natural Language

To return briefly to the difference between Tarski and Davidson's projects, Tarski's schema for a truth definition was intended for *formalised* languages and his test for the adequacy was intended as a test for truth theories for formalised languages.

Formalised languages can be roughly characterised as artificially constructed languages in which the sense of every expression is uniquely determined by its form.¹⁹

Tarski was prepared to call languages formal if they fulfilled the following criteria: (1) The full vocabulary of the language is available, (2) The syntactic rules are precisely formulated, (3) The syntactic rules refer only to the form of an expression (4) The function and meaning of an expression depend only on its form. Tarski thought that his truth definition applied only to formalised rather than natural languages for two reasons.

Tarski's first reason was that "the universal character of natural languages leads to contradiction (the semantic paradoxes)".²⁰ The "universal character" of natural language refers to the lack of restrictions on the sentences we allow to be formed in natural language. The quantifiers of natural language have a broad range of values in the following sense. We allow the quantifiers of natural language to range over sentences of natural language themselves and can derive paradoxes from this feature of natural language. We can form, for example, the sentence "This sentence is false" which is true if and only if it is false. Davidson, unlike Tarski, took the fact that such paradoxes arise as a reason to be suspicious that natural languages are universal in this way.

¹⁹ Tarski, 1935, sect. 2.

²⁰ 1967, pp.28.

Tarski's second reason for thinking that his work on truth did not apply to natural languages was that "natural languages are too confused and amorphous to permit the direct application of formal methods".²¹ Tarski thought we would have to reform natural language out of all recognition before we could apply his definition to them. Davidson saw that:

If this is true it is fatal to my project, for the task as I conceive it is not to change, improve, or reform a language, but to describe and understand it.²²

Davidson was more optimistic than Tarski about the application of formal methods to natural language. The following remarks illustrate Davidson's hopes that the ambiguity and amorphousness of natural language are not prohibitive:

As long as the ambiguity does not affect grammatical form, and can be translated, ambiguity for ambiguity, into the metalanguage, a truth definition will not tell us any lies.²³

To give an example: suppose success in giving the truth conditions for some significant range of sentences in the active voice. Then with a formal procedure for transforming each such sentence into a corresponding sentence in the passive voice, the theory of truth could be extended in an obvious way to this new set of sentences.²⁴

Davidson thought his suggestion paved the way for a "serious semantics of natural language, for it is likely that, many outstanding puzzles, such as the analysis of quotations or sentences about propositional attitudes, can be solved if we recognise a concealed demonstrative construction".²⁵ It is important to recognise the challenge that Davidson saw here because without the prospect of resolution to these problems his claim that a Tarskian truth theory could serve as a theory of meaning would begin to look implausible. It is not a goal of this paper to assess the prescience or plausibility of these claims about the possible regimentation of natural language.

²¹ *ibid*

²² 1967, pp.29.

²³ 1967, pp.30.

²⁴ *ibid*

²⁵ 1967, pp.35.

2.4 Satisfying the Conditions on a Theory of Meaning

Davidson's idea then was that a truth definition adapted to natural language, which satisfied Convention T and fitted all the empirical evidence could serve as a theory of meaning for that language. Let us look back at our three conditions on a theory of meaning. Our first condition was that the theory states the meaning of every sentence of the language in question. If we accept a very close relation between knowing the truth-conditions of an assertoric sentence and knowing what that sentence means²⁶ then this condition is met. Convention T ensures that a theorem stating the truth-conditions of every sentence of the language in question will be given in an adequate theory of truth:

the definition works by giving necessary and sufficient conditions for the truth of every sentence, and to give truth conditions is a way of giving the meaning of a sentence. To know the semantic concept of truth for a language is to know what it is for a sentence – any sentence – to be true, and this amounts, in one good sense we can give to the phrase, to understanding the language.²⁷

Within such a theory of truth it is possible to prove from a finite set of axioms and a set of procedural rules, a theorem, for each sentence of the language, that states the truth conditions of that sentence. The recursiveness of a Tarskian truth theory seems to ensure that the third condition on a satisfactory theory of meaning is fulfilled. The proof of each theorem from the axioms shows how the truth or falsity of the sentence depends on its composition from finite vocabulary. The compositionality of the theory is one way in which it answers the question of how words mean what they do²⁸. The procedural rules tell us how theorems can be constructed from the axioms;

²⁶ In the discussion of Foster's problem that follows we are looking at one way in which a sentence's truth-conditions and a sentence's meaning seem to come apart. This is then a challenge to the idea that a correct Tarskian truth definition will state the meaning of every sentence of a language.

²⁷ 1967, pp.24.

²⁸ The other way in which it provides an answer to the question of how utterances mean what they do concerns the theory's fulfilment of the second condition. In specifying the evidence from which we could come to know what the utterances of a language mean we are saying something about how those

once we have some axioms stating the semantically relevant features of sentence constituents we can generate indefinitely many T-sentences by applying minimal logical machinery. The logical machinery that Davidson originally had in mind was just first-order quantificational logic but he later wavered on this point:

My working assumption has been that nothing more than first-order quantification theory is available. Indeed, I was long convinced that many alternative approaches to semantics, employing for example, modal logics, possible world semantics, or substitutional quantification, could not be accommodated in a theory that met the demands of Convention T. I now know this was hasty. Convention T does not settle as much as I thought, and more possibilities for interesting theorizing are open than I had realized. The well known virtues of quantification theory still provide plenty of motivation, however, to see how much we can do with it.²⁹

Davidson thought that the axioms of a truth theory must be finite in number if the theory is to give us insight into the structure of the language and how the meaning of a sentence depends on its parts.³⁰ He may have been wrong. One needs only a finite number of axiom schemas, not a finite number of axioms for such insight. One needs only the axiom schemas such as “x denotes y” or “x is satisfied if and only if y” to be enumerable for insight into the structure of the language, not for there to be a finite number of axioms such as “‘Gareth’ denotes Gareth”. His point that a language is only learnable if the basic vocabulary stated in the axioms is finite would still stand.

If one were constructing a theory for a formal language one could stipulate the conditions under which the sentences of that language are true. Sentences of natural languages, however, already have truth-conditions. So a question can be framed about what those truth-conditions are and which sentence of one natural language has the same truth-conditions as a certain sentence of another. How do we know when we have framed a T-sentence in which p states the truth conditions of s? To get clear on

utterances can be meaningful to the interpreter. Davidson’s way of meeting this second condition is the focus of this thesis. It is in relation to this condition that we are introduced to DPOC.

²⁹ 2001b, pp.xvii.

³⁰ See 1965.

the kind of objectively available evidence that could confirm or disconfirm a T-sentence, Davidson suggests, following Quine's discussion of Radical Translation in *Word and Object*, that we think about the radical case in which we encounter a language that is entirely unfamiliar to us and spoken by people about whom we know nothing. Davidson assumes that the language in which the theory is couched is one that we understand. In Davidson's scenario we get interpretations rather than translations, with "is true" occurring on only the left hand side of the bi-conditional. The reference to a language we understand becomes unnecessary, Davidson suggested, when we view an overall theory of interpretation as a compound of "a structurally revealing theory of interpretation for a known language, and a system of translation from the unknown language into the known."³¹

On Davidson's approach the evidence available to the interpreter in the radical situation consists in only the sounds he hears uttered, what he sees the speakers do and the observable circumstances. It is a little misleading to attribute this view to Davidson unreservedly. It is unclear whether he wanted to delimit the relevant evidence quite so narrowly, even in the radical case. For example, Davidson mentions in connection with the evidence relevant to interpretation "simplicity, hunches about the effects of social conditioning, and of course our common sense, or scientific knowledge of explicable error".³² He also says that "interpretation must take into account probable errors due to bad positioning, deficient sensory apparatus, and differences in background knowledge"³³. On the other hand, in answering the question of whether a theory of truth can be verified by appeal to evidence available before interpretation has begun, Davidson says:

the relevant evidence can consist entirely of facts about the behaviour and attitudes of speakers in relation to sentences (no doubt by way of utterances).³⁴

³¹ 1973a, pp.130.

³² 1974b, pp.196.

³³ 2001b

³⁴ 1973a, pp.133.

What is clear is that Davidson claimed that an interpreter cannot start out by making use of the intentions that accompany speech because discerning these intentions relies on and is part of working out what has been said:

radical interpretation cannot hope to take as evidence for the meaning of a sentence an account of the complex and delicately discriminated intentions with which the sentence is typically uttered.....the central difficulty is that we cannot hope to attach a sense to the attribution of finely discriminated intentions independently of interpreting speech. The reason is not that we cannot ask necessary questions, but that interpreting an agent's intentions, his beliefs and his words are parts of a single project, no part of which can be assumed to be complete before the rest is.³⁵

2.5 Foster's Problem, Compositionality, Law-likeness and the Interdependence of Meaning and Belief

There is a well-known problem for Davidson's approach, originally presented by Foster, concerning whether or not the theorems of his theory are interpretive. If Foster's problem cannot be given a credible answer then we cannot be satisfied that Davidson's theory will fulfil his first condition on a theory of meaning for a language; namely, that it state the meaning of every sentence of the language in question. This problem is discussed here, not in order to provide a fully explicit and convincing resolution to Foster's problem, or to claim that one can be found in Davidson's writings. Rather, we turn to it because Davidson's reaction to the problem is illuminating of his method and the way in which he viewed the theorems of an interpretive truth theory.

Foster's problem involves the possibility of correct but non-interpretive theorems. The truth of the bi-conditional requires only that the sentences on either side have the same truth value. So, how can we exclude from our theory theorems such as:

"Snow is white" is true if and only if grass is green

³⁵ 1973a, pp.127.

In which the sentences to the left and right of the bi-conditional share truth-value but the theorem would not, if known, allow one to understand the object sentence? In such an example, only the formal restrictions on a theory of meaning are in place. An interpretive truth theory would also have to fit the empirical evidence given by a speaker's use of the language in observable circumstances. A true theorem of a theory of meaning must at least be implied by a recursive theory of truth that fits all of the evidence.³⁶ The description of interpretation we are given by Davidson is of a process. We may be able to get our speakers talking on the subject matter of these bi-conditionals, and in doing so eliminate some of the alternatives on the basis of new evidence. We exchange one theory of truth for another – always treating our theory as a work in progress.

The theorem for each sentence of the object language must be derivable from a finite number of axioms. A theory in which we could derive the theorem:

“Snow is white” is true if and only if grass is green

would include amongst its axioms:

“Snow” denotes grass.

This would entail that within our theory we could derive the false theorem:

“Snow is falling” is true if and only if grass is falling.

Furthermore, Davidson³⁷ came to think that if a theory of truth were to explain how utterances meant what they did it would not be sufficient for the theory to be true. Firstly, its axioms and theorems must be law-like. Davidson saw that this was a

³⁶ 1973a

³⁷ 1976

concession to intensionality³⁸ within what had been intended as an extensional theory. How much of a concession this was, he argued, depends on one's analysis of a law. One must concede, however, that the notion of a law is one drawn on in all the empirical sciences. Secondly, someone in possession of the truth theory could only understand a speaker of the object language, could only know what they meant, if they knew that the theorems of the theory were law-like. This second requirement may be excessive. It amounts to the requirement that the interpreter needs not only to know what the speaker's utterance means, he also needs to be aware that what he has is knowledge. An alternative response to this requirement, one defended by Larson and Segal³⁹, is that we are disposed to treat a correct truth theory as interpretive. They claim that if one came to possess an interpretive theory then "the knowledge gap, though present, would be irrelevant to understanding or action, since speakers would proceed as if they already knew that their truth theories were interpretive."⁴⁰

There may be harder instances of Foster's problem for the truth-theoretic approach that are not dealt with by the compositionality or law-likeness of the theorems. One such case involves counterfeit theorems into which have been substituted co-referential terms that differ in sense.⁴¹ How could we distinguish the interpretive:

"Hesperus is bright" is true if and only if Hesperus is bright

From the non-interpretive "Hesperus is bright" is true if and only if Phosphorus is bright'? Since Hesperus and Phosphorus are identical it doesn't seem that we can appeal to the law-likeness of the theorems to explain why one is interpretive and the other not so. The compositionality of the theory won't serve to distinguish them either: as "Hesperus" and "Phosphorus" have the same reference we would not get false theorems involving sentences in which "Hesperus" and "Phosphorus" occur.

³⁸ Laws seem to form an intensional context, a context in which the substitution of expressions with the same semantic value need not preserve the semantic value (truth-value) of the original sentence.

³⁹ Larson, R. and Segal, G., 1995.

⁴⁰ Larson, R. and Segal, G., 1995, pp.39.

⁴¹ Assuming we find talk of terms having a sense and a referent cogent.

There may be a way to distinguish the interpretive truth theory from the counterfeit available within the truth theoretic approach described by Davidson. The other half of constructing a theory of meaning for a speaker, of interpreting a speaker, is ascribing beliefs to the speaker. According to the interpretive theorem the speaker's utterance "Hesperus is bright" provides evidence that he believes that Hesperus is bright and according to the counterfeit it is evidence for a belief that Phosphorus is bright. But if we assign our speaker the belief that Phosphorus is bright then shouldn't we predict that he'll hold true the sentence "Phosphorus is bright"? If he doesn't know that Hesperus and Phosphorus are identical then he may not hold this sentence true and this would serve to distinguish the theory containing the interpretive theorem from the one containing the counterfeit theorem. Only the interpretive theorem sustains the result that one can understand the language without knowing that Hesperus is Phosphorus.

3. Davidson's Principle of Charity

3.1 The Origins of DPOC

DPOC is a methodological rule for constructing a theory of meaning for a speaker of some language with which we are unfamiliar. Before giving an account of the content of DPOC, the status of the principle within Davidson's account should be remarked upon. He saw DPOC as "unavoidable"⁴², "forced on us"⁴³, "not an option but a condition of a workable theory".⁴⁴ He was uncompromising on the necessity of DPOC for interpretation. In light of this status, Davidson was prepared to say that "it is meaningless to suggest that we could fall into error by endorsing it".⁴⁵ Davidson saw the adoption of DPOC as the only way in which to solve the interpreter's problem of disentangling questions of what a speaker means from questions of what he believes. To see how questions of meaning and belief are entangled consider an example. On hearing the unfamiliar utterance "Gavagai" in the presence of a rabbit, we could attribute the speaker the belief that there is a rabbit present and translate the utterance as "There is a rabbit present". But on the ascription of a different belief, perhaps about rabbit parts rather than rabbits, we should translate his utterance as expressing something different, perhaps "there is an undetached rabbit part". How we interpret a speaker's words will depend on what belief we take him to be expressing and similarly, which beliefs we think he is expressing will depend on how we interpret his words. DPOC serves to "rule out a priori massive error on the part of the speaker"⁴⁶ and in doing so, fixes, to some degree, the beliefs of the speaker to be interpreted. Davidson believed that this was the only feasible way to go about a "systematic correlation of sentences held true with sentences held true"⁴⁷. Prior to pairing up the speaker's sentences with sentences of one's own language it makes no

⁴² 1967, pp.27.

⁴³ 1974b, pp.197.

⁴⁴ *ibid*

⁴⁵ *ibid*

⁴⁶ 1975, pp.168-9.

⁴⁷ 1974b, pp.197.

sense to talk of a speaker as making a mistaken assertion because one cannot say in one's own language over what it is that the speaker is mistaken.

A noteworthy feature of DPOC is the considerable development that it underwent between 1967 and 1991 and this is brought out in the account that follows. On its appearance in 1967, DPOC looked a lot like Quine's principle; involving the maximisation of agreement and also consistency. But it ended up as something quite different; richer but also intriguing in terms of what it might involve. By 1984⁴⁸ Davidson claimed to find the notion of maximising over beliefs, and particularly maximising agreement over beliefs, to be an unfortunate one. He professed that his goal had always been intelligibility rather than agreement. The notion of agreement dropped out of the formulation of DPOC by 1991 to be replaced by the notion of shared responses to the external world.

Davidson originally saw himself as applying "across the board"⁴⁹ a Principle of Charity (henceforth, any Principle of Charity that is not Davidson's reads POC) similar to that which Quine had applied to the observation sentences and logical connectives in his account of Radical Translation (see 4.4 for discussion of Quine's views on translation). Quine's maxim of translation had been that:

assertions startlingly false on the face of them are likely to turn on hidden differences of language.⁵⁰

The plausibility of the maxim resided in the translator thinking absurdity in the speaker less probable than his having made a bad translation. In constructing analytical hypotheses Quine warns us that "the more absurd or exotic the beliefs imputed to a people, the more suspicious we are entitled to be of the translations".⁵¹ This amounts to finding agreement on the obvious which for Quine included

⁴⁸ 2001b, pp.xvii-xix.

⁴⁹ 1973a, pp.136n.

⁵⁰ Quine, 1960, ch.2 Sect.13.

⁵¹ Quine, 1960, ch.2 Sect.15.

observation sentences and logical truths. Quine thought observation sentences were uncontaminated by theory because they are directly conditioned to sensory stimulation.

Davidson did not see how to draw a strict line between observational and theoretical sentences and so “had less choice”⁵² about the sentences to which DPOC was to be applied. Rather, he did not think there was any particular class of sentences to which its application should be restricted. Quine’s distinction between observational and theoretical sentences turned on the notion of “stimulus meaning”. The stimulus meaning of a sentence is a pairing of its affirmative and negative stimulus meaning. The affirmative stimulus meaning consists in all those patterns of sensory stimulation that cause assent to the sentence and the negative stimulus meaning all those that cause dissent. Quine thought there was an important distinction between the observation sentences, whose significance is determined by their direct conditioning to sensory stimulation, and the more theoretical sentences. The distinction implies that belief in an observation sentence could be justified extra-linguistically by appeal to sensory stimulation: that there can be a direct comparison of language and the world. Davidson did not accept the significance of the distinction because he found he could not make sense of the idea that sensory stimulation could justify a belief, even a belief of which the stimulation in question is the cause.⁵³ Davidson held that only beliefs can justify a further belief.⁵⁴ More broadly, he claimed he could not make any sense of the idea of direct confrontation between language and un-interpreted world and recognising this point can serve to allay a misunderstanding. The application of DPOC which Davidson envisages is not a two part process. There is not a first application of DPOC to observational beliefs followed by a second application to the rest. All the ascribed beliefs get the character they have from their place amongst a host of beliefs discerned using DPOC.⁵⁵

⁵² 1983, pp.151.

⁵³ See 1983.

⁵⁴ See 1983.

⁵⁵ Smith, Unpublished.

DPOC, as presented in 1967⁵⁶, instructed the interpreter to (1) maximise agreement with the speaker to be interpreted and (2) to maximise the self-consistency of the speaker.⁵⁷ Davidson saw the first instruction as crucial to our understanding what the speaker is talking about: throughout his writings he claims that we cannot attribute a speaker too much erroneous belief because it will leave nothing for them to be in error about. Davidson saw the second instruction as crucial to our understanding *the speaker*:

Crediting people with a large degree of consistency cannot be counted mere charity: it is unavoidable if we are to be in a position to accuse them meaningfully of error and some degree of irrationality. Global confusion, like universal mistake, is unthinkable, not because imagination boggles, but because too much confusion leaves nothing to be confused about and massive error erodes the background of true belief against which alone failure can be construed.⁵⁸

Both of the instructions for understanding a speaker must be understood in connection to Davidson's views on the interdependence of the content of beliefs. Davidson thought that to attribute one belief involved attributing many other beliefs:

A belief is identified by its location in a pattern of beliefs; it is their pattern that determines the subject matter of the belief, what the belief is about. Before some object in, or some aspect of, the world can become part of the subject matter of a belief (true or false) there must be endless true beliefs about that subject matter. False beliefs tend to undermine the identification of the subject matter; to undermine, therefore, the validity of the description of the belief as being about that subject. And so, in turn, false beliefs undermine the claim that a connected belief is false.⁵⁹

The case-by-case applications of DPOC ramify throughout the holistically attributed set of beliefs. The strength of Davidson's holism about belief attribution should not go unmentioned. It is a contentious claim that one needs to have endless true beliefs

⁵⁶ It is worth dating and considering the development of DPOC.

⁵⁷ 1967, pp.27

⁵⁸ 1970b, p.221

⁵⁹ 1975, pp.168-9

about some subject matter for it to figure in one's beliefs. A less contentious way of putting Davidson's holism about belief attribution is to say that for an interpreter to attribute a speaker an incredible belief it must have some place within the overall interpretation of the speaker. This latter formulation may not take the interpreter as far as attributing the speaker endless true beliefs about the subject matter. On either way of making the point, to attribute a false belief the interpreter must have well confirmed hypotheses about the meaning of the relevant utterances but find no place for the corresponding belief amongst his own pattern of beliefs.

Davidson gave the following example of a belief seeming to have a certain subject matter but the purported subject matter being obscured by related false beliefs. Centuries ago some people believed that the Earth was flat. One can interpret their belief as about the same earth that we have in mind because they share a sufficient degree of mundane belief about the Earth with us. Davidson ventures that if we construe these people as wrong in a sufficient number of their related beliefs, it becomes questionable whether or not they did in fact believe that the Earth, the same Earth which we have in mind, was flat. The possibility of identifying their belief as about the Earth rests on attributing them a background of "largely unmentioned and unquestioned true beliefs"⁶⁰ such as the beliefs that we live on the earth and that the face of the Earth is covered by land and water.

Davidson thought that as DPOC contained two instructions, labeled (1) and (2) above, it would determine no single theory of meaning for a speaker:

No single principle of optimum charity emerges; the constraints therefore determine no single theory.⁶¹

This lack of complete determination stems in part from there being:

⁶⁰ 1975, pp.168

⁶¹ 1967, pp.27

no completely disentangling questions of what the alien means from questions of what he believes.⁶²

The interplay of the two instructions is another block on complete determinacy because our finding a particular belief of the speaker's in agreement with our own may compromise the speaker's overall consistency and vice-versa. So there may be a choice between two equally acceptable theories of meaning: one of which settled on more agreement and compromised consistency and the other on less agreement and more consistency.

3.2 Coherence and Correspondence

By 1991⁶³ DPOC was explicitly split into two components: a Principle of Coherence and a Principle of Correspondence (henceforth, POCoh and POCor):

The Principle of Coherence prompts the interpreter to discover a degree of logical consistency in the thought of the speaker, The Principle of Correspondence prompts the interpreter to take the speaker to be responding to the same features of the world that he (the interpreter) would be responding to under similar circumstances. Both principles can be (and have been) called Principles of Charity: one principle endows the speaker with a modicum of logic, the other endows him with a degree of what the interpreter takes to be true belief about the world.

It is worth considering why Davidson could not make do with only one of these principles. POCoh is by itself insufficient for effecting the disentanglement of belief and meaning because at the outset the interpreter doesn't have a scheme of beliefs for the speaker on which to impose a degree of consistency. It will also be the case that there are many competing sets of consistent beliefs and so finding consistency alone would be inadequate. Neither will POCor suffice; if we attributed speakers a host of beliefs in accord with POCor but littered with inconsistencies, then their beliefs would come to imply any further beliefs because anything follows from inconsistent

⁶² *ibid*

⁶³ 1991, pp.211

premises.⁶⁴ We would struggle to establish that a speaker's beliefs concerned had any particular subject matter given enough inconsistency on his part. So the dual aspects of DPOC are mutually reinforcing. Notice that in this dual aspect formulation of DPOC talk of maximizing agreement has disappeared to be replaced by the notion of responding to shared aspects of the environment. This shift is the focus of section 3.3. Davidson's talk of speaker and interpreter responding to the same aspects of the world does not afford perceptual beliefs a special status in his account. All beliefs need to be set in the context of other beliefs to be made sense of. What he is interested in is the importance of occasion sentences in beginning interpretation. Occasion sentences are those whose causes of assent and dissent come and go as circumstances change. So these sentences can be paired with observable circumstances when stating their truth-conditions.

There are on any given occasion many aspects of the observable circumstances that an interpreter might respond to. The question of which aspects of the observable circumstances are relevant to the speaker is answerable only against the background of a developing theory of the speaker's beliefs. The interpreter takes the speaker to be responding to the features *he would* respond to and there is obviously some notion of salience involved here. But the interpreter need not circumscribe these features too narrowly at the outset. He may start with a quite general statement of the observable circumstances involving many conjuncts. The interpreter progresses by narrowing down and revising the statement of truth-conditions that appear on the right hand side of the T-sentence.

The idea that the speaker is "responding to" his environment calls for clarification. When one says in this context that a speaker is responding to features of his environment, what is it that the speaker is doing? At the very least we are claiming some causal relationship between our speaker and his environment; that some aspect

⁶⁴ Clearly it is not the case that the inconsistencies exhibited in most people's belief sets mean that their beliefs imply any further beliefs. This should be explained by thinking of people's belief sets as divided into subsets, and of normal people as having some inconsistent subsets but many consistent subsets.

of his environment is causing a piece of verbal behaviour. It may turn out that in some cases all we can say is that a speaker's verbal behaviour is caused by his interaction with his environment: the speaker may not be interpretable as expressing beliefs about his environment. In this minimal sense, a monkey "responds to" his environment. Monkeys are caused to make certain distinguishable sounds in response to different aggressors. Yet one might not want to attribute to a monkey a range of beliefs about his environment because a monkey cannot apply the concepts necessary for the possession of beliefs. The idea of an interpretable speaker responding to his environment, his having a collection of beliefs and expressing them, involves more than mere causal interaction. It involves the speaker's thoughts being directed at some aspect of the environment with which he is in causal interaction: his finding some aspect of the environment, with which he interacts, salient. For Davidson, a speaker's possessing beliefs also requires his having a notion of objectivity; his grasping the distinction between something's being believed and its being the case.

One also needs to clarify what is involved in the interpreter and speaker responding to the same aspect of the environment. It does not mean that the interpreter cannot find the speaker to have beliefs incompatible with his own. Both may be responding to the same object or event by judge something different of it. But the POCor entails that the interpreter finds the speaker to have many beliefs in accord with his own on pain of undermining the assumption that the speaker is responding to a particular feature of the environment, an object or event say, and that it is the very same feature that the interpreter is responding to.

In understanding POCoh, one needs to work out what Davidson had in mind regarding "a degree of logical consistency" in such statements as:

Needless to say, there are degrees of logical and other consistency, and perfect consistency, and perfect consistency is not to be expected. What needs emphasis is only the methodological necessity for finding consistency enough.⁶⁵

⁶⁵ 1983, pp.150

One does not usually think of logical consistency as admitting of degrees. A set of sentences is normally said to be logically consistent or logically inconsistent not more consistent or less consistent. Yet Davidson clearly states that “Coherence is nothing but consistency.”⁶⁶ Davidson needs to be careful because, just as a logically inconsistent set of premises implies any conclusion, a logically inconsistent set of beliefs will imply any further beliefs. One thing Davidson could mean is that the more coherent believer, the more logically consistent speaker, has more consistent subsets of beliefs. This would involve our isolating sets of beliefs, perhaps by context or time, and examining their consistency individually. The degree of consistency would then be the proportion of consistent sets amongst a speaker’s sets of beliefs.

Another element of the dual statement of DPOC above that it is crucial to elaborate on is the term “endows”, and the distinction between what is endowed by DPOC and what DPOC is. That which is endowed by DPOC is a consequence of its application not a part of the content of DPOC. When Davidson says that POCor endows a speaker with a degree of true belief, one might mistakenly take Davidson to mean that beliefs being largely true, or generally true, is part of the very nature of belief. Some commentators have read Davidson as saying that we cannot make any sense of someone’s having beliefs if we do not first recognize that veracity is in the nature of belief and there must be a preponderance of true belief in any speaker’s belief set. This way of reading Davidson gets things the wrong way around. It is from an investigation of the conditions that make interpretation possible that Davidson derived the conclusion that there has to be “a large degree of truth and consistency in the thought and speech of an agent.”⁶⁷ He thought DPOC necessary to interpretation and that which is endowed by applying DPOC is a product of interpretation:

It is an artifact of the interpreter’s correct interpretation of a person’s speech and attitudes that there is a large degree of truth and consistency in the thought and speech of an agent. ⁶⁸

⁶⁶ 1987b, pp.155.

⁶⁷ 1983, pp.150.

⁶⁸ 1983, pp.150. The point is owed to Smith, Unpublished

It is not the instruction of DPOC that we must find a speaker's beliefs to be largely true. DPOC does not contain any general conclusions about interpretable agents. DPOC is a methodological instruction applied case-by-case in giving us new interpretations; it is a procedural rule and does not involve any general claims about the nature of belief or believing agents.⁶⁹

3.3 The Move from Maximizing Agreement to Optimizing Intelligibility

Davidson became aware that his earlier talk of “maximizing agreement” could not be taken literally because the possible sentences to be construed are infinite in number, and no sense can be given to maximizing over infinities. Moreover, as the theory begins to take shape – as we accumulate more well-confirmed T-sentences – it makes sense for the interpreter to entertain error and allow for more or less likely mistakes.⁷⁰

Davidson came to think that:

there is no useful way to count beliefs, and so no clear meaning to the idea that most of a person's beliefs are true.⁷¹

Therefore, talk of maximizing agreement could only be taken as a hint at how the interpreter proceeds. Neither can the interpreter's instruction be expressed in terms of a general presumption in favour of the truth of the speaker's beliefs. This presumption would be compatible with each individual belief that we attribute to an interpretable speaker turning out to be false, and it is precisely this which Davidson denies is possible (see 3.7).⁷²

Davidson decided that rather than “maximize”, “a better word might be *optimize*”.⁷³ In optimizing rather than maximizing the agreement one finds with a speaker one might

⁶⁹ *ibid*

⁷⁰ 1973a, pp.136.

⁷¹ 1983, pp.138-9.

⁷² This point is owed to McNeill, 2004.

⁷³ 2001b, pp.xix.

think, for example, that speakers are unlikely to be mistaken about their own feelings and very likely mistaken on questions of highly theoretical physics. Similarly, one might think speakers more likely to be mistaken in their perceptual beliefs when the weather is foggy. Optimizing agreement involves the interpreter making sense of explicable error, and also explicable successes.

Charity prompts the interpreter to maximise the intelligibility of the speaker, not sameness of belief. This entails, as Lewis says, that interpretation must take into account probable errors due to bad positioning, deficient sensory apparatus, and differences in background knowledge.⁷⁴

Davidson's views on how we go about making sense of error are a matter returned to in sections four and five.

In 1984 Davidson wrote:

the Principle of Charity counsels us quite generally to prefer theories of meaning that minimize disagreement. So I used to put the matter in the early essays, wanting to stress the inevitability of the appeal to charity. But minimizing disagreement, or maximizing agreement, is a confused ideal..... My point has always been that understanding can be secured only by interpreting in a way that makes for the right kind of agreement. The "right sort", however, is no easier to specify than to say what constitutes a good reason for holding a particular belief.⁷⁵

Davidson thought the kind of agreement (and also disagreement) that the principle counseled us to find between speaker and interpreter was sensitive to both the context and nature of belief in question and hence, nothing more general could be said about "the right sort" of agreement.

⁷⁴ *ibid*

⁷⁵ *ibid*

DPOC does not prompt the interpreter to find sameness of belief but rather “prompts the interpreter to maximize the intelligibility of the speaker”.⁷⁶ It is interesting to consider what “making intelligible” could involve. What Davidson seems to have meant by “making intelligible” is making beliefs and utterances explicable. The project of making a speaker’s beliefs and utterances explicable could give Davidson’s account a far wider remit, because a wide variety of information might be relevant to making someone’s beliefs explicable. Quine’s employment of a POC was not directed at making speakers more explicable. Quine’s POC involved maximizing agreement on observation sentences and reading our logic into the speaker’s language. It was probably this notion that Davidson himself first had in mind. But if Davidson no longer had Quine’s POC in mind, as talk of making people intelligible and explicable would suggest, then this raises questions about what Davidson’s account of interpretation might involve. More specifically, talk of maximizing, or rather optimizing, intelligibility could place interpretation in the setting of the wider empirical investigations of psychology, linguistics, optics and other disciplines. Empirical psychology has much to say on the matter of making people’s errors and successes explicable, and making their beliefs intelligible. Empirical linguistics has much to say with regard to explaining how people come to speak the language that they do. The employment of these sources of information in making a speaker’s beliefs and utterances intelligible might make room for a lot more difference in belief between speaker and interpreter. If, for argument’s sake, our speaker was in possession of peculiar sensory apparatus, it may take serious empirical work for the interpreter to be in a position to make the speaker’s beliefs intelligible.

Davidson, as quoted above, was concerned to make room for Lewis’ injunction that interpretation be sensitive to such factors as “probable errors due to bad positioning, deficient sensory apparatus, and differences in background knowledge”.⁷⁷ Davidson was concerned with these factors in connection with making a speaker explicable. It is an interesting question (see 4.1, 4.4, 4.5, 5) whether knowledge concerning the

⁷⁶ *ibid*

⁷⁷ 2001a, pp.xix.

factors Lewis mentions and related factors might be relevant to making good the assumption that the speaker is responding to the same aspects of the world as their interpreter. Davidson, however, equivocates over whether the employment of such knowledge is part of good interpretation. Shortly after his concession to Lewis, he says:

I do not take this to prove that the evidential base on which I depend is not in theory adequate. I grant, however, that it may not be.⁷⁸

3.4 Assertion: The Very General Attitude of Holding True

How we construe a speaker's words will depend on what we take him to believe. DPOC can disentangle the question of what a speaker's words mean from the question of what he believes by holding belief fixed (see 3.5). Davidson's view is that the knowledge relevant to ascribing beliefs can come only with the interpretation of someone's words⁷⁹:

If all we know is what sentences a speaker holds true, and we cannot assume that his language is our own, then we cannot take even a first step towards interpretation without knowing or assuming a great deal about the speaker's beliefs. Since knowledge of beliefs comes only with the ability to interpret words, the only possibility at the start is to assume general agreement in beliefs.⁸⁰

Davidson's idea was that prior to understanding a speaker's utterances and prior to knowing what he believes, we can nevertheless identify the sentences that a speaker holds true:

A good place to begin is with the very general attitude of holding a sentence true.... It is an attitude an interpreter may plausibly be taken to be able to identify before he can interpret,

⁷⁸ 2001a, pp.xx.

⁷⁹ I try to put Davidson's claim that the knowledge relevant to discerning someone's beliefs comes only with interpreting their words under pressure (see 4.1, 4.4, 4.5, 5).

⁸⁰ 1974b, pp.196-7.

since he may know that a person intends to express a truth in uttering a sentence without having any idea *what* truth.⁸¹

One then has a list of sentences held true by a speaker but we know neither what beliefs the sentences express nor what the sentences mean. Despite what Davidson says in the quotation above, one does not *know* that a speaker is holding sentences true. One assumes that one is observing assertion, the very general attitude of holding true. Davidson does not furnish the interpreter with a method by which to know, in the radical case, that one is observing speakers holding sentences true.

Davidson may have been relying on the method described by Quine.⁸² Quine thought that to find out the expression of assent in some unfamiliar speaker one should echo the speaker's utterances back at them in unchanging circumstances. The native may have two such short words as our "yes" and "no", of course it is conceivable that they do not. One would expect to be answered by their word for "yes" when the observable circumstances seemed unchanging to us. But even if the native does have two such words which they use at the times when we would use "yes" and "no", one cannot assume that the two words mean precisely the same as ours do. Quine's method then involves the assumptions that the native's have two short words for assent and dissent, and that natives will assent and dissent to their own sentences repeated back to them. Quine thought that one could spot the equivalent behaviour in advance of translating the speaker's words. Some sentences may be too complex to allow for the uncomplicated responses Quine relied on. Sometimes an asserted sentence is met with no clear assent or dissent on its being queried. Davidson's assumption that we are observing assertions is more risky than Quine's claim to be able to isolate signs of assent and dissent. Quine is seeking out only two forms of expression but assertion can take a whole host of forms and each particular assertion is a different expression. It seems that prior to working out what a speaker believes and what his words mean (and getting a grip on the force of his utterances) then there is always room for doubt about whether we are observing assertion. Establishing

⁸¹ 1973a, pp.135.

⁸² Quine, 1960, ch.2.

whether some creature is performing linguistic actions at all might involve us in learning something about, or making assumptions about, their psychology and Davidson would agree on this point.

Even if one credited oneself with being especially good at recognizing assertive behaviour in unfamiliar speakers, one still has a problem. One of Davidson's concerns is that the evidence for Radical Interpretation should be evidence available to any third person observer. So accessing the evidence cannot require of the interpreter any skill that is too highly specialized. There is also the question of how to justify limiting our interpreter to assertions at the outset. One might want to allow for a broader set of forces accompanying the utterances the Radical Interpreter hears. Allowing for a broader range of forces such as questions, commands and the rest, is not precluded by the entanglement of belief and meaning. The assumption that we observe, for example, questions and assertions would also be neutral as to the finer content of speaker's beliefs and the meaning of his utterances. There should be some further justification for the assumption that we observe assertions rather than jokes, or questions. If the only justification offered for starting with assertions was that assertions are neutral as to the precise belief expressed and the meaning of the utterance, then this is insufficient because other forces of utterance may also be neutral.

But there are further reasons why Davidson takes assertions as his starting point. The notion of someone's holding something true is supposed to be basic in two respects. Firstly, one might expect an interpreter to be able to grasp that a speaker is affirming or stating something prior to grasping that someone has one of the other more sophisticated attitudes towards that utterance. One might be dubious that assertive behaviour is basic in this sense. If the interpreter is capable of telling that a speaker is making an assertion then he might be capable of discerning further interesting facts about the force which a speaker attaches to his utterances. Secondly, we might think assertion a more basic, less sophisticated, achievement for the speaker himself. But

this explanation is also dubious. Perhaps demand behaviour - "Bring me a slab!" - is more basic than assertive behaviour - "There is a slab!".

The most convincing reason for starting with assertions is that they are all that is needed to begin constructing a theory of meaning. It is assertions rather than utterances of other force that have truth-conditions. Nothing more than its truth-condition need be paired with an assertion to arrive at its meaning; no theory of force is required. Davidson's thought may have been that the early introduction of other forces of utterance is an additional clutter to his methodology. Other forces of utterance are set aside to be dealt with later on the basis of a theory of the meaning for the speaker's assertions. One might still be dissatisfied if one thought that allowing the interpreter access to a wider range of forces of utterance from the start had significant benefit, or if one did not think that the meaning of utterances of other forces could be explained by a modification of one's account of the meaning of assertions.

3.5 Holding Belief Fixed to Solve for Meaning

A speaker makes an assertion as a result of the interplay of two things: what they believe to be the case and what the sentence means. If we could pin down either of these factors then we could solve for the other. If we knew what someone meant by an assertion then we could work out what they believed and similarly, if we knew what belief they expressed we could work out what they meant. In taking a speaker to be responding to the same aspects of the world that the interpreter would, Davidson saw the interpreter as pinning down the speaker's belief. Without DPOC, Davidson thought we would have to find out what a speaker believed before we could start working out the meanings of their sentences. But working out what the speaker believed would involve us interpreting their utterances. Davidson prohibited his interpreter from assuming anything about the speaker's beliefs and meanings prior to engaging in interpretation. There is nothing in this dictum that implies the inadmissibility of knowledge about a speaker's beliefs acquired prior to

interpretation; it is rather that Davidson did not think that such knowledge was possible prior to interpreting the speaker's words and vice-versa. One should be alert to the possibility that there is knowledge available prior to interpretation that can tell us something about a speaker's beliefs and does not involve us in interpreting their words, and vice-versa.

Davidson takes the fact that DPOC can effect a disentanglement of questions of meaning and questions of belief as justification for its adoption. The speaker's beliefs are fixed in close accord with the interpreter's through his finding their beliefs to be consistent and directed towards the same features of the world as his would be. There is no complete disentanglement on this approach as beyond a certain point there is no deciding between a speaker's using words as we do but with odd beliefs and their using words differently.⁸³ Davidson does not enter into much discussion of what degree of indeterminacy resides here and by what factors the degree of indeterminacy is fixed.

In applying DPOC the interpreter assigns truth-conditions to a speaker's utterances that make the speaker right when "plausibly possible"⁸⁴, according to the interpreter's surmising of the observable circumstances. Davidson was skeptical that anything general could be said about the kind of agreement that could be "plausibly" found between speaker and interpreter (see 3.3). We take a speaker's holding a sentence true as *evidence* for that sentence's truth in the observable circumstances.⁸⁵ Similarly, Davidson was skeptical that anything of a general nature could be said about the evidential relation between something's being asserted by a speaker and its being true on the occasion of utterance.⁸⁶

The interpreter can make room for more radical differences in the circumstances under which some sentences are held true than others. Some sentences have causes of

⁸³ 1968, pp.101.

⁸⁴ 1973a, pp.137.

⁸⁵ 1974a, pp.152.

⁸⁶ See 1974a, 2001b p.xvii-xix

assent and dissent that come and go whilst some sentences are assented to or dissented to come what may. A simple addition may be held true in any circumstance but the truth of a sentence about the weather is sensitive to the circumstances. Whilst neither the occasion sentences nor standing sentences were epistemologically privileged on Davidson's view, there is a priority in the order of interpretation. The interpreter enters into the speaker's language by giving the truth-conditions of a host of occasion sentences that express beliefs about the observable circumstances.

The interpreter should take into account individual differences among speakers and their environment that may have caused discrepancies in the circumstances under which a sentence is held true:

We get a first approximation to a finished theory by assigning to sentences of a speaker conditions of truth that actually obtain (in our own opinion) just when the speaker holds those sentences true. The guiding policy is to do this as far as possible, subject to considerations of simplicity, hunches about the effects of social conditioning, and of course commonsense, or scientific knowledge of explicable error.⁸⁷

With this mention of the use of scientific knowledge we are given a hint that Davidson is not delimiting the evidential base for interpretation in as narrow and behaviouristic a way as Quine had done in his account of Radical Translation. However, many of Davidson's statements of the available of evidence (see 4.1, 4.4) do not take these points seriously. With his mention of hunches about conditioning and common sense, one gets the uncomfortable feeling that there are elements creeping into interpretation which, even if pragmatically justifiable, are not substantiated by evidence. They also introduce an element of indeterminacy as there may be more than one defensible view of what our common sense amounts to and what it licenses.

3.6 Testing a Theory of Meaning

⁸⁷ 1974b, pp.196.

As the interpreter accumulates a body of hypothetical T-sentences he can begin testing his theory. The T-sentences are tested as an overall theory of what the speaker's words mean and as the theorist gathers more evidence so the level of speculation diminishes. The hypotheses about the beliefs and meanings of the speaker are constantly subject to revision:

Each case tests a theory and depends on one.⁸⁸

The interpreter tries to observe as many occasions of the speaker's utterances as possible. As he tests the T-sentences collectively, he hopes to shed light on the structure of the speaker's sentences. The repetition of words, strings of words and structural features of sentences are the data for axiomatisation. A hypothetical T-sentence maintains its place in a scheme of interpretation if it is not disconfirmed by the evidence or replaced by an equally well-confirmed but simpler T-sentence. Attributing one belief to a speaker involves attributing him many others and corroboration for the speaker having any of those other beliefs must be found elsewhere in the theory. The logical relations that hold between beliefs and the holistic nature of belief ascription mean that if a speaker and interpreter hold the same sentence true and share the same logic⁸⁹ they must have many more beliefs in common. So evidence that the speaker rejects one of the beliefs implied by the belief set the interpreter has attributed him is evidence against that attribution.

3.7 Making Disagreement Intelligible Against a Background of Agreement

Davidson's approach is often criticized for loading the deck in favour of too much agreement between speaker and interpreter. His response was that his method is not intended to eliminate disagreement. Rather his method accounts for disagreement; making disagreement between speaker and interpreter intelligible requires grounding

⁸⁸ 1970b pp.221.

⁸⁹ 1970b pp.221: "we have no choice but to read our own logic into the thoughts of a speaker".

it in agreement. There must be some agreement in subject matter between interpreter and speaker for there to be something about which they disagree. For Davidson, interpreting a speaker as pronouncing on a certain subject matter requires the interpreter to find much truth and in his surrounding beliefs.

Davidson was anxious that one should not think of DPOC's virtues as resting on an assumption about human intelligence that could turn out to be ill founded:

The methodological advice to interpret in a way that optimizes agreement should not be conceived as resting on a charitable assumption about human intelligence that might turn out to be false. If we cannot find a way to interpret the utterances and other behaviour of a creature as revealing a set of beliefs largely consistent and true by our own standards, we have no reason to count that creature as rational, as having beliefs, or as saying anything.⁹⁰

DPOC is justified by its indispensable role in disentangling questions of belief and meaning. The degree of truth and consistency with which it endows the speaker is justified by the thought that disagreement between interpreter and speaker is only intelligible against a backdrop of agreement and that we weaken the intelligibility of belief attributions when we fail to find a consistent pattern in the speaker's beliefs:

To see too much unreason on the part of others is simply to undermine our ability to understand what it is they are so unreasonable about.⁹¹

A background of mundane and widespread agreement is imposed on an interpretation scheme by repeated applications of DPOC to behavioural evidence (this point is explained at length in 4.1). The holistic nature of belief attribution implies that agreement in subject matter between interpreter and speaker cannot be established independently of a high degree of agreement in opinion. Beliefs are individuated by their position in a network of beliefs. Davidson was keen to emphasise the masses of

⁹⁰ 1973a, pp.137.

⁹¹ *ibid*

agreement on trivial matters that ensures interpreter and speaker are thinking about the same content. He thought that this underlying agreement easily escapes attention.

One might wonder why the interpreter finds a degree of truth in the speaker's beliefs rather than systematically rejecting them as false. This point turns on the incompatibility of, on the one hand, the interpreter taking the speaker to be responding to the same aspect of the observable circumstances and doing so with a degree of consistency, and on the other, taking the speaker's beliefs to exhibit a high degree of falsity. The interpreter finding enough of the speaker's related beliefs false undermines the instruction that he take the speaker to be responding to the same aspect of the observable circumstances that he would (see 4.1 for further discussion). By taking the utterances with which he is presented as merely expressions of false belief the interpreter does not gain himself any evidence for a T-sentence. The assumption that the utterance "Gavagai" is false in the presence of a rabbit does not imply the hypothesis that:

"Gavagai" is true if and only if there is not a rabbit.

It does not imply this hypothesis because the assumption merely that the utterance is false, that it expresses a false belief, does not establish any connection between the utterance and the observable circumstances. On assuming merely that the utterance is false, the interpreter would do as well to pair "Gavagai" with any truth-conditions that do not obtain. What the interpreter needs to establish a link between the verbal behaviour he observes and the observable circumstances is his speaker to be responding to the observable environment. This is what is ensured for the interpreter by POCor. But once we accept POCor then the idea that we could take our speaker to exhibit a high degree of false belief is a non-starter: the holistic nature of belief attribution entails that for the interpreter to make sense of the speaker's responding to the same feature of the world as the interpreter would, speaker and interpreter must exhibit much background agreement in belief.

4. The Available Evidence and the Status of Charity

4.1 DPOC as a Necessity of Interpretation

Davidson's argument for adopting DPOC was that we must adopt it. If we refuse to then, he claimed, interpretation becomes impossible. So if his opponent rejects DPOC but believes interpretation to be possible then he must provide an alternative account of how it can be carried out. With the correct understanding of Davidson's views in hand, the aim of this fourth section is to undermine the claim that good interpreters must adopt DPOC. Davidson's idea was that whatever the utterances with which an interpreter is initially presented mean, and whatever some speaker believes in uttering them, the utterances can only provide an evidential basis for a T-sentence if we accept that the utterances are responses to the observable circumstances. More specifically, we must accept that they are responses to those aspects of the circumstances the interpreter finds relevant or salient; the aspects *he would* respond to. The holistic nature of Radical Interpretation entails that this sharing of subject matter will amount to a degree of shared related beliefs. In accepting these assumptions we can begin to pin down what the speaker believes, and in doing so determine what it is his words mean.

In the interpreter's first encounters with the speaker, the observable circumstances on the occasions of utterances provide evidence for interpretations of the utterances only if the interpreter takes the observable circumstances to be realizations of the truth-conditions of the utterances under interpretation. This is not to say that Davidson's interpreter can attribute no false beliefs in the early stages of interpretation. The point is that when the interpreter attributes false beliefs in the early stages of interpretation, he undermines the instruction that he take the speaker to be responding to the same features of observable circumstances. The interpreter and speaker sharing a subject matter, namely some aspect of the observable circumstances, is contingent on some agreement in their beliefs about that subject matter. This agreement is only secured as the interpreter pairs the speaker's utterances with sentences of his own. Too much

early disagreement leaves no shared subject matter on which interpreter and speaker can disagree. Thus the POCor which instructs the interpreter to take the speaker to be world directed is said to endow a degree of true belief. Davidson was, of course, not ruling out that the interpreter could impute false beliefs to the speaker. Rather he was claiming that the interpreter can only begin to make sense of imputing true or false beliefs against accumulated T-sentences formed by applying DPOC. Davidson's view was that adopting DPOC was the only feasible way to construct a truth-conditional theory of meaning for a speaker. Davidson thought that if an interpreter applied DPOC to the behavioural evidence on a case-by-case basis and satisfied the formal constraints on a theory of meaning, he has done all that good interpretation requires.

A familiar objection to Davidson's reasoning⁹² is that the interpreter would do just as well to start off assuming that the utterances to which he is exposed are false. It is not an assumption of Davidson's method that the first utterances with which the interpreter is confronted must be counted true but application of the POCor and the holistic nature of belief ensure that the interpretable speaker must enjoy a degree of true belief. Nevertheless, let us consider what this assumption of falsity which has been proposed as a challenge to Davidson could amount to.

The falsity of an uttered sentence in and of itself implies neither that there is a particular connection nor that there is any connection at all between the sentence and the observable circumstances. So why would Davidson's opponent think that the mere assumption that an utterance is false could earn the interpreter any content for his theory to work on? Imagine a principle of "Uncharity" which instructed the interpreter to find a link between utterances and the observable circumstances such that when the interpreter hears the unfamiliar utterance "Gavagai" upon the appearance of a rabbit he takes the utterance to be false. It might then seem natural to adopt the hypothesis that:

"Gavagai" is true if and only if a rabbit has not appeared.

⁹² For example, McGinn, 1977.

But on the assumption merely that the utterance is false the interpreter is entirely unjustified in adopting such a hypothesis. The Uncharitable Interpreter is forging a link between the utterance and the circumstances where, on the assumption merely that the utterance is false, no such link obtains. The relationship that the interpreter finds between an utterance, about which all he knows is that it is false, and the observable circumstances should be a spurious one; not the negation of the relation that would obtain if the utterance were true. The utterance merely being false does not imply that it has anything to do with rabbits and their appearance or otherwise. It is only by invoking a POCor that one links the utterance to the circumstances, and we have seen Davidson's reasons for thinking that applying POCor endows the speaker with a degree of true belief. The Uncharitable interpreter has no more reason to assign "Gavagai" the truth-condition that a rabbit has not appeared than he has to assign to it any other truth-conditions that do not obtain. He has no complement to the Charitable interpreter's POCor but to provide him with one would be to endow the speaker with some true belief. When one offends against the POCor one seems to sever the connection between the subject's utterances and his environment, and when one adopts a POCor in interpretation one thereby endows the speaker with some true belief.

DPOC comes down to making such a link between speaker's utterances and the observable circumstances intelligible if interpretation is to proceed. When the interpreter takes the speaker to be responding to some feature of his environment, he endows the speaker with enough true belief to secure a shared subject matter and enough consistency to ensure that we can count the speaker as having beliefs at all. It is in this way that Davidson saw DPOC as indispensable to interpretation, rather than a means to choose between equally possible, rival theories of interpretation. Strictly in the first stages of interpretation it makes no sense to talk of agreement between the beliefs of the interpreter and the speaker. Beliefs and meanings are brought out as the interpreter begins to impose structure on accumulated evidence. Even the interpreter's methodological precept that the observable circumstances provide direct evidence for

the truth-conditions of any particular utterance is a hypothesis about that utterance that may be discarded when we have worked our way into the speaker's language. Observable circumstances are not direct evidence for the truth-conditions of mathematical or logical statements for example. It is occasion sentences that serve as the entering wedge into the speaker's language and the crucial part of Davidson's account is his proposal for how an interpreter can enter into the speaker's language:

I propose that we take the fact that speakers of a language hold a sentence to be true (under observed circumstances) as prima facie evidence that the sentence is true under those circumstances. For example, positive instances of "Speakers (of German) hold 'Es schneit' true when, and only when, it is snowing" should be taken not only to confirm the generalization, but also the T-sentence, "'Es schneit' is true (in German) for a speaker x at time t if and only if it is snowing at t (and near x)".⁹³

Davidson takes the evidential relation between utterances, their truth-conditions and the observable circumstances to be the following. An utterance in observable circumstances is evidence that the observable circumstances realize the truth-conditions of that utterance. It seems to me that there are different relations between utterances and their conditions of truth that might be serviceable as grounds for interpretation. One such relation is to take the utterance on an occasion as one source of defeasible evidence amongst several sources of evidence for truth-conditions. Davidson allows that the prima facie evidence is defeasible but only by more of the same, namely verbal and perhaps non-verbal behaviour in observable circumstances. The remainder of this thesis attempt to explain why Davidson had this single-source conception of the evidence and to explore what the alternative suggestion might come to.

Another evidential relation between utterances and their truth-conditions one might consider is to take the utterance on an occasion only as evidence that the utterance is a response to the observable circumstances, and then to resist Davidson's move from agreement in subject matter to agreement in belief (this suggestion is discussed in

⁹³ 1974a, pp.152

4.3). One might wonder whether, even in the first instances, a speaker's holding a sentence true in observable circumstances might be evidence for a T-sentence even if the sentence actually turns out to be false to those circumstances. This is possible on Davidson's method: we can use an utterance as evidence for a T-sentence and then, retrospectively, against the background of accumulated T-sentences see that the utterance was in fact false. It is a question of whether taking the speaker to be responding to his environment, being world directed, implies that the interpreter must at the outset attribute the speaker beliefs that are true by his lights. Davidson's problem with attributing false beliefs in the early encounters is the following. If in first encounters we take our speaker to be world directed but attribute him a false belief, say the belief that there is a white rabbit when in the presence of a black rabbit, then we may undermine our attribution to the speaker of a belief about the very same rabbit to which the interpreter is responding.

Within the framework of Radical Interpretation, making sense of a speaker's engaging with a particular subject matter involves attributing him some true collateral beliefs. If one could gather evidence about the subject matter of a speaker's beliefs independently of Radical Interpretation then one might plausibly claim that establishing a link between an occasioned utterance and the observable circumstances does not involve the attribution of collateral true belief. We might then take the utterance as one piece of defeasible evidence amongst others for the truth not of the sentence uttered but only of some T-sentence. A defeasible piece of evidence here meaning roughly: evidence, the importance of which may diminish in light of other evidence.

Davidson thought that if the utterance is to be evidence for the T-sentence then the utterance must also be evidence of its own truth. We have seen that in Radically Interpreting a speaker we ascribe beliefs to a speaker and meanings to his words, and that in ascribing him one belief we ascribe many. From this point about the holistic nature of ascriptions of belief and meaning we have been alerted to the fact that our interpreter must be careful in making his ascriptions not to ascribe such a degree of

false belief as would obscure his speaker responding to a shared subject matter. The possibility mentioned above and which this thesis focuses ^{on} is that an interpretable utterance need not be counted as evidence of its own truth in order to count as one piece of *mere evidence* amongst other sources of defeasible evidence for a T-sentence.

To take an example, imagine an interpreter in the radical situation who hears the words “s’ti gniwons” when it appears to him to be snowing. Our interpreter may go on to form the hypothetical T-sentence “<s’ti gniwons> is true if and only if it’s snowing”. One might take the utterance <s’ti gniwons> in the observable circumstances as evidence for this hypothetical T-sentence without taking the utterance as evidence for its own truth in those circumstances. One might take the utterance to be one source of defeasible evidence for the truth-conditions of the sentence, and claim that there are other sources of defeasible evidence. The question of whether the utterance is true would then be an independent question. The link between an utterance and the observable circumstances that obtains when an utterance is true is only one link between the speaker and the observable circumstances but there are others. It is not the case that unless we take an utterance as evidence of its own truth then we cannot make use of the observable circumstances of an utterance in constructing a T-sentence. The choice that Davidson presents us with is between, on the one hand, DPOC and, on the other, putting the content of a speaker’s beliefs entirely beyond the interpreter’s reach. This choice may be a misleading one. It may be misleading in its suggestion that if we do not employ a POC or and, in doing so, endow our speaker with much true belief, we are depriving a theory of interpretation of any connection to the conditions under which an utterance is made. We should be open to the idea that the connections between a speaker and his environment can be investigated in a range of ways, and that these investigations produce evidence that should be made available to and employed by the interpreter.

What is up for consideration is the claim that there are kinds of evidence relevant to interpreting a speaker which play only a peripheral role in Davidson’s methodology.

This contention is not novel in itself⁹⁴ but the rationale provided here for this conclusion is different and offers a new kind of challenge to Davidson. There are fundamentals on which this challenge to DPOC can be distinguished from some other prominent challenges and some preliminary remarks will help place this thesis in that context (4.2, 4.3 and 4.5 distinguish my position from the nearby relatives. 4.4 explains the source of my disagreement with Davidson). The account of DPOC given in section three is more comprehensive and accurate to Davidson than those to be found in Grandy, McGinn or Lewis.⁹⁵ My aim is not to offer an alternative master principle of interpretation, which is the issue to which Grandy directs his attention (4.2). There was some discussion in this section of the possibility of securing shared subject matter without shared belief, and McGinn pursues this route via a causal theory of belief (4.3). He claims that building a causal theory of belief into interpretation will turn the trick but it is suggested here that this is not the right way to go. Whilst the challenge presented here is congenial to Lewis's broader view of the available evidence it does not rely on the kind of reductionism implied by Lewis's paper (4.5).

It was Davidson's contention that without DPOC we would have to know something about a speaker's beliefs to begin interpreting their words. Since such knowledge is deemed unavailable by Davidson we must employ DPOC. It was mentioned above that there may be more sources of defeasible evidence for T-sentences than Davidson allows. We should not share Davidson's pessimism about the relevance of knowledge available prior to interpretation to helping discern a speaker's beliefs. Davidson's claim that no useful evidence or knowledge is available prior to interpretation should be put under the spotlight. There are sources of evidence marginalized by Davidson and they may constitute the kind of knowledge he deems unavailable. If one can make it plausible that knowledge helpful to discerning meanings or beliefs may be available prior to interpretation then one undermines Davidson's strong claims about

⁹⁴ Grandy, 1973. Lewis, 1974. McGinn, 1977.

⁹⁵ The differences are explained in the respective sections below. Grandy's argument is a response to Quine rather than DPOC. It should be borne in mind that these papers were published whilst DPOC was undergoing significant change and their characterisations of DPOC may have been more accurate to Davidson's writings at the time of publication.

the necessity of DPOC to interpretation. What follows in section four is a challenge to the claim that there is no evidence available prior to interpretation that is relevant to the disentanglement of meaning and belief. By making it plausible that we can gather such information about a speaker's beliefs or language as would pin down either to a significant degree, one undermines Davidson's claim that DPOC is a *sine qua non* of interpretation. If such evidence is admissible, and available in impressive quantities, then DPOC would take on a less central role in interpretation.

The major source of the extra evidence alluded to is the natural, human sciences. Empirical inquiries into psychology, linguistics, optics, and other branches of the sciences of human nature are concerned with speakers and their relations to environments. These sciences highlight the similarities and differences between humans in the way they interact with their environment and come to speak a language. Of course research in these fields is often concerned with similarities and differences between humans and other organisms, and these findings may also be revealing. It seems to me highly plausible that learning about the way a speaker interacts with his environment will be illuminating of, and a constraint on, the kind of beliefs which it is feasible to ascribe him. If we discover through empirical investigation that some speaker is a lot like us in the way that he interacts with his environment then this will be evidence that the speaker responds to the same aspect of the observable circumstances to which we, as interpreters, respond. We may then proceed by hypothetical maxims to interpretations that are much like those the POCor would imply. But our finding out about such similarity and employing that knowledge would not serve as a justification of DPOC as a precondition of interpretation. We might, on the contrary, find out that our speaker is dissimilar in some relevant respect. If a speaker could not interact with his environment in the way that we do this may put the formation of certain beliefs beyond him.⁹⁶ A crude example could involve a speaker who had no sense of smell. Our attaining the knowledge that an unfamiliar speaker suffered from this deficiency would be

⁹⁶ This same point could be made positively. If the speaker interacts with his environment in a way that we cannot, there may be beliefs that he could form that are beyond us.

evidence against the speaker responding to the some features of the observable circumstances that we do and holding the related beliefs. This information seems to be available prior to interpretation and available in ever more sophisticated varieties, and it is relevant to what the speaker's beliefs are.⁹⁷

There are several places where Davidson makes piecemeal concessions to a broader conception of the evidence. But he nowhere notices the full significance that marshalling such evidence could have with regard to the status of DPOC. Where such evidence serves to corroborate similarities between speaker and interpreter in terms of the way they respond to their environment then it can serve to make good the POCor. More generally such evidence as is relevant to pinning down a speaker's beliefs and is available to the interpreter can fulfill the role of holding belief to some degree fixed – precisely the role for which DPOC was invoked. Despite a few limited concessions, it is clear that Davidson did not agree that there were sources of evidence other than verbal behaviour available prior to interpretation. If Davidson had recognized that more evidence was both available and useful to the interpreter then he would not have proceeded in the directions he did. Had Davidson seen the sharing of subject matter between interpreter and speaker that the POCor instructs us to find as something we might investigate, rather than assume, then he would not have sought to underpin the POCor with his discussion of Triangulation. Moreover, if Davidson had recognized that there was evidence available to the interpreter relevant to effecting the disentanglement of meaning and belief then he would have not have defended DPOC as a *necessary condition* on constructing a theory of interpretation.

4.2 Grandy's Principle of Humanity

Grandy famously advocated a move from a POC to a Principle of Humanity. The POC with which he found fault came from Quine and several of his criticisms are highly plausible. Here is the principle that Grandy advocated:

⁹⁷ A more sophisticated example would be the recent mapping of the cochlea and the importance of this structure to the subject's placing of sounds.

If a translation tells us that the other person's beliefs and desires are connected in a way that is too bizarre for us to make sense of, then the translation is useless for our purposes. So we have a pragmatic constraint on translation, the condition that the imputed pattern of beliefs, desires, and the world be as similar to our own as possible. This principle I shall call the *principle of humanity*.⁹⁸

If we find DPOC unsatisfactory, we might think there is some other aprioristic principle of interpretation that we should adhere to. The purpose of this thesis is not to offer an alternative principle to Davidson's but to explore and question the status we should afford such principles, and to propose a more open-armed methodology; to relax Davidson's own approach. But it is worth saying something about Grandy's Principle of Humanity (henceforth, POH) because it is often thought to be an improvement on DPOC. As such, one might think it remedies the complaints made in the present thesis against Davidson. This thesis presses the claim that we should be cautious of principles that have the status not of hypothetical maxims of investigation but of strong, a priori conditions on interpretation

The POC which Grandy criticises is Quine's rather than DPOC, so one should be alert to the differences between the two. The POC that Grandy was challenging instructed the interpreter to maximize agreement on obvious truths; Quine included observation sentences and logical statements in this category. Davidson did not accept Quine's distinction between the observation sentences and the rest, and was concerned to optimize *intelligibility* rather than agreement. Several of Grandy's criticisms of Quine's POC do not apply to DPOC.

Looking at the quoted statement of the POH, one might notice that it instructs the translator to impute a higher degree of similarity than is required by the condition that a translation be intelligible. It is true that a translation that is too bizarre for the translator to make sense of will be useless. But this does not entail that we plump for the translation makes the subject *as similar as possible* to ourselves. It only follows

⁹⁸ Grandy, 1973, p.443.

that we discard translations that make someone unintelligible. There are degrees of similarity that a translator can find between himself and his subject. What degree of similarity the translator finds should be sensitive to the evidence he collects. It is implausible to suggest that only a translation that makes the subject as similar to us as possible could further our understanding of a subject. It is a plus point of DPOC that it does not impose such cognitive parochialism.⁹⁹

An interesting aspect of Grandy's paper is the very broad role he envisages for a POH. He introduces the POH in the context of translation, as a rival to Quine's POC. But he also thinks of POH as a means by which, once we already have a theory of a speaker's beliefs and desires and a translation manual, we might go about predicting a speaker's behaviour. Grandy thinks a POH is needed to prevent a translator attributing inexplicably held beliefs: beliefs that may be true by the translator's lights but inaccessible to the speaker. So in assessing Grandy's principle one is confronted by the question of what counts as a good reason for a subject's having a belief. One should bear in mind that Davidson saw DPOC as a necessary condition on getting a theory of the speaker's beliefs at all; a condition on attributing any beliefs, odd or otherwise, to a speaker:

When she (the interpreter) suddenly finds herself crediting the L-speaker with an inexplicable belief, she must have a specific semantic location in the structure of L for the sentence, in so far as she believes she knows what it means, but has no similar place for the odd belief in her pattern of beliefs... .. The oddness of the belief presupposes that the interpreter has assigned meanings to the words of the sentences.¹⁰⁰

DPOC is by Davidson's lights a necessary condition on constructing a theory of meaning for a speaker, and theories of a speaker's meanings and beliefs come together. So arguments for alternative principles that assume we already have theories

⁹⁹ Grandy's statement of a principle of translation may be erroneous in another way. A translator does not state any semantical truths. The translator's manual gives us a mapping of the sentences of one language onto sentences of another language with the same meaning. But it does not say what it is for any of those sentences to be true. So a translation manual does not relate beliefs and desires to the world.

¹⁰⁰ Ramberg, 1989, pp.78.

of meaning and belief in place will not cut to the heart of the claim that DPOC is a necessity of interpretation. Grandy has a rather different role in mind for his POH. The POH is employed on a step-by-step basis, first yielding a translation of the speaker's language into our own, then determining the speaker's beliefs and desires and then predicting actions:

The actual use of translation in this prediction process is only one of the intermediate steps. We can translate verbal behaviour into our own language and use this to determine what the person's beliefs and desires are, and then use that information to predict actions.

For Davidson, none of these activities can be carried out in isolation: one cannot give a theory of meaning, belief or action independently of one another. In Davidson's method DPOC was designed to solve the problem of the entanglement of meaning and belief. If one could translate a speaker prior to working out what their beliefs and desires were we would not need to employ DPOC. Grandy is working on the hypothesis that one can compile a translation manual prior to determining what the speaker's beliefs are. Davidson thought the construal of a speaker's words impossible prior to theorizing about a speaker's beliefs.

Grandy thought that in obtaining a theory of belief and desire, together with facts about non-verbal behaviour, one would not have sufficient evidence to give a determinate prediction of a subject's behaviour. Grandy is concerned that a subject's actions should be predictable and that the inputs he highlights are insufficient for such prediction. He takes it that, given this insufficiency, we need some further model of the subject and he moves for a POH.

In theory we could (perhaps) elicit the total belief-and-desire structure and use mathematical decision theory to arrive at the prediction, but this is not what we do in practice. And since it is the actual processes of communication that was our original concern, we must look for an alternative model. The most obvious alternative is that we use ourselves in order to arrive at the prediction: we consider what we should do if we had the relevant beliefs and desires.¹⁰¹

¹⁰¹ Grandy, 1973, p.442.

Grandy assumes that the mechanism for prediction that we use in conjunction with a theory of belief-and-desire must be readily available in communication: it must be one that we actually do or could make use of rather than one that is overly technical or sophisticated. He does not consider that decision theory might be a systematisation of what we actually do or a theory of which we have an implicit knowledge. For Grandy the natural procedure is to use ourselves as models for predicting the speaker's behaviour and this is enshrined in the POH.

But looked at in this way, POH is not a method for disentangling questions of meaning and belief, and construing a subject's words. Rather, Grandy's main concern is with a method for the prediction and explanation of behaviour, verbal or otherwise. He does not concern himself with the indeterminacy that arises from the entanglement of questions of meaning and belief (although a POH could perhaps be employed in holding belief fixed). Moreover, there are less ambitious theories than the one Grandy has in mind, theories that do not go as far as offering predictions of behaviour, which could claim to provide an interpretation of a speaker.

There are some puzzling aspects of the methodology that Grandy describes. He claims for example that:

Whether our simulation (*of a subject*) is successful will depend heavily on the similarity of his belief-desire network to our own.¹⁰²

But he has not provided us with any independent means to assess whether a belief-desire network is like our own. His instruction was simply: impute a pattern of beliefs and desires to the speaker which is "as similar to our own as possible". If the translator then witnesses behaviour that seems to disconfirm his attribution of a belief-desire network to a speaker it is unclear by what principle he should proceed. By Grandy's lights we cannot tell, independently of this simulation, whether the

¹⁰² Grandy, 1973, p.443.

speaker is similar to us and, hence, whether our simulation has been successful. What would help is a means to grade the similarity of belief-desire networks to our own. Then if the translator made an inaccurate prediction of behaviour based on attributing the speaker the pattern of beliefs and desires most similar to the translator's own, he could move to attribute the most similar pattern consistent with the new evidence. One should bear in mind that some of the plausibility of Grandy's POH resides in the fact that when one imagines how the project of translation will actually be carried out, one tends to imagine even unfamiliar speakers as being a lot like us. Thus, we find it plausible that imputing patterns of belief and desire similar to our own will yield acceptable results. One should bear in mind that however well POH may turn out to serve as a rule-of-thumb, finding out about the similarity or otherwise of the speakers with which one is confronted ought to be a goal of interpretation.

The POC that Grandy found in Quine's account of Radical Translation and wished to criticize was the following:

choose that translation which maximized agreement (at least of certain sorts) between ourselves and our translatee.¹⁰³

The agreement "of certain sorts" is agreement on obvious truths such as those of logic and of observation. It was worth stressing the differences between DPOC and Quine's POC. It is plain from the quotation that Grandy sees Quine's POC as a rule by which to choose between alternative manuals of translation. We should remember that, whether rightly or wrongly, Davidson saw DPOC as a necessary condition on obtaining a theory of meaning for a speaker, not as a means by which to choose amongst several possible rivals. The role of DPOC consisted in fixing belief by finding responsiveness to the environment and consistency in a speaker so as to solve for meaning. DPOC presents a moving target to Grandy's arguments: Davidson has moved away from the old Quinian dictum involving maximizing agreement on obvious truths.

¹⁰³ Grandy, 1973, pp.440.

Grandy considers a complicated theorem of quantification theory that can be proved but could, prior to one's seeing the proof, appear invalid. Grandy imagines showing the theorem to someone and asking them if it is valid. It is quite possible, perhaps likely, that they will give the wrong answer. But Quine's POC instructed us to maximize agreement on logical truths and translate the speaker so that he takes the theorem to be valid. It strikes me that in making the following sorts of injunction:

the past history of a speaker is quite relevant to the question of what is obvious to him.....
We should not go out of our way to find some complicated agreement on this logical truth, because the error is not only explainable but was predictable given some knowledge of his past history.¹⁰⁴

Grandy is accommodating a point that Davidson too is concerned to accommodate. Grandy suggests that if we take the subject's past history into account it is both explicable and predictable that they should give the wrong answer. They have never seen the proof or such complicated theorems before. But DPOC does not counsel us to find agreement on complex logical truths at any expense to the simplicity and intelligibility of the interpretation. DPOC aims to optimise the subject's intelligibility and consistency not simply his agreeableness.

Grandy's master argument is that he has recognized cases in which Quine's POC leads to unnatural translations which attribute a speaker true but inexplicable beliefs. Grandy's points, whilst solid against the maximizing version, do not carry over to DPOC. On DPOC there are cases even at the outset in which the interpretation of a sentence involves the attribution of a false belief. Grandy recognizes that even Quine's POC can accommodate the attribution of false beliefs to a speaker. But the cases Grandy has in mind are widespread and can be generated systematically, so Quine's POC cannot accommodate them. For Grandy these cases are best accounted

¹⁰⁴ Grandy, 1973, pp.443-4.

for using a POH that can take account of causal evidence afforded only a peripheral role by Quine's POC.

The case Grandy focuses on is the following. Paul arrives at a party. Paul asserts "The man with a martini is a philosopher". There is a man in plain view who is drinking water from a martini glass and who is not a philosopher. There is in fact only one man drinking martini at the party. This man is a philosopher and he is standing out of sight in the garden. The problem is over what we should take the term "the man" as referring to in Paul's assertion. Grandy notes that according to Quine's POC we should take the remark at face value and count Paul's utterance as true, which means that "the man" refers to the man in the garden. But the natural thing to do is to take the utterance as false because we cannot explain how Paul could have the true belief about the man outside whom he has not seen. The falsity of the utterance is:

predicted by the principle of humanity, of course, for that constraint instructs us to prefer the interpretation that makes the utterance explainable.¹⁰⁵

On Davidson's view, attributing Paul a true belief about the man in the garden may not be feasible by the POCor. The interpreter would not in those circumstances be responding to the man in the garden. Furthermore, attributing^{to} Paul the belief that the man in the garden is a philosopher and is drinking a martini would require attributing^{to} him some surrounding beliefs for which one would have to be able to find evidence. It is puzzling that people take Grandy's points to be a straightforward counter to DPOC. It is not peculiar to the POH to attribute a subject "explicable falsehood" rather than "mysterious truth". DPOC mutated from an instruction to maximize agreement into a principle in which agreement in belief plays a part in securing shared subject matter, within the wider scope of explaining the limited perspective of a speaker. The error on the part of the speaker is explicable given the speaker's perspective; given the features of the world he could be responding to and his other beliefs. X

¹⁰⁵ Grandy, 1973, pp.445.

Grandy draws the moral that the POH fits best with a causal theory of belief. As Grandy finds a causal theory of belief integral to our understanding others, he thinks this speaks strongly in favour of a POH which:

directs us to bear in mind that the speaker is a person and has certain basic similarities to ourselves when we are choosing between translations. When we look at the actual choices we make, guided by this principle, it appears that part of our epistemological view of ourselves is what could be described loosely as a causal theory of belief. In the example just given, the operative principle that makes the case forceful is that the object that would be referred to according to the charitable interpretation did not interact causally with the speaker. Thus the causal theory of belief accords much better with the principle of humanity than with the principle of charity.¹⁰⁶

Obviously, the notion of someone's sharing "certain basic similarities" with us requires elaboration. This may amount to something like a POCor which instructs the interpreter to take the speaker to be responding to the same features of the world as he would in the circumstances. If Paul's patterns of belief are similar to ours then, then without any causal interaction with the philosopher in the garden he could not have a belief about which drink that person had. It is plausible that we do reason in such a way but it would be wrong to think that DPOC suggests something to the contrary. Grandy also insisted on placing heavy emphasis on the speaker's past history, his verbal conditioning and his non-verbal stimulations. Davidson might not have thought so much information available in the radical case.

4.3 McGinn on DPOC and the Causal Theory of Belief

McGinn has criticized Davidson views on the role of DPOC on the following grounds. McGinn thinks that Davidson's reason for adopting DPOC is that without assuming that most of a subject's beliefs are true, we would not be able to work out what beliefs the subject had. McGinn claims to the contrary that observation of

¹⁰⁶ *ibid*

speakers causally interacting with objects in their environment gives independent grounds for ascribing them beliefs. McGinn claims that these independent grounds leave DPOC redundant.

McGinn is mistaken about the role of DPOC on three counts. Firstly, there is no general assumption within DPOC to the effect that most of a subject's beliefs are true – this is an artifact of interpretations that employ DPOC. This aspect of DPOC was most clearly brought out in writings that post-date McGinn's paper. It is clear that McGinn has misinterpreted Davidson in this way:

charity as a methodological precept is to be insisted on because we know in advance, by a transcendental argument of some sort, that most of what others say and believe is going to be true (according of course to our own view of the truth).¹⁰⁷

Secondly, strictly speaking Davidson does not think it makes any sense to talk of “most” of a speaker's beliefs, beliefs being infinite in number and there being no useful way to count them. Thirdly, we apply DPOC in ascribing beliefs to a speaker but that is not the whole story of belief ascription. One employs DPOC in holding belief fixed so as to discern a speaker's meanings, and we work out a speaker's beliefs by knowing what their words mean. There is no way, on Davidson's method, of working out what someone believes independently of interpreting their words and we attribute someone beliefs in tandem with interpreting their words. The Charitable belief ascriptions are revisable as the interpreter learns more about the speaker's language.

There is common ground between ~~the argument of this text~~ and the point McGinn is pressing. I agree that there are grounds independent of DPOC by which an interpreter may begin to discern what it is that someone believes. But McGinn thinks that merely observing the causal interaction of subject and environment can ground the ascription of fine-grained beliefs, whereas nothing that straightforward is defended

¹⁰⁷ McGinn, 1977, pp.522.

here. It is unclear exactly what falls under the umbrella of a speaker's causal interaction with his environment for McGinn. If it means merely observing, as one does habitually, the objects and events with which some speaker is confronted then McGinn's claim seems to me wrong. The speaker's senses are bombarded by information from the external world and the causes of this information are many. Habitual observation alone cannot serve to justify the ascription of finely discriminated propositional attitudes: we cannot be sure, on this basis, to what aspects of this information a speaker is responding. Davidson cuts down on this myriad information by assuming that the speaker responds to that aspect of the environment which the interpreter would respond to under those circumstances. This thesis suggests that some of this work may be done by the psychological sciences.

McGinn notices that, on Davidson's account, there are heavy burdens on attributions of false beliefs in the early stages of interpretation. McGinn takes this as a point against Davidson. McGinn thinks that Davidson's desired conclusion is that most of what others say and do must be true but this is incorrect. Davidson's point is rather that those interpretable using DPOC will exhibit much truth and consistency. McGinn thinks of DPOC as the assumption that most of a speaker's beliefs must be true, and seeks to find in Davidson's writings some argument to that effect that comes prior to and stands apart from the activity of interpretation.

McGinn notes the problem of entanglement but misses the importance of DPOC in that context. For Davidson it is only by holding belief fixed that we can effect a systematic pairing of utterances and their truth-conditions. Davidson's real argument for DPOC is not "transcendental", it is that there is no feasible way to accomplish the disentanglement other than by applying DPOC. McGinn thinks that in disentangling meaning and belief one does as well by taking a speaker to have false belief as a true one:

For we may equally provide a basis for deriving the meanings of sentences held true by uncharitably imputing false beliefs to our speaker¹⁰⁸

But for Davidson establishing some shared content between speaker and interpreter requires establishing some shared belief. We have also seen that the mere assumption of falsity cannot pin to a belief a particular content or range of contents. If we assume that our speaker has made a mistake and is expressing a false belief then which one is it? A high instance of falsity in his surrounding web of beliefs may leave us beggared for a subject matter shared with the speaker. By DPOC the speaker is responding to the same feature of the environment as the interpreter would in the circumstances. A merely false belief need not be a response to the observable circumstances at all. Knowing only that a belief is false and that all the beliefs implied by it are false does not serve to fix upon a content so as to facilitate the pairing of an utterance with its truth-conditions. What stands in the way of McGinn's claim is Radical Interpretation.

McGinn argues that we can make sense of a speaker's belief being about a certain object or event without taking that someone to have a collateral constellation of other true beliefs concerning that object or event. McGinn argues that what allows us to identify someone's belief set as a collection of "egregious misconceptions" is that attributions of false belief have a relational form:

in assigning such attitudes to a person, there is no presumption that the concepts (predicates) the reporter calls upon to pick out the entity the reported belief is said to concern, be themselves – those concepts – credited to the believer. The reporter employs *his* concepts to identify some object he takes the believer to be cognitively related to. That these concepts are not assumed to be possessed by the believer is indicated by keeping the vocabulary one uses to express the concepts outside the scope of the belief operator.... We are thus able intelligibly to impute preponderantly false beliefs to the ancients in relation to the earth and the stars without the implication that they believed these things of those entities as conceptualized by us.¹⁰⁹

¹⁰⁸ McGinn, 1977, pp.523.

¹⁰⁹ McGinn, 1977, pp.525

McGinn's idea is that the shared subject matter does not require surrounding true belief but is fixed causally. False believers are causally related to the same objects as us, but it is not required that they conceptualise them in the same way.¹¹⁰ A worry for this view of McGinn's is that it identifies the cause of a belief with the content of that belief. There are many cases in which the cause of a belief is not identifiable with its content. Imagine that some primitive people attributed mystical importance to squid and held some strange beliefs about what this animal is and the powers it has. It may be true that the appearance and activities of squid have caused many of their strange beliefs about a marine god but untrue that the subject matter of their belief can be characterized by the mundane marine activities. What allows us to hold the content and cause of a belief distinct is that the content of a particular belief is dependent not only on its causal genesis but on its place within a cluster of beliefs. If McGinn were right then those who held beliefs about phlogiston were really thinking about oxygen, because the substance oxygen was the real cause of their belief. Applications of POCor and POCoh do not have the implication that the content of a belief must be its cause. Ascribing intelligible beliefs about particular objects becomes impossible if those beliefs exhibit too much inconsistency and falsity. What gets in the way of McGinn's argument is Radical Interpretation. The holistic nature of Radical Interpretation seems to ensure that agreement in subject matter will involve some agreement in opinion.

The POCor ensures a causal aspect to the determining of the content of beliefs. McGinn takes Davidson not to affirm something that it seems to me he does affirm, namely;

it is a necessary (but not sufficient) condition of an object's being referred to in correctly specifying the intentional content of a relational mental state that it – that object – figure suitably in the causal genesis of that state.¹¹¹

¹¹⁰ Although it does seem that whatever the interpreter places inside the belief operator must be described using some set of the interpreter's own concepts.

¹¹¹ McGinn, 1977, pp.527.

Compare, for example, these statements concerning the conditions Davidson considers necessary for content individuation:

propositional attitudes and related events and states are in part identified in terms of their causal and other relations to events extraneous in time and place to the agent they characterize.¹¹²

Each of two people is reacting differentially to sensory stimuli streaming in from a certain direction. Projecting the incoming lines outward, the common cause is at their intersection.¹¹³

McGinn mistakenly takes Davidson to be denying a claim that McGinn finds plausible: that the intentional objects of belief are in part determined by the speaker's causal history. From the claim that Davidson denies this point, McGinn wants to motivate an alternative approach to interpretation on which we take this "independently plausible" thesis about the relational form of beliefs as a basis from which to begin interpretation.

It is because we observe that people causally interact with objects in their environment in such ways as enable them to have thoughts concerning those objects, paradigmatically in perception, that we are prepared to assign those objects to their beliefs as comprising their subject matter, notwithstanding the amount of bad theory they may bring to bear upon their objects.¹¹⁴

Davidson agrees that observing people interacting with their environment is the major source of evidence for ascribing them thoughts. The point of disagreement is over the autonomy of content and theory. McGinn thinks that the content of the speaker's beliefs is ascertainable independently of the speaker's theory and Davidson holism about belief ascription implies the contrary. Davidson recognized that merely observing this causal interaction with his environment will not serve to discriminate the contents of a speaker's propositional attitudes. Once we apply the POCor to this

¹¹² 1997b, pp.71.

¹¹³ 1991, pp.213.

¹¹⁴ McGinn, 1977, pp.530.

evidence we are ascribing not only a belief about the observable circumstances but also a cluster of beliefs needed to make sense of the speaker having the belief ascribed to him. Davidson was convinced that mere observation of causal chains underdetermines the content a mental state possesses:

the cause is doubly indeterminate with respect to width, and with respect to distance. The first ambiguity concerns how much of the total cause of a belief is relevant to content... The second problem has to do with the ambiguity of the relevant stimulus, whether it is proximal (at the skin, say) or distal.¹¹⁵

There can also be a range of causes that count as instantiations of one content. Davidson saw social interaction as resolving such issues and determining content:

I would introduce the social factor in a way that connects it directly with perceptual externalism, thus locating the role of society within the causal nexus that includes the interplay between persons and the rest of nature.¹¹⁶

On Davidson's picture social factors have a normative as well as a causal role in content individuation. Whether or not one agrees that social factors can be the factors that make the determining difference, we can take these remarks as showing that Davidson had thought through the role that a subject's causal interaction with his environment plays in fixing the content of his beliefs and had not ignored the kind of alternative that McGinn presents.

For McGinn, like Davidson, the attributions of thoughts are tested "by seeing whether the agent behaves as one would who possessed those attitudes" and the verification is holistic. McGinn suggests that we think of a theory that would explain the observed facts if it were true, then test it by working out its further consequences and seeing if they are realized in the speaker's behaviour. On this point, McGinn is in agreement with Davidson. In addition McGinn thinks that the theorist's conjectures should

¹¹⁵ 1997a, pp.129-30.

¹¹⁶ 1990, pp.201.

accord with whatever hunches and expectations one has about a subject. One might find this approach more compelling if the “hunches and expectations” were replaced by knowledge or models from the psychological sciences. McGinn notes that his method is not a method in the sense of a set of instructions that select the correct theory without imagination on the theorist’s part. He is dubious that what we require is a method in this sense:

the lack of such a procedure for extracting theory from evidence does not impede science; and indeed, it is to be assumed that radical interpretation has often been successfully undertaken without benefit of Davidson’s method. It may well be, as Davidson’s writings suggest, that there can be no method without charity: but there can be interpretation without either.¹¹⁷

McGinn’s comments on the lack of a uniform procedure for deducing theory from evidence are congenial to the more spontaneous and patchwork approach to understanding a speaker that ^{this} thesis proposes as an alternative to DPOC.

4.4 Quine and Davidson: A Behaviouristic Legacy

In the introduction to *Inquiries into Truth and Interpretation* Davidson says that:

Only by studying the *pattern* of assents to sentences can we decide what is meant and what believed. Depending on this evidence which, without the aid of theory, makes no distinction between the contribution of belief and meaning to linguistic behaviour, requires a method for effecting the separation to a degree sufficient for communication. Devices to this end are described and defended in the present essays. But all of them, in one way or another, rely on the Principle of Charity.¹¹⁸

Davidson believed that it is only by finding patterns in someone’s verbal behaviour that we can go about understanding them. As this evidence does not serve to distinguish the contribution of belief and meaning to the observed behaviour,

¹¹⁷ McGinn, 1977, p.535.

¹¹⁸ 2001b, pp.xix

Davidson requires a device to effect this discrimination and it is here that DPOC comes in.

Davidson acknowledged Quine's discussion of Radical Translation as the forefather his own account of Radical Interpretation:

The idea of a translation manual with appropriate empirical constraints as a device for studying problems in the philosophy of language is, of course Quine's. This idea inspired much of my thinking on the present subject, and my proposal is in important respects very close to Quine's.¹¹⁹

Although Davidson was opposed to Behaviourism as a doctrine about the reducibility of intentional to behavioural facts, his views on the evidence relevant to an interpreter derive directly from Quine's Behaviourism about language. Quine claimed that all the evidence relevant to correct translation consists in facts about speakers' overt dispositions to behaviour:

Language is a social art. In acquiring it we have to depend entirely on intersubjectively available cues as to what to say and when. Hence, there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable stimulations.¹²⁰

In compressed form this passage contains what will be discussed in a moment as "Quine's acquisition argument". The major premise of this argument is that language acquisition depends entirely on observation of behaviour and reinforcement of learned behaviour. One way of taking this premise is as a highly contentious empirical claim about language acquisition. The second suppressed premise of the argument is that translators can draw upon no more evidence than is available in first language acquisition. According to Quine, the evidence available to child and translator cannot differ in type. The conclusion of Quine's argument is that the only

¹¹⁹ 1973a, pp.129n

¹²⁰ Quine,1960, pp.ix

available evidence to the translator is observable behaviour. Quine recognized that he had provided his Radical Translator with a “meagre basis”¹²¹ but maintained that “the native speaker has had no other”.¹²²

Davidson seems to have accepted Quine’s characterisation of the evidence but not Quine’s claim that if there were facts about meaning they could amount to nothing more than facts about behaviour. Quine had two arguments for his Behaviorist view of the evidence; one concerned language acquisition and the other concerned what is manifest in linguistic understanding. By undermining Quine’s arguments for this construal of the available evidence, from which Davidson’s own view derives, we take a substantial step towards the conclusion that Davidson is too narrow-minded about the kind of evidence that might be relevant to his interpreter. It then becomes an interesting question whether non-behavioural evidence is neutral as to the contribution of meaning and belief. It is revealing to view Quine’s scientific-sounding account of language through Chomsky’s eyes.

Given their prominence in empirical science and the antithesis they provide to Quine’s views, it is also worth having Chomsky’s positive views on the nature of language in the background. Chomsky depicts language as a mental faculty, more specifically; a physically realised set of psychological facts, concerning the principles by which the mind attributes phonetic forms and determinate meanings to physical objects. Chomsky’s depiction dominates empirical linguistics and provides evidence for the view that knowledge of meaning is represented in some way in the mind or brain. Chomsky’s suggests that under many circumstances features of the language faculty change according to pragmatics rather than any change in belief. These cases constitute a change in language that can be sharply distinguished from epistemological considerations.¹²³ So Chomsky’s work is in conflict with the view that we could not have knowledge about meanings, knowledge of a speaker’s language, without working out what the speaker believed.

¹²¹ Quine, 1987, pp.5.

¹²² *ibid*

¹²³ These suggestions are owed to Horwich, 1992.

On Quine's account of Radical Translation, the translation of observation sentences - those sentences for which patterns of assent and dissent are closely conditioned to stimulations - serves as an entering wedge into the language. Having translated some observation sentences, the translator can begin compounding those sentences. He also undertakes translations of a different kind. When there are no obvious correlations between a particular sentence being uttered and the concurrent stimulations, the linguist must note and dismantle these unconstrued sentences. Some fragments of these unconstrued sentences will occur in observation sentences, others will be treated as words and tried out in analytical pairings with expressions of the translator's language. The translator tries translations that pair unconstrued non-observation sentences with sentences of his own with similar ostensible features. Thus, the translator accumulates a tentative vocabulary and grammar subject to constant tests. Given these details and the possibility of mounting behavioural data, one should allow that Quine is describing something richer than one might expect possible from a foundation of pairings of stimulus synonymous sentences.

Quine adhered to Physicalism: the view that ultimately the physical facts constitute all the facts there are. One might have expected that by Physicalist lights there would be more physical facts available to the translator than merely facts about verbal behaviour. But it is clear that Quine's views on the facts relevant to translation are more austere:

There is nothing to linguistic meaning beyond what is gleaned from overt behaviour in observable circumstances. In order to exhibit these limitations, I propounded the thought experiment of radical translation.¹²⁴

It would be fair to say that Quine's arguments for this austerity have not found many supporters.

¹²⁴ Quine, 1987, pp.5.

The crucial point for Quine is that all the determinants of meaning must be objectively available. What objectively available meant to Quine in the context of translation was available from the third person perspective to someone who did not already know the language and who knew nothing further about the speaker. It is the inclusion of this last desideratum which perhaps stands out most. The translator is denied any knowledge about the speaker bar knowledge of his behaviour and whatever knowledge of a speaker's language can be constructed from this material. Both Quine and Davidson accepted the further point that the behavioural evidence would not distinguish the contributions of meaning and belief to a speaker's assertions. Quine's behaviouristic view of the evidence, and Davidson's concurrence, meant neither would consider seriously the possibility of other sources of evidence with positive bearing on the contributions of meaning and belief. If there is evidence that may be brought to bear on interpretation that supplements the behavioural evidence, it may serve to constrain the meanings or beliefs that the interpreter can ascribe. We would then be able to begin solving our problem of two unknowns, the problem of entanglement, without recourse to DPOC.

When drawing the evidence for the study of language Quine thought of language as *nothing but* a social phenomenon. On Quine's view, what it is that a speaker means could not be illuminated by facts about a speaker's sensory apparatus, a speaker's mind, or a speaker's brain, or by any facts made available from empirical psychology and linguistics. There are two separate but related arguments extractable from Quine for the view that the only evidence relevant to correct translation comes from overt verbal behaviour. Here they are termed the acquisition argument and the manifestation argument. The term "manifestation" is suggestive of a point of contact to be found between Quine's argument and an argument to be found in Dummett and Wright.¹²⁵ The conviction they share is that there is nothing more to meaning than what is manifest in linguistic behaviour.

¹²⁵ Wright, 1993, pp.13-26.

Quine's acquisition argument for Behaviourism starts with a claim about how we learn a language. Quine's major premise is that we learn a language solely through observation of overt behaviour and our responsiveness to correction:

Each of us learns his language by observing other people's verbal behaviour and having his own faltering verbal behaviour observed and reinforced or corrected by others.¹²⁶

Quine's argument runs as follows:

- (1) Language is learnt by observation of verbal behaviour.
- (2) No more is available to the translator than is available in learning a language. (suppressed premise)
- (3) Therefore, no more evidence is available to the translator than observable verbal behaviour.

Both of the premises are highly dubious and certainly requiring of convincing argumentation, for which one would search Quine's writings in vain. One could view Quine's premise (1) as an empirical claim about language acquisition; it is hard to see how else one could view it. Chomsky saw it as having the status of an empirical assumption given that no evidence is presented in support of it.¹²⁷ If seen this way we have to question whether it has been fatally undermined by Chomsky's own work in linguistics; empirical research which seems to disconfirm the hypothesis that children could come to speak the languages they do on the basis of the limited behavioural data they receive. To be valid the argument requires the suppressed premise (2) and this premise is false. Even if one accepted premise (1) it would not follow that the translator could not find other information useful. We want to allow that there can be better and worse translators and in the radical case the best translators should make use of whatever evidence is useful to understanding speakers.

What one has acquired when one knows a language is, on Quine's view, a set of dispositions to verbal behaviour. Quine characterizes dispositions to verbal behaviour in terms assent and dissent to a network of sentences related to one another and, in

¹²⁶ Quine, 1987, pp.5.

¹²⁷ Chomsky, 1975, pp.53.

part, to stimuli. Chomsky was highly skeptical, for reasons that will become apparent, of Quine's account of how we learn a language. One can be impressed by Chomsky's points against Quine even if one does not wholeheartedly endorse Chomsky's own account of language acquisition. His early arguments against Behaviourism placed Chomsky himself in such a position:

it seems that there is neither empirical evidence nor any known argument to support any specific claim about the relative importance of "feedback" from the environment and the "independent contribution of the organism" in the process of language acquisition.¹²⁸

Quine famously claimed that in imputing meanings and beliefs to a speaker we go beyond what is implied by that speaker's verbal behaviour:

we impute our sense of linguistic analogy unverifiably to the native mind.¹²⁹

Quine thought the analogy unverifiable because there are alternative imputations inconsistent with one another but consistent with all the evidence. Chomsky took issue with Quine's indeterminacy thesis because it depended on his Behaviourist theory of acquisition. Quine and Davidson held that an element of indeterminacy in translation arises from the interplay of meaning and belief ascriptions and sought to alleviate this indeterminacy by applying a POC. According to both Quine and Davidson the indeterminacy survives application of a POC. Chomsky argued that there are further determinants to meaning than those identified by Quine, and that the indeterminacy can be resolved on an improved understanding of the factors involved in language learning. He claimed that there is one true and determinate meaning for each sentence of the language settled by the baggage human organisms bring to language use. It is not my intention to defend Chomsky's dismissal of the indeterminacy of translation but rather to sketch his alternative picture. On Chomsky's view our minds develop systems of knowledge that determine the status of arbitrary physical objects, assigning to these objects syntax, phonetics and

¹²⁸ Chomsky, 1959.

¹²⁹ Quine, 1960, pp.72.

meaning. On this picture a language is a specific function assigning meaning to a vast range of physical events, and this function is ultimately realized in physical mechanisms.¹³⁰ My focus is not on these positive claims, but on how arguments against Behaviourist theories of acquisition (and competence) undermine Quine's Behaviouristic delimitation of the evidence available in translation.

Chomsky was unhappy with the notion of similarity required by Quine's account. Speakers are required to find certain sentences and certain stimuli similar and we need an explanation of what this consists in. A traditional Behaviourist explanation had been that two stimuli are similar when the speaker makes the same sort of response to them. But we need to know when two responses are of the same sort. This explanation of similarity is in danger of turning into a tautology.¹³¹ If one wanted to define similarity of stimuli in terms of similarity of responses, one could not define similarity of response, in terms that involve stimuli as closing the loop would be unappealing. Quine had a different explanation of similarity. He postulated that we possess pre-linguistic, quality spaces. One needs to know more about these quality spaces and their nature for Quine's account to explain anything about how language is learnt. What is a quality space? Chomsky offered the following explanation on Quine's behalf.¹³² A quality space may be thought of as an in-built measure of distance, perhaps "restricted to dimensions with physical correlates such as brightness or loudness, and distance defined in terms of those physical correlates".¹³³ Chomsky's point is that we now have an empirical "doctrine of innate spaces" which must face the empirical evidence. In reply,¹³⁴ Quine claimed that there must be such things as innate quality spaces as there can be no habit formation without similarity dispositions to find things similar. Quine claims that it is the nature of these quality spaces that must be experimentally determined, not the issue of whether or not there are such things. But surely Quine did not mean to suggest that it is not an empirical

¹³⁰ Chomsky labels this conception of language "I-language" and claims that it is I-languages that are the proper objects of empirical study.

¹³¹ Chomsky, 1959.

¹³² Chomsky, 1975.

¹³³ *ibid*

¹³⁴ Quine, 1969b.

issue whether such quality spaces exist, and if they do exist, what role they play. Quine did not meet Chomsky's objection: his claim about "quality spaces" and the role they play in acquisition has not been made to face the evidence or compete with alternative hypotheses. Neither did Quine do anything to remedy Chomsky's concern that the nature of a "quality space" remains mysterious.

Chomsky is unhappy for similar reasons with the notion of analogical synthesis important to Quine's account of acquisition. Analogical synthesis is the process of abstraction and assembly of parts of sentences. Quine uses the notion to explain how we understand sentences not directly tied to stimuli, but:

this explanation is empty until an account of "analogy" is given, and none exists.¹³⁵

The claim that language learners make such synthesis is also mysterious until a basis for this abstraction and assembly is specified and empirically tested. Quine's refusal to look to the contribution of the child leaves us with a superficial account of learning on which we attribute a vast and uncomprehended contribution to a step called synthesis or generalization. This step includes a great deal of what is of interest in the study of language learning. Chomsky takes Quine's view to highlight the danger of neglecting the part that the structure of an organism, whether described in biological terms or at the level of abstraction Chomsky terms the "mind-brain", may play in the acquisition of language.

Chomsky saw it as an empirical problem to assess what properties of the mind or brain determine the nature of our experience, how experience contributes to language learning and what we come to believe on the basis of that experience.¹³⁶ He asks us to adopt an attitude of distrust towards the idea that properties of the mind open to empirical investigation "impose no conditions on language and theories".¹³⁷ Quine's

¹³⁵ Chomsky, 1990, pp.589.

¹³⁶ Chomsky, 1975, pp.64.

¹³⁷ Chomsky, 1975, pp.65.

position appeared to him nothing more than a dogma unsupported by evidence about language learning. Chomsky's positive views are in sharp contrast with Quine's:

it is clear that when we learn a language we are not "learning sentences" or acquiring a "behavioral repertoire" through training. Rather, we somehow develop certain principles (unconscious of course) that determine the form and meaning of indefinitely many sentences. A description of knowledge of language as an associative net constructed by conditioned response is in sharp conflict with whatever evidence we have about these matters.¹³⁸

What is the evidence which Chomsky finds so inexplicable for the Behaviorist?

We must deal with the crucial and easily demonstrated fact that what a person knows (*in knowing a language*) is vastly underdetermined by available evidence, and that much of this knowledge is based on no direct evidence at all. There is little doubt that this problem of "poverty of stimulus" is in fact the norm rather than the exception.¹³⁹

From an early date Chomsky stated in modest terms the importance that the work in linguistics with which he was involved could have for understanding why one possesses the language and theory one does:

Although such a study, even if successful, would by no means answer the major problems involved in the investigation of meaning and the causation of behaviour, it surely will not be unrelated to these. the general character of the results of syntactic investigations may be a corrective to oversimplified approaches ^{to} the theory of meaning.¹⁴⁰

This is not a defence of a Chomskian conception of language acquisition, or aspects of his views to do with innateness, universality or internality. Rather it is an attempt to make palpable the highly speculative and doubtful nature of Quine's account. Quine's reply to Chomsky illustrates well the level of speculation in which he was involved:

¹³⁸ Chomsky, 1975, pp.64.

¹³⁹ Chomsky, 1990, pp.593.

¹⁴⁰ Chomsky, 1959.

Language aptitude is innate; language learning, on the other hand, in which that aptitude is put to work turns on intersubjectively observable features of human behaviour and its environing circumstances, there being no innate language and no telepathy.¹⁴¹

To Chomsky this would have appeared to be merely a restatement of the unjustified empirical assumption to which he had accused Quine of adhering. Whether learning turns only on observation of behaviour is a matter to be investigated; Chomsky: a preponderance of empirical linguists strongly favour a negative answer and stress that in the case of language learning:

prediction of the behaviour of a complex organism would require, in addition to information about external stimulation, knowledge of the internal structure of the organism, the ways in which it processes input information and organizes its own behaviour.¹⁴²

Behaviourists who tried to explain language acquisition as the activation of pre-existing aptitudes (which themselves would require much explanation), by stimulus and reinforcement, omitted factors of first importance to the scientific study of language. They also faced a problem with the claim that regularities in dispositions to speech are formed through reinforcement:

It is not easy to find any basis (or, for that matter, to attach very much content) to the claim that reinforcing contingencies set up by the verbal community are the single factor responsible for maintaining the strength of verbal behaviour. The sources of the "strength" of this behaviour are almost a total mystery at present.¹⁴³

It is conceivable that there could be Behaviourists who do not say that reinforcement is responsible for consistency in verbal behaviour. However, Quine contends that observable behaviour determines the specific character of language acquisition because a child speaks the language of, and receives correction from, the group in which he grows up. The argument is superficial because at this level of speculation,

¹⁴¹ Quine, 1969, pp.306.

¹⁴² Chomsky, 1959, pp.27.

¹⁴³ Chomsky, 1959.

Chomsky points out, one may equally hypothesise that the human organism has evolved in such a way that on observing verbal behaviour he miraculously internalizes the rules of English grammar. Or that on observing the application of a term to certain instances he predicts the extension of that term to complexly related instances. Whatever theory of language learning one adopts it must be able to cope with the following sorts of data:

Study of the actual observed ability of a speaker to distinguish sentences from nonsentences, detect ambiguities, etc., apparently forces us to the conclusion that this grammar is of an extremely complex and abstract character, and that the young child has succeeded in carrying out what from the formal point of view, at least, seems to be a remarkable type of theory construction. Furthermore, this task is accomplished in an astonishingly short time, to a large extent independently of intelligence, and in a comparable way by all children. Any theory of learning must cope with these facts.¹⁴⁴

There is no explanation of these facts in terms of reinforcement or indication of how Behaviourist theory could explain these facts in Quine's writings.

Quine's manifestation argument, like his acquisition argument, aims to establish that the only evidence relevant to translation is behavioural. It starts from the premise that there is nothing more to understanding than what is manifest in our use of language. Quine has several takes on this argument. The first concerns the basis on which speakers are counted as understanding the language:

As long as our command of language fits all external checkpoints, where our utterances or our reaction to someone's utterance can be appraised in the light of some shared situation, so long all is well. Our mental life between checkpoints is indifferent to our rating as a master of the language.¹⁴⁵

In this last sentence, Quine has run two distinct points together. There may be a sense in which it is correct to say that our mental life is indifferent to our mastery of the

¹⁴⁴ Chomsky, 1959.

¹⁴⁵ Quine, 1987, pp.5.

language. Whatever is going on in our minds at the level of consciousness we will generally count as a master of the language if we perform correctly with the language.¹⁴⁶ One would not, for example, be counted an incompetent calculator if one calculated correct answers but had strange thoughts running through one's head concurrently.¹⁴⁷ We might put this by saying that generally verbal behaviour is a criterion of understanding language. This view of mastery is still contentious in one respect - think, for example, of a man who commands complete mastery of the language but has lost his expressive powers.¹⁴⁸ But Quine also implies something far less plausible which is that our mentality has no impact on, is not a determinant of, how we perform with the language.

When one claims that a speaker's mentality is a determinant of his mastery of the language what one is concerned with is not what Wittgenstein described as:

a wheel that can be turned though nothing else moves with it, is not part of the mechanism.¹⁴⁹

One is not concerned with the postulation of internal mechanisms that have no effect on verbal behaviour and our use of language. There may be properties of the mind – “internal mechanisms” if you like – which are crucial determinants of verbal behaviour and one's mastery of the language. Idle wheels are irrelevant to Non-Behaviourist and Behaviourist alike. This argument does not establish that there is nothing more to mastery, and hence meaning, than is manifest in verbal behaviour because it provides no good reason to believe that mentality is indifferent to mastery. On the contrary, one might expect our mindedness and our use of language to be intimately related, and the former to be a determinant of the latter. The argument that one should consider only behavioural evidence because one counts as a master of the language on the basis of behaviour, is not only invalid but rests on a doubtful premise about the basis on which one comes to count as a master of the language.

¹⁴⁶ Although there are interesting questions about whether zombies, creatures that behave as competent language users do but have no inner life, count as masters of the language.

¹⁴⁷ Wittgenstein makes this point several times in *Philosophical Investigations*.

¹⁴⁸ Chomsky, 1990.

¹⁴⁹ Wittgenstein, 2001, sect 271.

The second tack Quine takes on manifestation is to consider how we habitually understand speakers of our own language. One only comes to know what others mean, he claims, on the basis of their behaviour. So how, the argument runs, could the translator be needy of any other evidence?¹⁵⁰ We have seen reason already for thinking that the premise that knowledge of language is best explained in terms of the observation of behaviour is dubious. If Quine's claim about how one knows what others mean were true then aspects of the speaker that cannot be reflected in their behaviour are not relevant to what one learns by observing their behaviour. The problem with this argument is that its conclusion might be accepted even by someone who rejected Behaviourism. The Behaviourist's opponent may agree that aspects of the speaker that have no resonance in verbal behaviour are not relevant to linguistic understanding whilst maintaining that there is an awful lot that plays a determining role in language use which is not merely behaviour. In other words there may be a lot of non-behavioural factors reflected in our use of language. Aspects of the speaker that are irrelevant to the kind of language they speak are irrelevant to the study of language; no one denies this.

There may be another strand to Quine's manifestation argument. Suppose we can come to know what it is that someone means on the basis of unsophisticated evidence about their behaviour. Allow Quine, for the sake of argument, that this is an accurate portrayal of what is involved in understanding: we ascribe everyday, simple meanings to people on the basis of readily available, unsophisticated behavioural evidence. Quine might think that the possibility of bringing more sophisticated evidence to bear on what someone means or believes is precluded by the facts that we come to know what familiar people mean on the basis of unsophisticated evidence and that the meanings we ascribe to them are generally simple. Therefore, the argument would go, more sophisticated scientific knowledge could not be relevant. But this is a mistake. More sophisticated evidence could be relevant to getting at the very same simple

¹⁵⁰ On the I-language conception, the discovery of I-meanings may require sophisticated techniques. I-languages do not have the epistemological properties affiliated to pre-scientific conceptions. See Horwich, 1992.

meanings and beliefs, especially in more radical cases where the speaker is very unfamiliar. There may be expert, as opposed to poor, translators who command a greater wealth of information about the speaker and his environment, and who are more ingenious in translating speakers of varying degrees of familiarity.

There may be an even more serious obstacle for Quine's manifestation argument which can only be touched upon here. McDowell has argued that manifestation arguments cannot be successful because taking them seriously makes understanding impossible.¹⁵¹ This argument is worth mentioning because its premises are in keeping with some prominent ideas in Davidson's writings¹⁵² and yet Davidson accepts some version of the manifestation argument. McDowell's argument depends on the idea that understanding and meaning are normative notions. Normative notions draw on standards of correctness, and distinctions between being and merely seeming right. One's meaning or understanding something by an expression requires there being a distinction between correct and incorrect uses of that expression. The lack of such a distinction compromises the expression meaning anything at all because without it, any use of the expression would count as one's having understood it and used it in accord with its meaning. Quine may not have been impressed by this kind of talk. But Davidson agreed that thought and meaning are normative notions requiring a distinction between seeming and being right. A premise of Quine's manifestation argument is that understanding can amount to nothing more than what is manifest in our linguistic behaviour. The thesis must characterize understanding without appealing to a prior command of the language otherwise one could say that a speaker manifests their understanding of the language simply by speaking the language. So manifestation of understanding must consist in behaviour such as assenting to a sentence in observable circumstances – much as Quine describes it. But what is actually manifest in episodes of behaviour? Behaviour can confirm the attribution of dispositions,

¹⁵¹ McDowell, 1984, sect 10.

¹⁵² 2001c in particular.

but how can we extrapolate to a determinate conception of what the disposition is a disposition to do?¹⁵³

How can we extrapolate a determinate disposition to do something from the behaviour we observe? We could only extrapolate from the behaviour inductively: x has seen the sign-post three times and gone left three times, by induction he is disposed to go left at the sign-post. What is not manifest is commitment to a determinate pattern. If meaning must be fully manifest in behaviour then no determinate extrapolation can be fixed upon. The manifestation argument rules out the idea that in understanding language one grasps a determinate pattern of use that extends beyond finite instances of behaviour – the idea that there can be correct and incorrect uses of expressions. Without this distinction we cannot make sense of meaning and understanding.

Behaviourism engenders a lack of curiosity about the relations between human nature and both the meanings of the expressions of human language and the beliefs held by its speakers. This lack of curiosity is not justified by the fallacious arguments for Behaviourism to be found in Quine's writings. Davidson rejected Quine's notion of stimulus meaning and certainly saw himself as breaking from the Behaviourism he associated with that notion. But Davidson's delimitation of the evidence for interpretation closely follows Quine's; the evidence comes from speaker's behaviour.

The crucial point on which I am in agreement with Quine might be put: all the evidence for or against a theory of truth (interpretation, translation) comes in the form of facts about what events or situations in the world cause, or would cause, speakers to assent to, or dissent from, each sentence in the speaker's repertoire. We probably differ on some details. Quine describes the events or situations in terms of patterns of stimulation, while I prefer a description in terms more like those of the sentence being studied; Quine would give more weight to a grading of sentences in terms of observability than I would; and where he likes assent and dissent because they suggest a behaviouristic test, I despair of Behaviourism and accept frankly intensional attitudes towards sentences, such as holding true.¹⁵⁴

¹⁵³ McDowell, 1984, p.68.

¹⁵⁴ 1979, pp.230-1.

Davidson placed the evidence in much the way Quine had done; even though they differed over the terms in which they thought the evidence should be described. The point of agreement is apparent in remarks such as:

Only by studying the *pattern* of assents to sentences can we decide what is meant and what believed.¹⁵⁵

The ultimate evidence for the correctness of a truth theory we must presume lies in available facts about how speakers use the language. When I say available, I mean publicly available – available not only in principle, but in fact to anyone who is capable of understanding the speaker or speakers of the language.¹⁵⁶

Davidson's condition that the evidence must be available, in fact, to *anyone*, may make the Radical Interpreter's task a mundane one. How much evidence is really available to absolutely *anyone* who is capable of speaking a language? What carries over from Quine is an assumption of "observationality". This is the assumption that meanings and propositional attitudes must be discoverable by interpreters privy to nothing but observable verbal behaviour, even if one holds, as Davidson does, that the meanings and attitudes are not reducible to that behaviour:

I assume that an observer can under favourable circumstances tell what beliefs, desires, and intentions an agent has. the observationality assumption does not imply that it is possible to state explicitly what evidence is necessary or sufficient to determine the presence of a particular thought; there is no suggestion that thinking can somehow be reduced definitionally to something else. Nor does the observationality assumption amount to Behaviourism. Propositional attitudes can be discovered by an observer who witnesses nothing but behaviour without the attitudes being in any way reducible to behaviour.¹⁵⁷

One might try to defend Davidson's views on the evidence relevant to interpretation by claiming that they represent merely the stipulation of a radical case. One might say

¹⁵⁵ 2001b, pp.xviii.

¹⁵⁶ 1988, pp.176.

¹⁵⁷ 1982, pp.99-100.

that Davidson has described the radical case such that this is all that is available to the Radical Interpreter. But one should bear the following in mind. Davidson's intention was to describe the radical case so as to get clear on the objectively available evidence, putting subjective elements to one side. Even if beliefs and meanings are in everyday cases discoverable by observing verbal behaviour, the fact that they are so discoverable does not imply that no other objective evidence would have been efficacious in their discovery. One should question the wisdom of the way Davidson has stipulated the evidence to which we are exposed in the radical case. Rather than accepting that verbal behaviour in observable circumstances is all that is objectively available in the radical case because that is how the radical case has been described, one should be suspicious that Davidson has described the evidence available to the Radical Interpreter in a way that makes his project less interesting than it might be.

4.5 Davidson and Lewis: Lewis's Radical Interpreter, Physicalism and Determination

David Lewis conceived of the task of Radical Interpretation in a rather different way to Davidson. Lewis did not think that Radical Interpretation is a matter of how *an interpreter* could work out a speaker's beliefs and meanings as a real-life task. He saw this as the epistemological counterpart of the real issue. He thought that the real issue was how the facts could determine what that speaker meant by his words and what he believed. Lewis conceived of the relevant facts as the totality of physical facts about a speaker. So Lewis' question was: given the constraints of all the physical facts how can we solve for a speaker's meanings and beliefs? He thought that discerning some general principles that could take us from all the physical facts to an interpretation would amount to stating a general theory of persons. This general theory would offer a set of principles that tell us how belief, meaning, behaviour and sensory input relate to one another.

Lewis's principles are five-fold and we need not enter into the details of each of them here. Amongst them, Lewis includes a POC. He calls it the "Improved POC".

“Improved” because there is a provision to make room for explicable error. Davidson found this provision in keeping with DPOC as he had conceived of it:

The improved principle of charity says why assume that Karl is right when you can see that he’s behind a post and can’t notice what’s going on behind it, and so forth. Surely we can do better by allowing for explicable error. I agree. The improved principle of charity, insofar as it says there are cases where you can make exceptions right from the beginning, is what I espouse.¹⁵⁸

On Davidson’s view the attributions of belief we make at the outset are “something we do in order to get our method working and not something independently testable belief by belief.”¹⁵⁹ Davidson requires of interpreters that in order to share subject matter with speakers they must share many beliefs and it may be the case that an explicable but false belief on the part of the speaker ensures more understanding, more shared belief with the interpreter, than inexplicable falsehoods.

Lewis adopts a broader construal of the evidence than Davidson’s. He claims that they relevant facts are all the physical facts. But even Lewis draws the evidence too narrowly and this is one way in which the suggestions of this thesis may be distinguished from what Lewis was pressing for. Firstly, he admits only evidence described in physical terms, despite his own suggestion that, as ontological parsimony is not the topic, we may as well assume any entities that will be of use. Secondly, he claims that as the concepts of belief, desire and meaning are common concepts “the theory that implicitly defines them had better be common property too”.¹⁶⁰ He thinks that the theory must be platitudinous, only putting common sense in a coherent and perspicuous way. Therefore his theory cannot draw on evidence that would only be available to those with special knowledge, because this goes beyond common sense. He thinks that scientific evidence must be barred “on pain of changing the subject”.¹⁶¹ Whilst we may not need any such knowledge to acquire the concepts of belief, desire

¹⁵⁸ 1974c, pp.283.

¹⁵⁹ 1974c, pp.283.

¹⁶⁰ Lewis, 1974.

¹⁶¹ Lewis, 1974.

and meaning, such knowledge may be relevant to working out what it is that someone believes, desires or means, especially if that person is very unfamiliar to us. The concept of belief may be a commonly owned one but working out what exactly Karl Marx believed may be a very tricky matter, and similarly working out what a Martian believes, if anything at all, may draw on knowledge not available to all those who have the concept of belief. Why should there not be better and worse interpreters, where the best may be distinguished by their command of a wealth of information about human nature and their ingenuity in interpreting unfamiliar people? Where explicable error is Lewis's concern, his claims for the insuperability of folk psychology amount to the claim that error can only be explained in ways that do not draw on special knowledge. If this were straightforwardly correct then he would put empirical psychologists out of a job.

Davidson's reply to Lewis illuminates his position on the evidence available to interpreters and the way it is to be utilised.¹⁶² Davidson notes that in common with Lewis he looks to make speakers intelligible rather than maximize agreement between speaker and interpreter. Davidson is sympathetic to Lewis's claim that this will involve taking into account probable errors due to bad positioning (equally, probable successes due to good positioning), deficient (or super-efficient) sensory apparatus, and differences in background knowledge. Davidson is amenable to Lewis suggestion that interpretations be checked against non-verbal as well as verbal behaviour:

there is no chance of telling when a sentence is held true without being able to attribute desires and being able to describe actions as having complex intentions.¹⁶³

As has been pointed out already, Davidson equivocates over whether any of this shows that the evidential base as he construed it was inadequate. He admits that it may be. But this is confusing. Perhaps it is a point of emphasis and Davidson thought that the kind of factors Lewis was making explicit were not excluded from the

¹⁶² 1974c

¹⁶³ 1975, pp.162.

framework he had described. It is hard to square this with the behaviouristic statements about the interpreter's evidential base quoted in 4.4.

A clearer instance of disagreement lay in their respective answers to the question of how physical descriptions of speakers are related to descriptions of their beliefs and meanings. It was Lewis' contention that a correct and complete physical description of the speaker *determines* that speaker's beliefs. Davidson distinguished two senses that determine could have here. In Lewis' sense, determination means that given a complete physical description of the speaker one could in some way read off an agent's beliefs and meanings. Davidson did not accept that knowledge of a complete physical description of a speaker could yield knowledge of a speaker's beliefs and meanings. He did not accept that a speaker's beliefs were determined by the totality of physical facts about a speaker in this sense.

concepts like those of belief and meaning are, in a fundamental way, not reducible to physical, neurological, or even behaviouristic concepts.¹⁶⁴

What Davidson did believe was that beliefs supervened on the totality of physical facts about a speaker and his environment; one way of putting this is to say that there could be no change in a speaker's beliefs without some change in the physicality of the speaker and his environment. It is interesting to consider how Lewis' method of interpretation would work. One might wonder how his view of the determination of belief and meaning by the physical facts fits with his dictum that the interpreter cannot make use of esoteric scientific knowledge. He would have to say that belief and meaning were determined by the mundane physical facts, but it is far from clear how this could be so. One might wonder what role his principles of interpretation can play when what he claims is really required by the interpreter is a survey of the physical facts.

¹⁶⁴ 1974a, pp.154.

My main concern is Davidson's openness or otherwise to the view that there may be information given in a natural description of a speaker and his environment that *his method wastes*.¹⁶⁵ One need not take a position on the question of determination to agree with Lewis on this point. All one need accept is the broad and seemingly uncontroversial thesis that knowledge of a speaker's nature may be *relevant* to assessing what his beliefs may be. This is consistent with Davidson's Anti-Reductionist requirement. Davidson, referring first to theories of meaning and belief and then to the physical and other sciences, takes it that the requirement:

sets these theories forever apart from those that describe mindless objects, or describe objects as mindless.¹⁶⁵

Perhaps he sets them too far apart, further apart than Anti-Reductionism requires. The psychological sciences, as practiced, seem to make use of the notions both of intentional notions, and of more physicalistic notions.

There is empirical evidence that Davidson has effectively thrown away. Against Lewis, this claim should not be restricted to evidence described in scrupulously physical terms. Davidson seems not to have taken this idea very seriously. Perhaps this is because he denied that the whole evidential basis for interpretation could come from the physical description of a speaker. But he does not seem to notice that finding physical and other non-behavioural evidence relevant does not commit one to saying that we can give a reduction of belief and meaning. Just finding some physical evidence relevant does not even commit us to saying that all the evidence relevant to interpretation can be described in non-intentional terms. Some of the empirical evidence relevant to understanding a speaker may come from the human sciences. Such evidence may not turn out to be describable in brute physical terms; fields such as empirical psychology and linguistics have theoretical terminology of their own and may make use of intentional and non-intentional language.

¹⁶⁵ *ibid*

Of course, there are questions about where exactly such extra evidence may come from, what form it may take and how it might be utilised by the Radical Interpreter. Let us bear in mind that there are many areas of empirical work in which people take themselves to be furthering our understanding of minded, language users. It would be surprising to say the least if they had nothing to say to us on such matters as when and how certain beliefs were formed in subjects, what factors are relevant to the kinds of language spoken by subjects and how human physiology relates to a subjects' interaction with their environment. It seems careless if the interpreter is discarding such information in forming his hypotheses and implausible that such evidence is irrelevant. There is a question about how such evidence is utilised in discerning and explaining a subject's beliefs; we should not assume that there is a single, uniform answer. The claim is not that the findings of empirical psychology or any other science, actually or potentially, give us unproblematic access to the fine grained propositional attitudes of speakers. What is being claimed is that such knowledge can be relevant to pinning down a speaker's beliefs; it can constrain us in the beliefs that we can plausibly attribute. To the extent that empirical knowledge can hold belief or meaning fixed it leaves DPOC redundant.

In making erroneous beliefs and utterances explicable we are, as Quine noted, doing applied psychology:

we're thinking: How would this man have learned his own language? How likely is it that he as a child would have picked up a language that had such funny complex twists when translated over into English? How would he learn it, from the point of view of conditioned reflexes and the rest?¹⁶⁶

Empirical work that sets out to answer such questions can, at a minimum, act as a constraint on the beliefs and meanings the interpreter attributes and to the extent that it does so it is fulfilling the role of DPOC.

¹⁶⁶ Quine, 1974.

Davidson saw that bringing in information from descriptions of a speaker's non-verbal behaviour can be important in understanding verbal and non-verbal behaviour:

if the way I described my method left that out, then that is indeed a fault.¹⁶⁷

But he implies in the remark below that such information is only to be used as a method of checking once one already has a theory of interpretation. Davidson's thought is that the speaker's actions described in our language rest on a certain interpretation of that speaker's beliefs and desires; and this seems correct.

In other words there's something to test the theory against besides just speech behaviour and that's the rest of behaviour. And I don't say that in passing as if it were only a little difference. I think it's a big and important one.¹⁶⁸

The difference is bigger still when we view the relevant evidence more broadly as including non-verbal, non-behavioural evidence. The difference in the theory we get will be bigger when we draw on verbal and non-verbal behaviour, the findings of empirical psychology, linguistics, optics, biology and the rest.

¹⁶⁷ 1974, pp.283.

¹⁶⁸ *ibid*

5. Some Positive Suggestions and Some Conclusions

Reviewing the remarks of the fourth section what is presented is largely a negative claim about Davidson's views on interpretation and the nature of the evidence available to the interpreter. The major claim is that Davidson has characterized the evidence too narrowly and that there may be more scope for the use of interesting information about speakers than his method makes room for. There are positive claims that can be extracted from this critique; namely that the evidence for interpretation should be viewed in an open-minded and broad way, and that we should be alert to the fact that evidence from the human sciences often discriminates between the contributions of meaning and belief to verbal behaviour. If the contributions can be distinguished, even if one employed some principle like DPOC on pragmatic grounds when other evidence is lacking, DPOC no longer has the status of an a priori necessity of interpretation. The critique is also suggestive of a certain way of thinking about bringing evidence to bear on interpretation. Putting the discernment of meaning and the Chomskian approach to one side, here is a sketch of one way this might go.¹⁶⁹

The English language has an apparatus of individuation including, amongst other things, expressions of identity and demonstratives. An expression can only be regarded as a predicate if it interacts with the apparatus of individuation, in other words, if it is predicated of something. Quine claimed that the stimulus meanings of the sentences in which predicates occur provide the only evidence for the extension of the predicate. On Davidson's view the observable conditions of assent and dissent behaviour provide the only evidence but there are constraints on the extension we could construe the predicate as having owing to the compositionality of a theory of meaning. Both found the identification of the native's apparatus of individuation to be

¹⁶⁹ My argument has been that one should be open-minded about the sources of evidence for interpretation. My examples, whilst loosely drawn from empirical sciences, lack the depth and subtlety that some background in those fields would provide.

empirically underdetermined – with suitable adjustments elsewhere we could treat, say, the identity predicate as some distinct equivalence relation.¹⁷⁰

A cognitive psychologist might form the empirical hypothesis that a speaker's individuating apparatus is dependent on, amongst other things, his perceptual mechanisms. So whether some expression is that of identity or of some distinct equivalence relation would be dependent, in part at least, on factors to do with the subject's vision. If this hypothesis could be supported by empirical data, then in determining the relation of a native's individuating apparatus to our own we ought to find out such things as how that native's eyes function: do they function in much the way that ours do? Are the mechanics of our eyes that allow us to see objects in three dimensions, to see in colour and to have uninterrupted sight all reproduced in the native? Are their eyes and other sensory organs connected up with their brains and nervous system in much the way ours are? How does the language spoken by the native vary when these factors are varied – do peculiarities or deficiencies have a discernable effect? With Behaviourism out the way, the answers to all these questions and more are relevant to understanding a speaker, and interpreting his individuating expressions. Whether such factors are determinants of the way a speaker individuates becomes an issue to which empirical knowledge is admissible. It may turn out that people with very similar perceptual mechanisms, psychological capabilities and neural make-ups can come to individuate in vastly different ways, but we should not prejudge such questions.

If one learns that a native speaker's eyes work so as to track enduring, medium-sized objects against contrasting backgrounds then that may be relevant to the matter of what the individuating apparatus of his language is and which beliefs one could plausibly attribute to him on given occasions. If one discovered that the native's eyes work so as to track enduring and relatively homogenous objects as wholes against contrasting backgrounds, as opposed to tracking the background or the gaps between objects, we might take this as a hypothetical maxim of interpretation or as one piece

¹⁷⁰ 1979b. Quine, 1969a.

of defeasible evidence for interpretation. Here is what Quine had to say about the translator who tries to make use of such information:

If he were to become conscious of this maxim, he might celebrate it as one of the linguistic universals, or traits of all languages, and he would have no trouble pointing out its psychological plausibility. But he would be wrong; the maxim is his own imposition on what is objectively indeterminate. It is a very sensible imposition and I would recommend no other. But I am making a philosophical point.¹⁷¹

This remark is hard to construe but the “philosophical point” must be one about Behaviourist scruples. Quine says that the hypothetical maxim would be an imposition on what is indeterminate but put like that it seems he is taking the indeterminacy of translation as a premise from which to argue that no other information could be objective evidence for translation. But then he has made a mistake: the indeterminacy thesis is a consequence of his Behaviourism not the grounds for it. Consider what would follow if it were true that such hypothetical maxims about speakers could only ever be our own imposition. The point would extend not only to interpretation of speakers’ apparatus of individuation but to many of the generalizations of empirical psychology and linguistics. The question to which we are owed an answer is why the knowledge acquired in empirical psychology can never be objective evidence in translation or interpretation and only ever be rule of thumb denied objective status. It is hard to explain why the philosopher who thought epistemology should fall into place as a chapter of natural science was so dogmatic about language.

Davidson’s method suffers as a consequence of Quine’s behaviouristic influence. But what Davidson has provided is a framework for thinking about language that can be separated from this behaviouristic view of the relevant evidence. To return to the three conditions on a theory of meaning which Davidson set out; it is his attempt to meet condition two that has been criticised. The requirement was that a theory of meaning be confirmable without detailed knowledge of the propositional attitudes of

¹⁷¹ Quine, 1969a, pp.34.

the speaker. This thesis took issue with Davidson's claim that there is no way into a speaker's meanings or beliefs independently of one another but left the over-arching Davidsonian framework in tact. So there is room for an approach that accepts the framework provided by Davidson for thinking about issues in the philosophy of language whilst taking a broader view of the evidence relevant to the study of language.

Bibliography

- Chomsky, N. 1959 "A Review of B. F. Skinner's Verbal Behaviour". In *Language*, 35 No.1: pp.26–58
- 1975 "Quine's Empirical Assumptions". In *Words and Objections*. Revised ed. D. Davidson and J. Hintikka: pp. 53-69. Boston: D. Reidel Publishing.
 - 1990 "Language and Problems of Knowledge". In *The Philosophy of Language* 4th edition, ed. A. P. Martinich: pp. 581-598
- Davidson, D. 1965 "Theories of Meaning and Learnable Languages". In *Inquiries into Truth and Interpretation*: pp. 3-15. Oxford: Clarendon Press.
- 1967 "Truth and Meaning". In *Inquiries into Truth and Interpretation*: pp. 17-36. Oxford: Clarendon Press.
 - 1968 "On Saying That". In *Inquiries into Truth and Interpretation*: pp. 93-108. Oxford: Clarendon Press.
 - 1970a "Semantics for Natural Languages". In *Inquiries into Truth and Interpretation*: pp. 55-64. Oxford: Clarendon Press.
 - 1970b "Mental Events". In *Essays on Actions and Events*: pp. 207-224. Oxford: Clarendon Press.
 - 1973a "Radical Interpretation". In *Inquiries into Truth and Interpretation*: pp. 125-40. Oxford: Clarendon Press.
 - 1973b "In Defence of Convention T". In *Inquiries into Truth and Interpretation*: pp. 65-75. Oxford: Clarendon Press.
 - 1974a "Belief and the Basis of Meaning". In *Inquiries into Truth and Interpretation*: pp. 141-154. Oxford: Clarendon Press.
 - 1974b "On the Very Idea of a Conceptual Scheme". In *Inquiries into Truth and Interpretation*: pp. 183-198. Oxford: Clarendon Press.
 - 1974c "Replies to Lewis and Quine". In *Inquiries into Truth and Interpretation*: pp. 281-285. Oxford: Clarendon Press.
 - 1975 "Thought and Talk". In *Inquiries into Truth and Interpretation*: pp. 155-170. Oxford: Clarendon Press.

- 1976 "Reply to Foster". In *Inquiries into Truth and Interpretation*: pp. 171-179. Oxford: Clarendon Press.
- 1977a "Reality without Reference". In *Inquiries into Truth and Interpretation*: pp. 215-25. Oxford: Clarendon Press.
- 1977b "The Method of Truth in Metaphysics". In *Inquiries into Truth and Interpretation*: pp. 199-214. Oxford: Clarendon Press.
- 1979 "The Inscrutability of Reference". In *Inquiries into Truth and Interpretation*: pp. 227-241. Oxford: Clarendon Press.
- 1982 "Rational Animals". In *Subjective, Intersubjective, Objective*: pp. 95-107. Oxford: Clarendon Press.
- 1983 "A Coherence Theory of Truth and Knowledge". In *Subjective, Intersubjective, Objective*: pp. 137-153. Oxford: Clarendon Press.
- 1986 "A Nice Derangement of Epitaphs". In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, ed. E. Lepore. Cambridge: Basil Blackwell
- 1987a "Knowing One's Own Mind". In *Subjective, Intersubjective, Objective*: pp. 15-38. Oxford: Clarendon Press.
- 1987b "Afterthoughts". In *Subjective, Intersubjective, Objective* pp.154-157. Oxford: Clarendon Press.
- 1988 "Epistemology and Truth". In *Subjective, Intersubjective, Objective*: pp. 171-191. Oxford: Clarendon Press.
- 1990 "Epistemology Externalised". In *Subjective, Intersubjective, Objective*: pp. 193-204. Oxford: Clarendon Press.
- 1991 "Three Varieties of Knowledge". In *Subjective, Intersubjective, Objective*: pp. 205-220. Oxford: Clarendon Press.
- 1992 "The Second Person". In *Subjective, Intersubjective, Objective*: pp. 107-123. Oxford: Clarendon Press.
- 1994 "Radical Interpretation Interpreted". In *Philosophical Perspectives*, vol. 8; Logic and Language: pp. 121-128.
- 1996 "The Folly of Trying to Define Truth". In *The Journal of Philosophy*, vol. 93 no.6: pp. 263-278

- 1997a "The Emergence of Thought". In *Subjective, Intersubjective, Objective*: pp. 123-34. Oxford: Clarendon Press.
- 1997b "Indeterminism and Antirealism". In *Subjective, Intersubjective, Objective*: pp. 69-84. Oxford: Clarendon Press.
- 2001a *Essays on Actions and Events*. 2nd ed., Oxford: Clarendon Press.
- 2001b *Inquiries into Truth and Interpretation*. 2nd ed., Oxford: Clarendon Press.
- 2001c *Subjective, Intersubjective, Objective*. 1st ed., Oxford: Clarendon Press.
- Davidson, D. and Hintikka, J. 1969 *Words and Objections*. Revised ed., Boston: D. Reidel Publishing.
- Evans, G. 1975 "Identity and Predication". In *The Journal of Philosophy* vol.72 no.13 pp.343-363
- Evans, G. and McDowell, J. (ed.) 1976 *Truth and Meaning. Essays in Semantics*. Oxford: Clarendon Press.
- Foster, J. 1976 "Meaning and Truth Theory". In *Truth and Meaning: Essays in Semantics* ed., Evans, G. and McDowell, J.
- Grandy, R. 1973 "Reference, Meaning and Belief". In *The Journal of Philosophy* vol.70 no.14: pp. 439-452
- Horwich, P. 1992 "Chomsky Versus Quine on the Analytic-Synthetic Distinction". In *The Proceedings of the Aristotelian Society*: pp.95-108
- Larson, R. and Segal, G. 1995 *Knowledge of Meaning*. Cambridge: M.I.T. Press.
- Lewis, D. 1974 "Radical interpretation". In *Synthese* vol. 27: pp. 331-344
- McDowell, J. 1984 "Wittgenstein on Following a Rule". In *Rule-Following and Meaning* ed. Miller, A. and Wright, C.: Acumen.
- McGinn, C. 1977 "Charity, Interpretation and Belief". In *The Journal of Philosophy* vol. 74 no.15: pp. 521-535.
- McNeill, W. 2004 *Triangulation* MPhil Thesis
- Miller, A. 1998 *Philosophy of Language*. UCL Press.
- Quine, W. V. 1960 *Word and Object* 1st ed., Cambridge: M.I.T Press.
- 1969a *Ontological Relativity and Other Essays*. 1st ed., New York: Columbia University Press.
- 1969b "Reply to Chomsky". In *Words and Objections*. Revised ed.: pp. 302-311.

- 1970 "On the Reasons for the Indeterminacy of Translation". In *The Journal of Philosophy*, vol.67 no.4: pp. 178-183
 - 1974 "Comment on Donald Davidson". In *Synthese* vol. 27: pp. 325-329
 - 1987 "The Indeterminacy of Translation Again". In *The Journal of Philosophy* vol. 84 no.1: pp. 5-9
 - 1996 "Progress on Two Fronts". In *The Journal of Philosophy*, vol.93 no.4: pp. 159-163
- Ramberg, B. 1989 *Donald Davidson's Philosophy of language: An Introduction*. 1st ed., Oxford: Basil Blackwell.
- Smith, B. 1992 "Understanding Language". In *Proceedings of the Aristotelian Society*: pp.109-141
- Unpublished "Comments on Peter Johnson's 'Omniscient Interpreters and the Principle of Charity'".
- Tarski, A. 1935 "The Concept of Truth in Formalized Languages". In *Logic, Semantics and Metamathematics*. Oxford: Clarendon Press, 1956.
- 1944 "The Semantic Conception of Truth and the Foundations of Semantics". In *The Philosophy of Language* 4th edition, ed. A. P. Martinich: pp. 69-91
- Wittgenstein, L. 2001 *Philosophical Investigations* 3rd ed., Oxford: Basil Blackwell.
- Wright, C. 1993 *Realism, Meaning and Truth* 2nd ed., Oxford: Basil Blackwell.