

Davidson on belief, truth, and the sceptic

MSc Thesis (*Afstudeerscriptie*)

written by

Guus Willem Eelink

(born 7 October 1990 in Winterswijk, The Netherlands)

under the supervision of **Prof. Dr Martin Stokhof**, and submitted to
the Board of Examiners in partial fulfillment of the requirements for the
degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense: **Members of the Thesis Committee:**
8 September 2014

Dr Maria Aloni
Dr Paul Dekker
Prof. Dr Robert van Rooij
Prof. Dr Martin Stokhof



INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

In this thesis the reader finds a critical discussion of Donald Davidson's conception of truth, and of Davison's claim that a proper understanding of the interrelatedness of meaning, truth, and belief leads to the conclusion that scepticism is unintelligible. Davidson argues that, instead of defining or analysing the notion of truth in isolation, we should rather try to understand how it relates to other fundamental notions, like that of meaning and of belief, and what role truth plays in an understanding of rationality. Understanding this for Davidson gives rise to the idea that our conception of the world must be globally adequate (since belief is by nature veridical) and it explains how we can have knowledge of the world and why scepticism, in as far as it goes against the idea that belief is veridical or that knowledge is possible, is unintelligible. In the thesis we try to reconstruct in detail how Davidson approaches these concepts and arrives at the anti-sceptical conclusion. We shall do this on the basis of all of Davidson's works pertaining to this issue, paying attention to chronological development in Davidson's ideas. The general aim of the thesis is to accomplish a critical understanding of this crucial part of Davidson's thought and to show that Davidson succeeds in offering a systematic account of the interrelatedness of truth, meaning, and belief by adopting a third person perspective on epistemology and by explaining the role of these notions in a theory of rational behaviour, and particularly in a theory of linguistic communication. We shall also try to show why Davidson wants to understand truth and the related notions from this perspective.

Keywords: Donald Davidson, truth, belief, meaning, scepticism, knowledge, philosophy of language, epistemology.

Acknowledgements

I would like to thank my supervisor, Martin Stokhof, for his useful advice and the nice and interesting meetings we had to discuss my thesis work. Martin is always very kind and encouraging. I am also grateful to him for his wonderful course called Radical Interpretation, Hermeneutics, and Forms of Life, which inspired me to write my thesis on Davidson.

I also want to thank my family, friends, and fellow students for their support.

In particular I want to thank Seyed, with whom I walked up and down the long corridors at Science Park so many times in the period I was writing this thesis, on our way to and back from the most distanced coffee machines, discussing numerous funny and interesting topics.

Truth is one of the clearest and most basic concepts we have, so
it is fruitless to dream of eliminating it in favor of something
simpler or more fundamental.

The truth of an utterance depends on just two things: what the
words as spoken mean, and how the world is arranged. There is
no further relativism to a conceptual scheme, a way of viewing
things, or a perspective. Two interpreters, as unlike in culture,
language, and point of view as you please, can disagree over
whether an utterance is true, but only if they differ on how
things are in the world they share, or what the utterance
means.

What Tarski has done for us is show in detail how to describe
the kind of pattern truth must make, whether in language or in
thought. What we need to do now is say how to identify the
presence of such a pattern or structure in the behaviour of
people.

—Donald Davidson

... no man is the lord of anything,
Though in and of him there be much consisting,
Till he communicate his parts to others;
Nor doth he of himself know them for aught
Till he behold them formed in th'applause
Where they're extended.

(*Troilus and Cressida*, III. iii. 115-120. Cited by Davidson in
Thought and Talk.)

Contents

| | |
|--|-----------|
| Introduction | 1 |
| 1 Overview: Davidson on meaning, belief, and truth | 7 |
| 1.1 Introduction to chapter 1 | 7 |
| 1.2 Truth in the theory of meaning: Tarski's truth theories | 8 |
| 1.2.1 Tarskian truth theories in Davidson's work | 8 |
| 1.2.2 The use of truth theories as meaning theories | 9 |
| 1.3 Truth theories and radical interpretation | 14 |
| 1.3.1 Davidson's thesis that meaning is publicly accessible and radical interpretation | 14 |
| 1.3.2 Davidson's version of Convention <i>T</i> and the empirical constraints on truth theories | 19 |
| 1.3.3 Breaking into the interdependence of meaning and be- lief: the principle of charity | 25 |
| 1.4 Causes, triangulation, and veridicality | 29 |
| 1.5 The concept of truth | 32 |
| 2 Radical interpretation: evidence and constraints | 36 |
| 2.1 Introduction to chapter 2 | 36 |
| 2.2 Hold true attitudes as primitive evidence | 37 |
| 2.2.1 Primitive evidence: hold true attitudes | 37 |
| 2.2.2 The identification of hold true attitudes | 38 |
| 2.3 From evidence to interpretation | 41 |
| 2.3.1 Holistic constraints | 41 |
| 2.3.2 Causal constraints | 42 |
| 3 The veridicality of belief | 45 |
| 3.1 Introduction to chapter 3 | 45 |
| 3.2 Interpretation and veridicality | 45 |
| 3.2.1 The agreement thesis | 45 |
| 3.2.2 The principle of charity and its justification | 47 |
| 3.2.3 From agreement to veridicality | 50 |
| 3.3 Triangulation, objectivity, and veridicality | 55 |

| | | |
|----------|---|------------|
| 3.3.1 | Individuation of the contents of thought | 55 |
| 3.3.2 | No belief without language | 55 |
| 3.3.3 | Triangulation and the common cause | 58 |
| 3.3.4 | The need for an external criterion | 61 |
| 3.4 | Radical interpretation and triangulation | 68 |
| 3.4.1 | A tension? | 68 |
| 3.4.2 | A unified account | 70 |
| 4 | The concept of truth: pattern and content | 72 |
| 4.1 | Introduction to chapter 4 | 72 |
| 4.2 | Rejection of existing conceptions of truth | 73 |
| 4.2.1 | Sentences as truth bearers | 73 |
| 4.2.2 | Rejection of realist, anti-realist, and pragmatist con- ceptions of truth | 74 |
| 4.2.3 | Rejection of deflationary conceptions of truth | 78 |
| 4.3 | Tarski's truth theories and the concept of truth | 79 |
| 4.3.1 | The philosophical significance of Tarski's work | 79 |
| 4.3.2 | Davidson's behavioural approach to truth | 83 |
| 4.3.3 | Truth as primitive | 86 |
| 4.3.4 | First truth, then ontology | 88 |
| 4.4 | Davidson's realism | 89 |
| 4.4.1 | Verification-transcendence | 89 |
| 4.4.2 | Indeterminacy and realism | 93 |
| 5 | Davidson's attitude towards the sceptic | 95 |
| 5.1 | Introduction to chapter 5 | 95 |
| 5.2 | Veridicality and scepticism | 96 |
| 5.3 | The veridicality thesis as argument against scepticism | 97 |
| 5.4 | Davidson's new <i>Cogito</i> against the sceptic: context and mo- tivation | 98 |
| 5.5 | Does Davidson leave room for the sceptic? | 105 |
| 5.5.1 | Two readings of the veridicality thesis | 105 |
| 5.5.2 | Veridicality and the possibility of knowledge | 108 |
| 5.5.3 | Assuming that one thinks | 112 |
| | Conclusions and suggestions for further research | 116 |
| | Abbreviations of works by Davidson | 119 |
| | Bibliography | 123 |

Introduction

Donald Davidson has left us a colossal intellectual legacy. His works are a treasury of deep ideas on almost all aspects of philosophy. Our concern in this thesis is Davidson's conception of truth. We are particularly interested in Davidson's ideas on the connections between truth and belief, and in his view on scepticism.

Davidson's programme in philosophy can in a sense be seen as an attempt to 'clean up' philosophy by attacking a range of presuppositions ingrained in the tradition that Davidson thinks stand in the way of a common sense realism and give rise to scepticism. Particularly, Davidson spent a lot of ink to attack the remnants of empiricism. He followed his mentor W.V. Quine in rejecting the analytic-synthetic distinction, but then proceeded to attack another (and presumably the last) tenet of empiricism, that Davidson also found in Quine's thought: the dualism of scheme and content. In many of his essays Davidson argues against these and other ideas that he thinks block our being directly 'in touch' with reality: the idea that there are conceptual schemes, epistemic intermediaries in between our beliefs and the world, or 'objects of the mind' that we grasp while we think. Davidson tries to show that these philosophical constructions rest on untenable assumptions.

Further, and more in general, Davidson thinks we should give up the attempt to provide a grounding of knowledge claims in something certain, or something independent of the realm of beliefs. Davidson holds that a belief can only be supported by another belief. Our beliefs (as a whole) are not *based* on something: they are merely *caused* by things. Davidson thinks that it is neither needed nor possible to give a foundation for knowledge claims on something certain outside the realm of belief.

Davidson argues that, once we abandon the ideas and attempts mentioned above, we can adopt a common sense realism, according to which there is nothing in between the mind and the world that prevents us from directly knowing the world: the mind (or language) neither creates nor distorts the world. In particular Davidson thinks that this will allow us to stop the threat of scepticism: the idea that reality is not fully knowable for us, or that reality as we perceive it may be a construct of thought.

This attempt to vindicate a common sense realism is mirrored by Davidson's conception of truth, which we shall be concerned with. Davidson tries

to find a ‘middle position’ on truth in the philosophical landscape. He rejects both metaphysical realism and anti-realism. For Davidson truth is independent of what we believe or what we (potentially) have evidence for, but our beliefs are not independent of the true and the real: our beliefs are globally veridical.

This brings us to the positive side of Davidson’s programme, which we are interested in in this thesis. Davidson has more to offer than just an attack on old ideas. The best way to illustrate this is by means of the following citation from the *Afterthoughts to A Coherence Theory of Truth and Knowledge*, in which Davidson contrasts his ideas with Rorty’s view on philosophy:¹

I set out not to ‘refute’ the sceptic, but to give a sketch of what I think to be a correct account of the foundations of linguistic communication and its implications for truth, belief, and knowledge. If one grants the correctness of this account, one can tell the sceptic to get lost.

Where Rorty and I differ, if we do, is in the importance we attach to the arguments that lead to the sceptic’s undoing, and in the interest we find in the consequences for knowledge, belief, truth, and meaning. Rorty wants to dwell on where the arguments have led: to a position which allows us to dismiss the sceptic’s doubts, and so to abandon the attempt to provide a general justification for knowledge claims—a justification that is neither possible nor needed. Rorty sees the history of Western philosophy as a confused and victorless battle between unintelligible scepticism and lame attempts to answer it. Epistemology from Descartes to Quine seems to me just one complex, and by no means unilluminating, chapter in the philosophical enterprise. If that chapter is coming to a close, it will be through recourse to modes of analysis and adherence to standards of clarity that have always distinguished the best philosophy, and will, with luck and enterprise, continue to do so.

Davidson claims that his account of the foundations of linguistic communication and its implications for truth, belief, and knowledge imply that he can tell the sceptic to get lost. For Davidson this means that his account implies that he can adopt a ‘middle position’ on truth: belief is by nature veridical, but truth is verification-transcendent. The veridicality of

¹Cf. AF, p. 157. This passage particularly reflects on *A Coherence Theory of Truth and Knowledge*, but we think it is programmatic for Davidson’s thinking on truth and scepticism in general. When referring to works by Davidson in the footnotes, we shall use abbreviations, like ‘AF’ (=Afterthoughts). These abbreviations are specified below, see: Works by Donald Davidson, with abbreviations. As we explain there, page numbers refer to the volumes with collected essays.

belief explains how our beliefs are connected to the world and how we can have knowledge, without the need for a grounding of knowledge claims in something certain.

Our aim in this thesis is to accomplish a critical understanding of Davidson's ideas on the notions of truth and belief and on how they are related. We are particularly interested in three aspects of Davidson's position: his claim that belief is by nature veridical, his conception of truth, and the consequences of Davidson's position for certain kinds of scepticism. We shall try to show how Davidson succeeds in offering a systematic account of the coherence of meaning, belief, and truth by explaining their role in a theory of rational behaviour. Davidson, like Quine, adopts a third person perspective on epistemology. Unlike for Quine, this does not mean for Davidson that we must explain how knowledge is based on sense data. Davidson offers a different kind of third person perspective: as suggested by the above citation, Davidson wants to illuminate the notions of meaning, belief, and truth on the basis of their role in a theory of linguistic communication (which again is part of a theory of rational behaviour). The reason to do so for Davidson is that he thinks that the notions involved (belief, truth, and knowledge) have their most primary application in such a theory. Davidson holds that linguistic communication is a precondition for having thought, and that therefore one can only have beliefs if one communicates. One of our aims is to try to show how Davidson justifies this perspective on epistemology by offering an argument against the conceivability of a 'solitary thinker' (someone with thought who never communicated). If this argument succeeds, Davidson's argument against the sceptic is far more successful.

Our task will not be easy. Davidson's ideas are scattered throughout his many essays and his only (short) book. Davidson never wrote down his ideas on any issue in a comprehensive, systematic treatise (with the possible exception of *Truth and Predication*). It is hard to appreciate his essays in isolation. This is not so because of obscure or idiosyncratic terminology (Davidson's sentences are mostly clear and short, and Davidson is even slightly allergic to technical philosophical vocabulary), but because of the extreme argumentative density of his writings. Another complicating factor is the holism of his ideas. Davidson's account of a specific problem in one of his essays is always informed and supported by what he has already written down in other essays. Davidson's essays often need many intensive readings before one can really appreciate what he is after. In trying to understand Davidson's view on truth we therefore have to take into account all the relevant sources, to arrive at a holistic interpretation.

Marc Joseph writes the following in the introduction to his overview of Davidson's philosophy:²

²Cf. Joseph (2004), p. 9.

The reader will find in the following chapters a sympathetic reading of Davidson's philosophy, not because he has the right or final answers to all the outstanding questions that define the contemporary scene in philosophy, but because, first, he offers what few other philosophers do today, namely, an over-arching theory of persons as rational animals. Secondly, this interpretation is undertaken in the spirit of the principle of charity—about which the reader will hear a great deal in what follows—in that the best way to understand a difficult thinker's work is to see it as making the best overall sense we can.

We follow Joseph by interpreting Davidson in the spirit of the principle of charity: our foremost interest is not to criticize specific points of Davidson's account, but rather to achieve a critical understanding. For an extremely critical, detailed, and comprehensive interpretation of Davidson's philosophy, including many of the issues we shall discuss, we refer the reader to Lepore and Ludwig's impressive book on Davidson's thought.³ We are interested in offering a more holistic perspective on Davidson's ideas on meaning, truth, and belief, and in detecting the basic principles that inform these ideas, and the implications they have for scepticism.

We shall do this by first giving an overview of central ideas in Davidson's philosophy of language and epistemology. This will form the background to the following chapters. In the following chapters we shall then discuss in more detail various aspects of Davidson's conception of truth and belief, and (finally) their implications for scepticism (and for Davidson's conception of the possibility of knowledge).

In the first part of the first chapter we shall start by giving an overview of Davidson's ideas on philosophy of language, particularly of his conception of a theory of meaning, which we shall need as a background to our discussion in the following chapters. We shall start by discussing Davidson's conception of what a theory of meaning should accomplish and how a theory of meaning for Davidson must be seen as part of a theory of rational behaviour: a normative explanation of behaviour in terms of beliefs and desires. Davidson argues that a theory of meaning for a language should have the form of a semantic, Tarski-style theory of truth for that language. Davidson explores by means of the notion of radical interpretation how an interpreter can find out which truth theory applies (or, given indeterminacy, which truth theories apply) to a speaker's language, without presupposing semantic facts. Whereas Tarski took for granted an understanding of meaning and was interested in exploring the notion of truth, Davidson takes truth to be a primitive concept and uses it to shed light on meaning. We shall see in chapter 4 that Tarski's theories play a crucial role in Davidson's understanding of truth.

³Cf. Lepore and Ludwig (2005).

In the last part of the first chapter we shall briefly discuss two issues that are of importance to an understanding of the following chapters. The first is Davidson's causal theory of content, according to which our perceptual beliefs are about what causes them: the conditions under which we are typically caused to entertain a perceptual belief are the conditions under which such a belief is true (and hence are constitutive of what the belief is about). Davidson's causal theory implies that our perceptual beliefs are globally veridical: most of the times they are true. These ideas will be discussed in more detail in chapter 3. Further, we shall briefly discuss some of Davidson's ideas on truth as a background to chapter 4.

In the second chapter we shall discuss Davidson's idea that an interpretive theory must be based on non-semantic evidence that is available to the radical interpreter. We shall go into the nature of Davidson's primitive behavioural evidence, the *hold true attitude*, and discuss its availability to a radical interpreter. We shall also discuss the formal and empirical constraints that need to be put on truth theories for them to be interpretive.

In the third chapter we focus on Davidson's conception of belief, particularly his idea that belief is by nature veridical. Our concern in the first part is the *veridicality thesis*, the idea that even though each of our beliefs may be false, our beliefs are globally veridical. We shall discuss Davidson's argument for this thesis, which rests on his causal theory of content.

In the second part of the third chapter we shall discuss Davidson's idea that thought and language are interdependent and that having thought depends on having a public language, which for Davidson means a language interpreted by others. Davidson explains this in terms of his 'triangulation argument': the idea that there could be no fact of the matter what our beliefs are about if we did not communicate with others in order to intersubjectively locate what in our environment we are responding to (and which causes inform our beliefs). This argument is controversial and has received a lot of criticism in the secondary literature. We shall nevertheless try to defend the argument and further argue that it essentially reinforces Davidson's veridicality thesis and his argument against the sceptic.

The central concept of the fourth chapter is truth. Our aim in this chapter is to show how Davidson defines a 'middle position', in between realist conceptions of truth that imply that we may be completely mistaken about the world and epistemic conceptions of truth that restrict truth to what is verifiable. We shall further go into Davidson's conception of truth as a primitive, undefinable concept and his ideas on the philosophical significance of Tarski's work for understanding truth, particularly in terms of the role it can play in the theory of linguistic communication.

In the fifth and final chapter we shall discuss the 'application' of everything discussed in the previous chapters to scepticism. In this chapter the need for interpretation goes further than in the previous ones, because it is difficult to formulate precisely what exactly Davidson's position im-

plies for scepticism. We shall distinguish several kinds of scepticism and the consequences Davidson's position has for them. We think that Davidson is particularly interested in the sceptic who denies the possibility of knowledge. We shall try to show how Davidson's position on the one hand undermines a presupposition in the sceptic's challenge (the demand for an ultimate justification) and on the other hand implies that the position of the sceptic is inconsistent.

Chapter 1

Overview: Davidson on meaning, belief, and truth

1.1 Introduction to chapter 1

In the first chapter we shall give a general overview of the most important aspects of Davidson's philosophy of language and (to a lesser extent) of his epistemology.

We shall first discuss Davidson's idea that truth theories can be used as theories of meaning. (1.2) We shall then explain the role of the notion of radical interpretation for Davidson's conception of meaning. (1.3) Following Davidson's formulation, we shall discuss how in terms of these two notions (truth theories and radical interpretation) Davidson provides an answer to the following two questions:¹ (1) What knowledge would enable an interpreter to interpret the language of a speaker? (2) How could the interpreter come to know this? We shall discuss the evidence that the radical interpreter relies on and the constraints that she should satisfy in order to construct an interpretive theory on the basis of this evidence. Evidence and constraints will be discussed further in chapter 2. We shall then briefly discuss Davidson's causal theory of content and Davidson's 'triangulation theory'. Particularly, we shall discuss how these ideas give rise to the veridicality thesis, the idea that it is a conceptual truth that most of our beliefs are true, a thesis that will be elaborated on in detail in chapter 3. (1.4) In the last section we shall briefly treat Davidson's view on the concept of truth, particularly his criticism of (fact-based) correspondence theory in his earlier works. Against this background Davidson's positive ideas on truth will be the subject of chapter 4. (1.5)

¹Cf. RI, p. 125.

1.2 Truth in the theory of meaning: Tarski's truth theories

1.2.1 Tarskian truth theories in Davidson's work

The most obvious way to start a discussion of the concept of truth in Davidson's philosophy is to discuss the importance that Davidson ascribes to Tarski's seminal work on truth and the role that Tarski's truth theories (or Davidson's appropriation of them) play(s) in Davidson's philosophical programme. Tarski's semantic truth theories are a cornerstone of all Davidson writes on truth, which accumulates in his last work, the posthumously published book *Truth and Predication* (2005). The role of Tarskian truth theories is significant (and surprising) for at least the following two reasons.

Truth theories play a fundamental role in Davidson's theory of meaning. Davidson's first appeals to Tarski in *Truth and Meaning* (1967), the programmatic paper for the part of his philosophical project that is concerned with semantics. In this paper Davidson's interest is not in an exploration of truth as such, but he is concerned with the theory of meaning and the way in which natural languages, with no a priori upper bound to the number of sentences they have, can be learned and interpreted by finite creatures.² Davidson argues that an adequate theory of meaning for a language is a definition of the truth predicate for that language (in a metalanguage).

Tarski's work can also shed light on the nature of truth for Davidson. Davidson conceives of truth theories as being useful formalizations of the most important application of truth in our life: its use in interpretation (and more in general in the theory of rationality). Appreciating this leads for Davidson to a deeper understanding of the nature of truth. Some of Davidson's views conflict with Tarski's own views (given that Tarski thought of his theory as a form of correspondence theory).³ It is surprising that Davidson ascribes such importance to Tarski's work on truth, since, if Davidson is right, this importance (both for doing semantics and for understanding truth as such) has been largely overlooked or misunderstood by the whole philosophical tradition after Tarski.⁴

Even though we are mainly concerned with Davidson's conception of truth and not with Davidson's conception of meaning, we have to start at least by elaborating in some detail on Davidson's programme in semantics and how this gives rise to the use of semantic truth theories, which we then shall see are important to Davidson's understanding of truth.

²Cf. TM, p. 17.

³Cf. Tarski (1944), section 3.

⁴Cf. 4.3.1.

1.2.2 The use of truth theories as meaning theories

For Davidson a theory of meaning must show how the meaning of a sentence depends on the meaning of its parts. He endorses the familiar argument that, since someone who speaks a language understands a ‘potential infinity’ of sentences, there must be some way to explain the meaning of sentences in terms of the meaning of their parts, in order to explain how a finite creature can (potentially) understand (and produce) infinitely many sentences. More precisely, it must be possible to explain the meaning of all sentences in terms of a finite vocabulary and a finitely stated set of rules.⁵

Davidson elaborates at some length on various (historical) approaches to the problem of providing a theory of meaning in *Truth and Meaning*, all of which he rejects.⁶

Davidson is particularly critical of approaches that take meanings to be entities, like Frege’s. His main reason for rejecting this approach is simply that he does not see what use meanings have for the theory of meaning. Particularly Davidson thinks it is not clear what entities should be associated to which kind of words (what kind of entity does for example correspond to a predicate?) and what role such entities can play in revealing the structure of a sentence.⁷ Davidson writes about this: ‘Paradoxically, the one thing meanings do not seem to do is oil the wheels of a theory of meaning—at least as long as we require of such a theory that it non-trivially give the meaning of every sentence in the language.’⁸ In chapter 4 it will be further clarified why Davidson thinks meaning-entities are explanatory vacuous.

A more primitive approach Davidson thinks is promising but fails in the end, is to identify meaning with reference. Davidson argues that—on some reasonable assumptions—it would entail that each true sentence refers to the same entity (this is in fact what Frege thought: each true sentence refers to The True, but for Frege that was not the *meaning* of a sentence, but its reference). We shall briefly discuss a version of this argument in 1.5.

Another consideration that plays an important role in Davidson’s requirements for a theory of meaning is meaning holism. Davidson’s holism does not just embrace the relationship between the meaning of words and the meaning of sentences, but also the relationship between the meaning of individual sentences and the meanings expressible by the language as a whole. Words obtain their meaning from the role they play in the sentences in which they occur. For Davidson the same goes (to a great extent) for individual sentences: they only obtain meaning in the context of the meaning of all the other sentences of the language.⁹ Beliefs feature exactly the same

⁵Cf. TM, p. 17.

⁶Cf. TM, pp. 17-22.

⁷Cf. 4.2.1 and 4.3.4.

⁸TM, pp. 20-21

⁹Cf. TM, p. 22.

sort of holism. Davidson therefore requires that a theory of meaning entail the meaning of *all* sentences in a language: only in the context of such an exhaustive theory of meaning can the meaning of individual sentences be fully understood.¹⁰

Given these considerations, we shall now turn to Davidson's motivation for the use of Tarskian theories. In *Truth and Meaning* Davidson presents the following train of thought. We shall not go into all of its steps (but present it in a very coarse-grained way), we are mainly interested in the resulting idea. (We assume at this point, like Davidson, that the object language is part of the language of the theory.)¹¹

Davidson had started considering the idea that we require that a theory of meaning entail all sentences of the form '*s* means *m*' for each sentence *s* (structurally described) of the language for which we give a theory (the object language henceforth). Because of his considerations against meanings as entities, Davidson thinks that *m*, which is supposed to denote a meaning, is a 'troublesome term', which we should get rid of. He therefore tries to consider sentences of another form as theorems of an adequate theory of meaning: '*s* means that *p*', where *p* is replaced by a sentence. Davidson also rejects this version, because of the intensional context 'means that *p*', which he thinks will cause logical troubles at least as complex as providing the theory itself. A theory entailing sentences with 'means that', Davidson thinks, must have a logical structure being able to deal with what are traditionally called intensional contexts, whence the structure of the theory is presumably at least as obscure as the phenomenon that we need the theory for (namely to illuminate meaning). Davidson wants a theory whose entailments can be transformed into sentences specifying meaning, but he wants to give such a theory without explicitly using the concept of meaning. Davidson then offers what he describes as a 'radical solution'. Davidson suggests that the theory has done its work if it provides for every sentence *s* a sentence *p* that in some way gives the meaning of *s*. Davidson suggests to 'sweep away' the non-extensional 'means that' and to replace it by a truth-functional sentential connective, namely *if and only if*, turning the position occupied by *p* into an extensional context. Of course, this does not yet suffice, since the outcome would be of the form '*s*-sentential connective-*p*', which is not a sentence. The reason is that *s* names a sentence, and can therefore not without more ado (i.e. without adding anything to it) be connected with the other sentence. Davidson therefore suggests to supply the description replacing *s* with its own predicate, *T*. What we obtain then *is* a sentence, of the form:

(*T*) *s* is *T* if and only if *p*.

¹⁰Cf. RII, p. 124.

¹¹Cf. TM, pp. 22-23.

Where s , as before, is a structural description of a sentence and p that sentence itself (i.e. the sentence that s describes). Such sentences are called T -sentences by postulation.

Davidson requires that a theory of meaning put enough restrictions on the predicate T to entail all sentences of form (T).

Davidson then observes that a theory that entails all sentences of form (T) is materially equivalent with a definition of truth-in-a-language, because it will characterize the predicate T in such a way that it only applies to all true sentences of the object language L . The requirement that a theory entail T -sentences is materially equivalent with Tarski's Convention T for the material adequacy of formal truth definitions.

In other words, for Davidson a theory of meaning for L , showing 'how the meanings of sentences depend upon the meanings of words' contains a definition of true-in- L .¹² Davidson characterizes this as the somewhat surprising outcome of his search for an adequate theory of meaning: the concept of truth played no role at the outset, but it turned out that a theory of meaning characterizes a predicate that is extensionally equivalent to the truth predicate (assuming that the extension of the predicate T is limited to the sentences of the object language). This is, of course, not really the outcome of an argument. It is rather based on the prior conviction that what is essential to the understanding of the meaning of a sentence is to understand under what conditions it is true. Davidson writes elsewhere that 'having a truth value is the simplest and clearest mark of the unity of sentences'.¹³ In other words: what is the essential feature of sentences for Davidson is that they are entities with a truth value. For Davidson there are no other entities with a truth value (except utterances, which sentences are abstractions over, as we shall discuss in chapter 4).¹⁴ It is therefore a natural idea to Davidson that 'to know the semantic concept of truth for a language [a Tarski-style truth definition] 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'.¹⁵ Knowing a truth theory for a language means knowing the truth conditions of all its sentences, and this, for Davidson, amounts to knowing what is essential to sentences, and hence, Davidson argues, to knowing their *meaning*. In order to be able to use the sentences of a language, we must therefore know under which conditions they are true.¹⁶

Tarski has shown how to construct truth theories for formal languages,

¹²Cf. TM, p. 23.

¹³Cf. TP, p. 120.

¹⁴Cf. 4.2.1.

¹⁵TM, p. 24

¹⁶Cf. TP, pp. 123-124. Cf. 4.2.1. Cf. 4.3.4. Sentences with another mood than the indicative mood have no truth value, but their unity for Davidson consists of their relationship to sentences with a truth value, cf. 2.2. Davidson rejects the anti-realist idea that some indicative sentences have no truth value, cf. 4.2.2.

by appealing to the precise recursive specification of their semantics. He does this by defining a notion of *satisfaction* of a sentence by an assignment function, a function that assigns objects to variables. Satisfaction is axiomatized by a set of non-recursive axioms and a set of recursive axioms. The non-recursive axioms specify the reference of names (if the language contains names) and the satisfaction conditions of predicates (specifying what (ordered sequences of) objects the (relational) predicates are true of). The recursive axioms specify the satisfaction conditions of complex formulas in terms of the satisfaction condition(s) of the simpler formula(s) they are constructed from and the recursive operation applied to the simpler formula(s) to obtain the complex formula(s). (By definition the well-formed expressions of the language are formulas, and only and all formulas with only bound variables are sentences.) Tarski then defines that a sentence is true if and only if it is satisfied by *all* assignment functions. Given these definitions, one can derive the truth conditions for all sentences.¹⁷

Davidson's interest, however, is in a truth theories for natural languages. He therefore faces at least two obvious problems (apart from the fact that it is not obvious that *all* sentences of natural language can be incorporated in a truth theory).¹⁸ The first is that natural language does not have a clearly defined recursive structure, which for Tarski is a requirement to define a theory of truth. Davidson's programme therefore requires that we somehow first systematically translate natural language into a formal language / reveal the recursive structure of natural language.¹⁹

Another problem is that natural languages contain their own truth predicate and are therefore inconsistent: they contain liar sentences. These sentences cannot be dealt with straightforwardly by a truth theory. Arguably this problem can be avoided by defining the theory for a consistent part of the language (the part without the truth predicate), where the inconsistent sentences can be understood on the basis of the results delivered by a theory for the consistent part.²⁰

These problems will not concern us in this thesis, although arguably Davidson's conception of truth is closely intertwined with his idea that truth plays a central role in the theory of meaning (and more in general, in the theory of rational behaviour), and in that sense it can be argued that some of the more general views Davidson proposes on truth cannot stand (or will be weakened) if his semantic programme fails.²¹

The power of truth theories for Davidson is in the fact that they reveal

¹⁷Cf. TF, pp. 47-48.

¹⁸Cf. 2.2.

¹⁹Cf. TM, pp. 29-31. Another problem is of course how to uncover the logical form of intensional verbs (cf. TM p. 30) and (related to that) the need to deal with embedded content clauses ('*that*-clauses'). Cf. 4.2.1.

²⁰Cf. RI, p. 133.

²¹See Conclusions.

semantic structure. If we take the object language to be contained in the metalanguage, the sentences entailed by the theory are all trivial, and they should be. What is ingenious about Tarski's idea is in the fact that the infinitely many trivial *T*-sentences, one for each sentence *s* of *L*, are entailed by a finite list of axioms. In this way the theory provides an explanation of how the truth conditions of complex sentences depend on the truth conditions of simpler sentences or formulas, the reference of names, and the satisfaction conditions of predicates. Tarski used this machinery to define the truth predicate for a formal language in a metalanguage.

Davidson's claim is that a truth theory (given certain conditions) can function as a theory of meaning. Davidson's claim is that knowledge of a truth theory for a language (and knowledge that that truth theory is the (or a) correct one for that language) amounts to knowledge of the meaning of the sentences of that language.²²

In the formulation of Convention *T* above it is assumed that the object language is included in the metalanguage, whence we could simply require that the sentence used at the right hand of the biconditional be the sentence mentioned at the left hand. However, for Davidson the case where the metalanguage includes the object language is just a special case. In case we do not assume that the object language is part of the metalanguage, we cannot require anymore that the sentence at the right hand of the biconditionals is the sentence mentioned at the left side. We could of course require that the sentence at the right side *translates* the sentence at the left side. Davidson thinks that this should be the case for a theory of meaning, but he does not want to put this as an explicit constraint on truth theories, because (as opposed to Tarski) his very aim is to shed light on the nature of meaning and translation. We shall see that his aim is to formulate constraints on truth theories, without appealing to semantic notions like translation, in such a way that the *T*-sentences that are entailed provide translations of the sentences in the object language. In other words: he wants to illuminate what the requirement of translation (in combination with formal requirements) amounts to, instead of assuming an understanding of it. This brings us to another crucial part of Davidson's project: radical interpretation.

²²Cf. RI, p. 125 and RI, pp. 138-139. Davidson is careful to stress that knowledge of a truth theory for a language (and knowledge that it is a truth theory for that language) *would* enable us to understand a language, which does of course not mean that we actually understand a language on the basis of knowledge of a correct truth theory.

1.3 Truth theories and radical interpretation

1.3.1 Davidson's thesis that meaning is publicly accessible and radical interpretation

Clarifying how the semantic structure of natural language can be finitely and recursively specified is only one part of Davidson's project. More in general, Davidson wants to clarify how an interpreter could come to know what is needed to know to interpret a speaker. The first essay in which Davidson systematically elaborates on this question is *Radical Interpretation* (1973).

Radical Interpretation starts with the following two questions concerning an imaginative interpreter's ability to interpret the German sentence 'Es regnet', uttered by an imaginative German guy called Kurt: (1) What could the interpreter know that would enable him to interpret Kurt's utterance? (2) How could he come to know this?²³ *Truth and Meaning* is (mainly) concerned with the first of these questions. The answer provided there by Davidson is that if the interpreter knew a of a certain truth theory that it is a theory for Kurt's language, she would be able to interpret Kurt. The second question concerns what Lepore and Ludwig call Davidson's 'extended project': how can an interpreter find out which truth theory is the correct one for a speaker's language?²⁴ The latter question is explored by means of the concept of radical interpretation, in which the interpreter does not have any knowledge yet regarding the meaning of the speaker's words when she sets out to construct a truth theory for the speaker's language. In this case the test whether the *T*-sentences entailed by a given truth theory are adequate is no longer trivial. In case the object language is part of the metalanguage and the interpreter knows this, she can read off from the form of the *T*-sentences whether they are true or not. The radical interpreter, on the contrary, does not know beforehand which *T*-sentences the truth theory must entail.

Davidson conceives of the task of the radical interpreter as part of the task of making sense of the speaker's behaviour. The theory of meaning for Davidson is part of the theory of rational behaviour. In *Thought and Talk* Davidson explains that he conceives of the theory of rational behaviour as a theory implicit in common sense psychological reasoning. Such a theory provides teleological explanations of why agents act the way they do, they explain why it is *reasonable* for agents to act that way.²⁵ Explanations of behaviour appeal to beliefs and desires to account for the behaviour of an agent. The evidence (behaviour) is observable, and is explained as the prod-

²³Cf. RI, p. 125.

²⁴Lepore and Ludwig (2005), pp. 74-76. To be more precise, the question is rather: which truth theories (*plural*) fit the evidence we have concerning the speaker. Due to the indeterminacy of interpretation (see below), there is no unique theory. Cf. IR, pp. 239-240.

²⁵Cf. TT, pp. 158-161.

uct of two forces not directly observable: beliefs and desires.²⁶ Linguistic behaviour is a specific part of behaviour in general, namely the behaviour consisting (in the most common case) of the intentional movements of the lips and the larynx.²⁷ A meaning theory cannot be used directly as an explanation of speech behaviour, but is rather a prerequisite for a teleological explanation of speech behaviour (and a teleological explanation of speech behaviour again is a prerequisite for a teleological explanation of behaviour in general, since, as we shall see in chapter 3, teleological explanations only apply to Davidson to creatures with speech). To return to the example of Kurt and his utterance about the rain: a theory of interpretation for Kurt does not explain *why* Kurt utters (by intentional movements of his lips and larynx) ‘Es regnet’, but it does explain something that is needed to make sense of Kurt’s intentionally moving his lips and larynx in the way he did. A theory of interpretation reveals how the disturbances of the air caused by Kurt’s intentionally moving his lips and larynx are meaningful. To understand what Kurt wants to achieve, we first have to understand the meaning of his utterance. However, knowing the meaning of the utterance does not tell us anything yet about the *aim* of Kurt’s utterance (i.e. Kurt’s extralinguistic purpose): this is to be explained in terms of (i.e. as being a product of) Kurt’s desires and beliefs.²⁸ Once we know the meaning of the utterance, we are in a position to explain why the agent made the specific utterance by means of a teleological theory, in terms of his beliefs and desires (provided we have sufficient information about Kurt’s beliefs and desires to do so, the availability of which again partly depends on having interpreted many of his other utterances).²⁹ Davidson conceives of a theory of meaning as redescribing actions (the utterance of sounds) in a revealing way, i.e. in such a way that a teleological theory (in terms of beliefs and desires) can explain the interpreted speech behaviour.

It is clear from the above considerations that for Davidson meaning and use of language are strictly separated: an utterance with one meaning can be employed for achieving (or trying to achieve) several different purposes. Davidson calls this idea the autonomy of meaning and he considers it an essential feature of language.³⁰ Another important point following from the above considerations is that Davidson conceives of the theory of (radical) interpretation as part of a more general theory of rational behaviour.

For Davidson psychological and intentional notions like belief, desire,

²⁶Cf. TT, p. 160.

²⁷Cf. RI, p. 126.

²⁸He can convey a message, give a command, ask a question, etc. Cf. MP.

²⁹For Davidson we can only ascribe beliefs and desires to an agent if the agent has a language, cf. 3.3.1. Only very simple beliefs concerning the surrounding environment may be identified prior to interpretation of the speaker’s language. Therefore exactly such beliefs form the starting point of radical interpretation, but only on the assumption that the creature entertaining those beliefs has a language.

³⁰Cf. TT, p. 164.

and meaning cannot be reduced to behaviour, but he does hold that they supervene on behaviour. By this he means that these psychological states can only be ascribed in as far as they explain behaviour (an important part of which is speech behaviour).³¹ This external approach to meaning has important consequences for the theory of meaning. Davidson holds that all there is to know about meaning must in principle be publicly accessible. All that a speaker means must be reconstructible on the basis of observable evidence that is not semantical in nature. Such evidence may not be observable in the context of particular utterances, but it must be available in the speaker's history and/or it must be possible for the speaker to make it available to an interpreter.

Davidson holds that intentional notions like meaning can and should be fully understood from a third person perspective: the role they play in an explanation of intentional behaviour; and in case of meaning specifically of social behaviour (communication). Internal applications of these notions (like an individual's entertaining beliefs without ever expressing them) is secondary for Davidson. Davidson has various reasons for this external approach, which will come to light in the following chapters. Specifically in case of meaning, we shall call this idea *interpretationism*.³² Interpretationism therefore is the idea that the principles governing interpretation are constitutive of meaning as such.

This is not to mean that Davidson is a behaviourist or even a sceptic regarding other minds. On the contrary: he is not at all, as we shall see. In fact, interpreting a creature as speaking a language for Davidson implies that the creature is not merely reacting to its environment in a programmed way, but has a mind, and that therefore its behaviour is open to explanation in terms of reasons.

How a theory of interpretation can be constructed on the basis of observable evidence is explored by reflecting on the position of the radical interpreter. The starting position of the radical interpreter is one of complete ignorance regarding the meaning of a speaker's words. The radical interpreter tries to find out what the speaker's words mean merely on the basis of observable evidence. Davidson takes as fundamental evidence the speaker's attitude of holding true certain sentences in her language (relativized to time) combined with the environmental conditions obtaining when the speaker holds true those sentences. Davidson assumes that this evidence is observable prior to knowing the meaning of those sentences (or, at least, can be reconstructed from what is observable without knowing the meaning of those sentences). We shall further consider this claim in chapter 2.

Davidson thinks that philosophers have long failed to see what the consequences are of the idea that all evidence as to what a speaker means must

³¹Cf. TT, p. 159.

³²Cf. Glüer (2007), p. 278.

be publicly available.³³ For Davidson, Quine was the first to systematically explore the consequences of this idea, most notably by means of his thought experiment of radical translation in chapter 2 of *Word and Object*.³⁴ Davidson throughout his work frequently looks at meaning from the perspective of the radical interpreter. He emphasized aspects different from the ones Quine was most interested in. For Quine the most important consequence following from considerations on radical translation is the large scale indeterminacy of translation. Davidson focusses on another aspect: the need for large scale agreement between interpreter and speaker. Davidson holds that radical interpretation leaves less room for indeterminacy than radical translation, partly because the radical interpreter imposes a logical structure on the speaker's language by reading quantificational theory into it. This is necessary since the speaker must construct a truth theory for the speaker's language.

For Davidson thinking about meaning from the perspective of the radical interpreter can illuminate the nature of meaning.³⁵ Radical interpretation, however, *as Davidson conceives of it*, is an imaginative process with constraints on the evidence that are imposed a priori. One might think that Davidson would argue that since radical interpretation in some form *must* be possible (for example, because field linguists succeed to interpret previously uninterpreted languages, which amounts to successful radical interpretation), it must be possible to base an interpretive theory on publicly available evidence. However, as is very clear in *Radical Interpretation Interpreted*, Davidson does not claim that it follows from the a priori assumption that radical interpretation is possible that the evidence is publicly accessible. It is rather the other way around: assuming that the evidence is publicly accessible, he wants to argue that radical interpretation, *as he specifies it*, is possible (and he does not assume that it is possible in that way). Davidson therefore grants that he may be wrong about how radical interpretation is possible: a field linguistic may use different (or more) evidence. Apparently contrary to this, Davidson does seem to base important conclusions on his considerations on the stance of the radical interpreter (for example the need for agreement between interpreter and speaker), which seem to be unsupported if radical interpretation as specified by Davidson turns out *not* to be possible. It seems, however, that Davidson thinks that in as far as such conclusions follow, they would follow regardless of how radical interpretation is done, which particularly goes for the need for agreement, which for Davidson constrains *all* interpretation.³⁶

In *Radical Interpretation Interpreted* Davidson writes the following on the role of radical interpretation (and the constraints and the assumptions

³³Cf. TP, p. 56.

³⁴Chapter 2 of Quine (1960).

³⁵Cf. RII, pp. 126-127.

³⁶Cf. 3.2.1.

that are part of it, which we shall further elaborate on below) in his philosophical project:³⁷

The story I tell about how radical interpretation is possible should be viewed as an informal proof that the evidence assumed available plus the constraints on the structure of the pattern of a person's beliefs, values and language, suffice to yield an interpretive theory for understanding that person.

It seems therefore that Davidson's aim is the following: Assuming that meaning determining facts are observable and assuming that there is a certain structure within language and between intentional notions (as e.g. described by decision theory), Davidson argues that interpretation on the basis of the evidence described is possible, namely by means of radical interpretation. If Davidson is right about this, then this would arguably make it more plausible that meaning determining facts are indeed publicly available in the way assumed by Davidson, because the facts specified would suffice for interpretation. At least, it would make it more plausible if we assume that the facts specified by Davidson as evidence are indeed publicly available, a matter we shall further investigate in chapter 2.

A crucial aspect of the truth theories constructed by the radical interpreter is the asymmetry between theoretical and empirical primacy. The empirical primacy is on the level of the *T*-sentences entailed by the theory. These sentences are tested by the radical interpreter by testing whether their predictions are corroborated by the speaker's use of language. The crucial evidence that the interpreter relies on when testing hypothetical theories is the speaker's holding true of sentences under specific conditions, at specific times. The subsentential machinery on the other hand, i.e. the reference axioms for names, the satisfaction axioms for predicates, and the recursive axioms for the sentential connectives and quantifiers, is theoretically/explanatory primary but postulated not on the basis of direct evidence, but on the basis of its success in generating *T*-sentences that fit the linguistic behaviour of the speaker. In fact, the theory is not tested by the predictions made by individual *T*-sentences, but by *all* the *T*-sentences together. Due to Davidson's meaning holism, there is no absolute fact of the matter as to the correctness of an individual *T*-sentence: it can only be tested holistically whether a theory is successful in predicting the speaker's linguistic behaviour.³⁸

This brings us to the indeterminacy of interpretation: there are several truth theories that are empirically adequate for a speaker (at some point Davidson suggests to identify a theory with a language, in which case we may say that there are several languages that the speaker could correctly be

³⁷RII, p. 127

³⁸Cf. 4.4.2. Cf. IA on this.

interpreted to be speaking)³⁹. There are two kinds of indeterminacy. The first consists of the fact that if a given theory is empirically adequate, any theory that provides equivalent truth conditions for all sentences is adequate as well. For a given adequate theory there is always another equivalent theory, with a different axiomatization but equivalent T -sentences: given that for Davidson empirical evidence only concerns the sentential level, the adequacy of the axioms is fully defined in terms of the truth conditions they generate; this is the doctrine of the inscrutability of reference.⁴⁰ Davidson thinks of this kind of indeterminacy relatively uninteresting: he compares it to the idea that several scales of measurement can be used to apply to the same measurable phenomenon.⁴¹ A more interesting, second kind of indeterminacy concerns the idea that sometimes adequate truth theories may differ concerning a given sentence of the speaker whether it is true or false at a given time. For Davidson this is a consequence of the entanglement of meaning and belief.⁴² It happens when in some case there is no way to find out whether the speaker has a different belief (particularly a different opinion on what falls under a certain concept) or whether the speaker means something different. In such cases an interpreter can choose either of these, by making ‘compensatory adjustments elsewhere’ in the theory, where depending on which he chooses the speaker is right or wrong.⁴³ In such cases there is no fact of the matter which interpretation is right. This is a direct consequence of the idea that meaning is publicly available in the form of non-semantic facts.⁴⁴

1.3.2 Davidson’s version of Convention T and the empirical constraints on truth theories

Davidson has an interest different from Tarski’s. Whereas Tarski assumes an understanding of the notion of meaning and wants to explore the notion of truth, Davidson takes an understanding of truth for granted and wants to give an account of translation and interpretation (and by doing so of meaning).⁴⁵ Davidson therefore does not want to appeal to translation (or sameness of meaning) in Convention T , in order not to assume what needs to be explained. His Convention T (Davidson’s Convention T henceforth) does not require that the sentence used at the right side of the biconditional be a

³⁹Cf. IR, pp. 239-240.

⁴⁰One adequate theory can easily be transformed into another by redefining the reference of proper names and compensating this by redefining the satisfaction definitions of predicates. Cf. IA, pp. 78-80. However, for some problems concerning how to interpret this idea see our footnotes on this in 2.3.2 and 4.4.2.

⁴¹Cf. IA, p. 75.

⁴²Cf. IA, pp. 80-81.

⁴³Cf. IA, pp. 80-81.

⁴⁴Cf. 4.4.2.

⁴⁵Cf. Davidson (2001b), introduction, p. xvi.

translation of the sentence at the left side of the biconditional.⁴⁶ This does of course not mean that Davidson does not think that the *T*-sentences entailed by a theory of truth for a language do (or should) not in fact translate (interpret) the sentences at the left side of the biconditional: Davidson wants to establish that this will be so, *if* the theory of truth satisfies certain formal *and* certain empirical constraints (which we will come back to below) equivalent to a translation requirement.⁴⁷

In Tarski's work, T-sentences are taken to be true because the right branch of the biconditional is assumed to be a translation of the sentence truth conditions for which are being given. But we cannot assume in advance that correct translation can be recognized without pre-empting the point of radical interpretation; in empirical applications, we must abandon the assumption. What I propose is to reverse the direction of explanation: assuming translation, Tarski was able to define truth: the present idea is to take truth as basic and to extract an account of translation or interpretation.

Davidson wants to argue that a truth theory satisfying certain constraints does not merely entail a true *T*-sentences for each sentence in the object language but interprets the object language. For a truth theory to be *interpretive* the sentences used at the right side of the biconditionals entailed by the theory must be translations into the metalanguage of the sentences mentioned at the left side of the biconditional (or, in case the object language is included in the metalanguage: the sentence used at the right side must be the sentence mentioned at the left side). The radical interpreter initially does not know how sentences in the speaker's language are to be translated into sentences of her own language. Davidson conceives of her task as constructing a truth theory for the speaker's language such that it gives true *T*-sentences for *all* the sentences of the speaker's language. He wants to argue then, that if this is so (given certain other conditions, to be discussed below) the *T*-sentences of that theory are interpretive.⁴⁸

There is no difficulty in rephrasing Convention *T* without appeal to the concept of translation: an acceptable theory of truth must entail, for every sentence *s* of the object language, a sentence of the form: *s* is true if and only if *p*, where '*p*' is replaced by any sentence that is true if and only if *s* is. Given this formulation,

⁴⁶Or, in case that object language is part of the metalanguage (which the interpreter need not know in advance), that the sentence used at the right side is the sentence mentioned at the left side. The latter is what Tarski's adequacy criterion in Tarski (1944) requires. Cf. Tarski (1944), pp. 343-345.

⁴⁷RI, p. 134

⁴⁸RI, p. 134

the theory is tested by evidence that T -sentences are simply true; we have given up the idea that we must also tell whether what replaces ‘ p ’ translates s .

Therefore, Davidson’s Convention T only requires that the T -sentences are true. In other words, Davidson’s Convention T only requires that the theory invariably pair truths with truths and falsehoods with falsehoods, which is the only thing needed to make biconditional sentences true.

The absence of the translation requirement is not the only feature that distinguishes Davidson’s Convention T from Tarski’s. Another, closely related difference is a consequence of the fact that Davidson’s truth theories are to be applied to natural rather than to formal languages. Since natural language contains indexicals (e.g. the tense of verbs and demonstratives), truth conditions must be relativized to speaker and time. T -sentences therefore have (something like) the following form:⁴⁹

‘Ich bin müde’ is true as (potentially) spoken by p at t if and only if p is tired at t .

Also given this additional complication, we could still say that Davidson’s Convention T only requires that the theory invariably pairs truths with truths and falsehoods with falsehoods, i.e. it pairs instantiations with a given value for p and t of the form “Ich bin müde’ is true as (potentially) spoken by p at t ’ with the correlating instantiations of the form ‘ p is tired at t ’, where it is required that both instantiations have the same truth value. All Davidson’s Convention T requires is that the for any biconditional the sentence at the left side have the same truth value as the sentence at the right side.⁵⁰

As we will see, the relativization of truth conditions to speaker and time is of great importance to Davidson’s ‘extended project’: the clarification of how an interpreter comes to know what she needs to know to interpret a speaker. The evidence available to the radical interpreter consists of the fact that speakers of the language to be interpreted hold various sentences true at certain times under specified circumstances. Correlations between the speaker’s holding true certain sentences at a certain times and the circumstances in the speaker’s environment allows the interpreter to tie the speaker’s words to objects and events in the environment, and hence to connect the speaker’s words to the world. This works particularly for what (after Quine) are called *observation sentences*, sentences whose truth value

⁴⁹Example adapted from TM, p. 34. The qualification ‘potentially’ is needed to deal with sentences not actually uttered by the speaker p at time t .

⁵⁰So, for example, if “Ich bin müde’ is true as (potentially) spoken by Kurt at noon on 21 September 2014” is a true (false) sentence, then the Convention requires that the following sentence be true (false) as well: ‘Kurt is tired at noon on 21 September 2014’, and vice versa.

vary depending on the context of utterance according to a systematicity that can be discovered without knowing their meaning.⁵¹ Notwithstanding the importance, in *Truth and Meaning* Davidson addresses the indexical features of natural language as (apparently) a ‘fly in the ointment’ for his project. The reason is that it is a controversial issue whether formal semantics can deal with such features (or rather, if it cannot, formal semantics is impossible). This is presumably only possible if the indexical features are finitely specifiable. Davidson implicitly assumes this is so, because if meaning is so local as to not allow any generalizations over specific contexts of use, it would be impossible to have a semantic theory in the form desired by Davidson.⁵²

The question that now arises is whether the formal requirement that Davidson put on truth theories suffices for them to be meaning theories, i.e. interpretive theories. Tarski’s Convention *T* required (or implied) that each *T*-sentence be interpretive. How can we be sure that the same goes for theories satisfying Davidson’s Convention *T*? Davidson writes on this:⁵³

What we have been overlooking, however, is that we have supplied an alternative criterion: this criterion is that the totality of T-sentences should (in the sense described above) optimally fit the evidence about sentences held true by native speakers. The present idea is that what Tarski assumed outright for each T-sentence can be indirectly elicited by a holistic constraint. If that constraint is adequate, each T-sentence will in fact yield an acceptable interpretation.

Davidson’s hope is that if we require that the truth theory *as a whole* optimally fit the evidence, it will be interpretive. What this requirement exactly amounts to depends on what ‘optimally fitting the evidence’ means, i.e. on how the radical interpreter takes the available evidence to support a theory. As part of an answer, we first have to go into the methodology of the radical interpreter, as sketched by Davidson in *Radical Interpretation*.

The evidence consists of the speaker’s holding true sentences, relative to time, combined with the environmental conditions obtaining when the speaker holds true those sentences.

The radical interpreter gathers evidence in support of a truth theory for the speaker. On the basis of patterns among hold true attitudes towards sentences the interpreter tries to read logical structure into the language (by identifying sentential connectives and quantificational machinery). Assignment of satisfaction conditions starts with the predicates in observation sentences. More theoretically loaded predicates (like ‘bachelor’) are given

⁵¹Cf. RI, p. 136 and CTTK, p. 149.

⁵²Cf. RI, p. 33 and p. 35. Cf. 4.2.1.

⁵³RI, p. 139

satisfaction conditions on the basis of how the sentences containing them are evidentially supported by sentences already interpreted and/or by observable conditions (constrained by the need to find quantificational structure).⁵⁴

In Radical Interpretation Davidson describes the evidence for the interpretation of observation sentences as follows:⁵⁵

- (E) Kurt belongs to the German speech community and Kurt holds true ‘Es regnet’ on Saturday at noon and it is raining near Kurt on Saturday at noon.

Instances of this type of evidence support the following kind of generalizations:

- (GE) $(x)(t)$ (if x belongs to the German speech community then (x holds true ‘Es regnet’ at t if and only if it is raining near x at t)).

Such generalizations again support the following related T -sentence:

- (T) ‘Es regnet’ is true-in-German when spoken by x at time t if and only if it is raining near x at t .

In the above description Davidson takes for granted that the interpreter has already identified a group of speakers as belonging to the same linguistic community. To establish this itself requires interpretation: it requires that the interpreter verify that the same utterance type (sentence) uttered by different speakers has the same meaning (abstracting away from the influence of indexical features). Verifying this in a sense is ipso facto interpreting the speakers. This is one of the reasons that in later works Davidson comes to adopt a more individualistic approach to meaning, based on the idea that conventions (obeyed by a linguistic community, or by an individual through time) are neither necessary nor sufficient for successful linguistic communication, but are only a matter of convenience, making interpretation easier (that is normally obeyed to quite an extent by speakers).⁵⁶

In going from (GE) to (T) a crucial step is taken: *held true* is replaced by *is true*. In other words: the interpreter assumes that the sentences are mostly true when they are held true. This assumption is part of the principle of charity. This principle is Davidson’s answer to a challenge that the interpreter faces: the interdependence of meaning and belief. In order to find out what the speaker means by ‘Es regnet’ (or any other of the sentences he holds true) the interpreter needs to know what belief the speaker expresses by means of the sentence (we assume here that the speaker expresses a belief by means of the sentence (which is presupposed by the nature of the evidence: holding true), but even utterance is not an assertion, finding out

⁵⁴Cf. RI, p. 136 and TP, pp. 65-75.

⁵⁵Cf. RI, pp. 135-136.

⁵⁶Cf. NDE.

its meaning depends on finding out what the speaker believes the utterance means). But as long as the speaker's language is yet uninterpreted, the interpreter has no access to the speaker's beliefs by means of communication. Davidson thinks, however, that the interpreter can identify some of the speaker's beliefs without having interpreted the speaker's language yet. These are exactly the beliefs expressed by sentences like 'Es regnet', that concern phenomena identifiable by observation, without knowledge of semantic facts (which Quine called 'observation sentences'). However, determining the truth conditions of such sentences on the basis of environmental conditions inevitably means that the interpreter has an opinion as to whether the speaker's sentences are true or false. In fact, there seems to be no option in order for interpretation to get off the ground but to assume that the speaker is mostly right. That is what is done by taking (GE) to support (T): the interpreter (implicitly) assumes that if speakers (admitting of individual (explicable) exceptions) *hold true* a certain sentence *if and only if* a certain condition obtains, that then that sentence *is* true if and only if that condition obtains. This is what the *principle of charity* requires the interpreter to do.

What we have not been paying explicit attention to yet is that in applying charity the radical interpreter assumes that (T) is interpretive, i.e. that 'es regnet' is not just true if and only if it is raining, but that it *means* that it is raining: the interpreter assumes that 'es regnet' expresses the belief that it is raining. But on what evidence is this assumption based? The obvious answer seems to be that it is *because* it is raining (and because of nothing else) that the speakers of German hold true 'Es regnet'. This, however, is not required by Davidson's Convention *T*, so it seems that we cannot be sure that Davidson's Convention *T* guarantees that a truth theory is interpretive. This brings us back to the holistic constraint cited above.

What Davidson writes is ambiguous between two readings. On a weaker reading, it merely requires that the radical interpreter find a theory such that *all* *T*-sentences are true, in which case it is not a genuinely additional constraint, since it already follows from Convention *T* that this is required. On the weaker reading, the fact that all sentences are true guarantees/implies that they are interpretive. On a stronger reading something is required in addition: not only must all *T*-sentences be true, but they must satisfy additional requirements in order to 'optimally fit the evidence'. Whichever Davidson intended when he wrote this down, clearly the stronger reading is needed, as can easily be shown by counterexamples to the sufficiency of the weaker reading, which we shall address in chapter 2.⁵⁷ Indeed, in a footnote that was added to *Truth and Meaning* in 1982 Davidson suggests that the

⁵⁷We shall discuss one counterexample in chapter 2. In fact, what we suggest there is a third reading: the very way in which a radical interpreter finds out that a theory is *true* ('optimally fits the evidence'), guarantees that it is interpretive (and this excludes some formally but not empirically correct theories). Cf. 2.3.

following closely related empirical constraints should also be satisfied:⁵⁸

- The sentences of the theory must not only be true but also lawlike and supported by appropriate counterfactuals.
- The evidence for accepting the (time and speaker relativized) truth conditions for [sentences like] ‘That is snow’ is based on the causal connection between a speaker’s assent to the sentence and the demonstrative presentation of snow.

What exactly ‘appropriate counterfactuals’ are will not concern us here, but we shall assume that Davidson means that the sentence used at the right side of the biconditionals specifies what it would be that would *cause* the sentence mentioned at the left side to be true, in case it were true: the *T*-sentence should be true because of *causal* connections. The fact that the two sentences of the biconditional have the same truth value should therefore not be arbitrary: there should be a causal connection.⁵⁹

Davidson’s idea seems to be that satisfying the second requirement with regard to the evidence for the meaning of observation sentences guarantees that the biconditional are lawlike (elsewhere Davidson writes that they are natural laws).⁶⁰ In other words, in *T*-sentences based on such causal connections we can substitute ‘because’ for ‘if and only if’. The idea implicit in formulating this constraint (as a constraint necessary for a theory to be interpretive) is that what causes the speaker to assent to a sentence is constitutive of the meaning of that sentence. We shall come back to this important idea below.⁶¹

In general, Davidson’s hope is that we can formulate a set of constraints (part of which are causal constraints) such that if the interpreter satisfies all these constraints and Convention *T*, she will end up having an interpretive theory for the speaker.

1.3.3 Breaking into the interdependence of meaning and belief: the principle of charity

In order to get radical interpretation off the ground, it is needed to dissolve—in as far as possible—the entanglement of meaning and belief. This is where Davidson thinks a principle is needed to guide the interpreter, the so-called

⁵⁸TM, p. 26, footnote 11. Cf. Davidson (2001b), introduction, p. xvi.

⁵⁹If appropriate counterfactual conditions are conditions under which the meaning remains the same, there must be a way to identify these conditions without presupposing semantic facts about the speaker’s language (in as far as this is not itself a semantic fact), otherwise the constraints on radical interpretation could not be satisfied. Cf. Pagin (2013), section 4.

⁶⁰Cf. Davidson (2001b), introduction, p. xvi and p. xx.

⁶¹Cf. 1.4.

principle of charity. The principle of charity tells the interpreter—roughly—to fix the speaker’s beliefs and to infer something about what the speaker means. The principle of charity tells the interpreter to interpret the speaker in such a way that what the speaker says makes sense and is mostly right. This formulation is vague, which is deliberately so, because Davidson has given various formulations of the principle, none of which are very precise.

For example, in *Radical Interpretation* Davidson writes:⁶² ‘This [to solve the problem of the interdependence of belief and meaning] is accomplished by assigning truth conditions to alien sentences that make native speakers right when plausibly possible, according, of course, to our own view of what is right.’

The principle of charity foremost concerns reading logical structure into the speaker’s sentences, in order to find the appropriate recursive axioms (for sentential connectives and quantificational devices) for a theory of truth. This is done by identifying sentences (always) held true or false (possible logical truths or falsehoods) and patterns of inference among them.⁶³ For such evidence to suffice for identifying logical structure, the interpreter must assume that the speaker is rational (from her own perspective, of course, but there is no other) to the extent that the speaker in normal cases obeys the laws of (the interpreter’s) logic. The idea that the interpreter could find the speaker using a fundamentally different logic (or being fundamentally illogical) is inconceivable: the identification of the speaker’s logic (and with that of the logical structure of his language) requires the assumption that the speaker uses the interpreter’s logic. As long as the interpreter has not found logic in the use of, for example, what she thinks is a quantifier in the speaker’s language, she cannot formulate an axiom for it.⁶⁴

Identifying a truth theoretical structure in the speaker’s language is one part of the job of the interpreter. Davidson sometimes suggests this can be done prior to assigning truth conditions to any of the speaker’s sentences, i.e. prior to formulating the non-recursive axioms (which specify the reference of proper names and the satisfaction conditions of predicates).⁶⁵ Also in assigning truth conditions to sentences the principle of charity is needed. In *A Coherence Theory of Truth and Knowledge* Davidson writes: ‘Something like charity operates in the interpretation of those sentences whose causes of assent come and go with time and place: when the interpreter finds a sentence of the speaker the speaker assents to regularly under conditions the interpreter recognizes, the interpreter takes those conditions to be the

⁶²RI, p. 137

⁶³Cf. RI, p. 136.

⁶⁴Cf. RI, p. 136. In fact, Quine only appealed to the principle of charity in order to identify logical connectives, cf. RI, p. 136, footnote 16.

⁶⁵Cf. RI, p. 136 and TP, p. 73. Cf. 2.2 for more discussion, particularly on whether this also concerns the identification of moods prior to interpreting any sentence.

truth conditions of the speaker's sentence.'⁶⁶

The two aspects of the principle (structure and content) are distinguished in following formulation from Three Varieties of Knowledge:⁶⁷

The process of separating meaning and opinion invokes two key principles which must be applicable if a speaker is interpretable: the Principle of Coherence and the Principle of Correspondence. 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.

The Principle of Coherence enables the identification of structure, as discussed above.

The Principle of Correspondence urges the interpreter to take the speaker to be responding to the same features as the interpreter would be responsive to. It may seem that the Principle of Correspondence overlaps (or even coincides) with the causality constraint: isn't the *cause* of a speaker's assent the same as the *feature of the world that the speaker responds to* (in which case the causality constraint is equivalent to the Principle of Correspondence)?

Things are not this easy, however. As we suggested above when we first mentioned charity, charity also constrains the interpretation of observation sentences. There are two main ways in which charity plays a role.

First, the interpreter assumes that what typically causes a speaker to hold true a sentence is what the sentence actually is about, i.e. she makes the speaker right. We shall see that Davidson thinks that this is a natural assumption: what causes a speaker to hold true a sentence—given certain provisos—is, in many cases, what makes the sentence true.

As for the second way, the Principle of Correspondence points to some other problem that occurs when basing interpretation on causes, namely the question *what* the cause of the hold true attitude is. Many things may be taken to have caused a speaker to hold true 'Gavagai': light rays striking the speaker's retina projecting a rabbit-image, the presence of the rabbit in the environment, or the rabbit's birth.⁶⁸ All of these (and many more) are part of the causal chain leading to the speaker's holding true the sentence. The first is presumably a prerequisite to cause a hold true attitude (and the

⁶⁶In RI, p. 137 and CTTK, p. 150 respectively.

⁶⁷TVK, p. 211

⁶⁸Cf. EE, pp. 201-202. Cf. 1.4 and 3.3.1.

presence of a rabbit-like retinal image does not require the actual presence of a rabbit), but it is not *natural* to say that the speaker held true the sentence *because* a rabbit-image was projected onto his retina. This is what is taken care of by the Principle of Correspondence: we should take the speaker to be responding to the ‘features of the world’ that *we* would also respond to ‘under similar circumstances’, and such features—under the given circumstances—are rabbits scurrying by, not the births of those rabbits or the projection of their images onto our retinas.

Davidson was fascinated by the problem of locating the cause of a hold true attitude, and his thoughts on this problem (in particular his rejection of Quine’s idea that sensory stimuli must be the evidence for the theory of interpretation) gave rise to his triangulation theory, which we shall discuss in the third chapter.⁶⁹

The qualification that the interpreter make the speaker right *when plausibly possible* in the first citation we give is important: especially when identifying the truth conditions of observation sentences, the interpreter must take into account that the speaker may sometimes be mistaken. The interpreter must allow for explicable error. In detecting what a speaker is talking about various considerations play a role, and making the speaker reasonable (taking into account explicable error) is more important than making her right as often as possible, although these two are closely related for Davidson. For example, if the interpreter observes that a hare scurrying by causes the speaker to hold true ‘Gavagai’ in quite some cases, it may be reasonable to assume that the speaker is mistaken, assuming that there is evidence that normally rabbits cause the speaker to hold true that sentence and that the speaker is potentially able to see the difference. (It is also easy to construct examples showing that the observational situation of the speaker must be taken into account to allow for error.) Therefore, making speaker’s behaviour rational and explicable is more important than making the speaker right as often as possible, although Davidson is convinced that the latter is a consequence of the former.⁷⁰

The principle of charity (together with the other constraints) for Davidson are principles that are constitutive of meaning. In this sense the name ‘charity’ is misleading, in as far as it suggests a charitable assumption: in order to be a speaker, for Davidson, it must be the case that we are interpretable (and that we interpret) according to the principle of charity: it is a principle of rationality. Nevertheless, commentators disagree on the justification of charity and whether it is actually forced on the interpreter, as Davidson suggests. We shall further discuss this issue in the third chapter.⁷¹

A consequence of the need for charity is that an interpretive truth theory

⁶⁹Cf. 1.4 and 3.3.1.

⁷⁰Cf. 3.3.5. Cf. TT, p. 169. Cf. RLQ, pp. 282-283.

⁷¹See 3.2.2.

for the speaker makes most of the speaker's beliefs true in the eyes of the interpreter: speaker and interpreter disagree on the world only in details. In general, according to what we shall call Davidson's *agreement thesis* all communication is constrained by agreement, which will be the topic of further discussion in the third chapter as well.⁷²

1.4 Causes, triangulation, and veridicality

One of Davidson's central ideas is that what typically causes us to believe something is what that belief is about. In the simplest cases words and thoughts refer to what causes them.⁷³ More precisely, this means that 'the situations which normally cause a belief determine the conditions in which it is true'.⁷⁴ This is so in particular for *perceptual* beliefs, beliefs about what we presently observe (i.e. beliefs expressed by observation sentences like 'I see water' and 'that is water').⁷⁵ Entertaining such beliefs is restricted to the occasion of observation. It is crucial that those causes are *typical*: what causes a belief in any particular case need, of course, not determine the conditions under which it is true; what matter is what has *typically* caused instances of the same type of belief in past cases. We can therefore have the *same* (perceptual) belief at several times and places (expressible by the same sentence), and on any particular occasion what causes the belief need not be what makes the belief true, but what *typically* causes the belief (or instances of it) determines under which condition the belief is true. Davidson writes:⁷⁶

We must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief.

The 'basic cases' are primarily cases of beliefs engendered by observation. Of course, in many cases beliefs prompted by observation are not *directly* caused by what they are about. For example, someone's saying something may typically cause someone else to entertain a particular belief. Also apart from such linguistic causes, perceptual beliefs are often not typically directly caused by what they are about. An example that Davidson gives is that a streak in the cloud chamber may typically cause us to hold true 'there goes an electron'.⁷⁷ In such cases a reason can be given why the perceived caused us to entertain a certain belief. The same dependency on information (that can be made public by the believer) pertains to beliefs that are directly

⁷²Cf. 3.2.1.

⁷³Cf. EE, pp. 196-197 and TR, p. 34 (in the latter Davidson uses the qualification 'typical').

⁷⁴Cf. EE, pp. 196-197.

⁷⁵Cf. EE, p. 199 and RTN, p. 208.

⁷⁶CTTK, p. 151

⁷⁷Cf. STL, pp. 137.

caused by what they are about for a speaker, but the meaning of which cannot (easily) be identified by an interpreter in terms of typical causes without prior knowledge of the speaker's beliefs (like the belief expressed by the speaker's sentence 'A bachelor is walking over there'): that such beliefs are engendered by observation is to be explained in terms of beliefs on the speaker's side that cannot be observed directly as in simple cases like the speaker's holding true the sentence 'it is raining'.⁷⁸

For Davidson all beliefs, also those easily identifiable in terms of typical causes for an interpreter, are informed by their relationships to other beliefs, and hence depend on theory. Understanding beliefs cannot be reduced to typically being caused to entertain them by what they are about: understanding rests on knowing how the belief is connected to many other beliefs that inform its meaning. This is the reason that Davidson rejects Quine's idea that observational beliefs and theoretical beliefs can be clearly separated on epistemological grounds. The only distinction that does make sense is the distinction between beliefs entertained because of what is observed and those entertained independently of what is currently observed.⁷⁹

A consequence of this causal theory is that what an individual means or believes does not supervene on the physical constitution of the individual: meaning is partly determined by what is external to the individual, particularly by how the individual has been in touch with his environment in the past. This means for Davidson that two physically indistinguishable individuals, producing sounds of the same type, may have to be interpreted differently, depending on (the history of) the circumstances external to them.⁸⁰

For Davidson the connection between beliefs and reality is causal.⁸¹ Davidson rejects the idea of epistemic intermediaries in between our beliefs and what they are about: something that our beliefs are *based on*, as empiricist have tried to find. There are merely causal intermediaries. Perceptual beliefs are caused spontaneously.⁸² Therefore, a belief can only be supported by another belief. Perception constitutes evidence for beliefs by engendering perceptual beliefs.⁸³

Davidson has further elaborated his causal externalism by his idea of triangulation, which adds a social element to it. Davidson uses the metaphor of triangulation for the idea that two minds are needed to locate the objects of belief in an objective world. Davidson's triangulation theory is supposed

⁷⁸Cf. STL, pp. 136-138.

⁷⁹Cf. CTTK, p. 149. Cf. STL, pp. 136-138.

⁸⁰Cf. 3.4.1. Cf. IA, pp. 70-73: this implies that only a token-token reduction of mental events or states to physical events or states (in our bodies) is possible, not a type-type reduction. Our thoughts partly derive their content from factors external to our body.

⁸¹Cf. CTTK, pp. 143-147.

⁸²Cf. STL, p. 136.

⁸³Cf. STL, pp. 136-138.

to provide an answer to the question what the relevant causes are that give content to our beliefs and what counts as the *typical* cause. Davidson argues that it is only when two (or more) persons interpret each other that there can be an answer to the question what their words are about: in that case their holding true sentences (and the beliefs expressed by those sentences) has (have) a *common cause*. It is only by communicating with others, by ‘*triangulating*’, that there can be a *common* cause and hence an objective criterion as to what our words mean. A private language (in the sense of a language not interpreted by someone other than the speaker) is not possible. In fact, Davidson holds that thought depends on language, and that therefore having beliefs is not possible without having a language that is interpreted by others. We shall discuss Davidson’s views on this in the third chapter.⁸⁴

For Davidson the causal theory gives rise to the veridicality thesis: the thesis that belief is by nature veridical. According to the veridicality thesis our beliefs are mostly true, since we are mostly (directly or indirectly) caused to have those beliefs only under the conditions under which they are in fact true. Or, expressed in terms of sentences: ‘the conditions that typically cause us to hold sentences true *constitute* the truth conditions, and hence the meanings, of our sentences.’ (Davidson’s emphasis)⁸⁵

In particular, the causal theory concerns the extension of the concepts we use to classify things in the world. This is so because in basic cases many of our (perceptual) beliefs are of the form ‘That is an F’ where an F is an observable thing or property. The extension of F is therefore determined by the objects we typically apply it to.⁸⁶ We cannot be fundamentally wrong about how we subsume the objects in our environment under concepts (and hence the conception we have of our environment), since the content of those concepts is determined by how we are typically caused to apply them. Davidson writes on this:⁸⁷

[I]t cannot happen that most of our plainest beliefs about what exists in the world are false. The reason is that we do not first form concepts and then discover what they apply to; rather, in the basic cases the application determines the content of the concept. An interpreter who starts from scratch—who does not already understand the language of a speaker—cannot independently discover what an agent’s beliefs are about, and

⁸⁴Cf. 3.3.1.

⁸⁵FTDT, p. 34.

⁸⁶Cf. DCT, p. 75. Given the inscrutability of reference, there is only a fact of the matter as to the extension of a speaker’s predicates relative to an interpretive theory for the speaker’s language, but that does not threaten this idea for Davidson: the inscrutability of reference does not make ontology relative (there is just no fact of the matter how to employ ontological categories to give the correct axioms). Cf. IR.

⁸⁷EE, pp. 196-197

then ask whether they are true. This is because situations which normally cause a belief determine the conditions under which it is true.

We shall further discuss these ideas in the third chapter. In the third chapter we shall also try to show how Davidson's ideas on triangulation are connected to his ideas on radical interpretation, and how Davidson's ideas concerning these two central notions if combined reinforce the veridicality thesis.

1.5 The concept of truth

We have seen that for Davidson the concept of truth is a crucial to the theory of meaning. Davidson does not analyse the concept of truth, and he is convinced that it is a more transparent concept than the concept of meaning.⁸⁸ Radical interpretation relies on an intuitive understanding of the concept of truth, which particularly features in the interpreter's intuitions as to what the truth conditions of certain sentences are and when they are true or false. Even though Davidson does not analyse truth (he wants to take it as a basic concept) he does write some interesting things on the nature of truth and on philosophical theories of truth.

Davidson's fundamental idea concerning truth, which he tries to defend throughout a number of publications, is twofold. On the one hand, truth has nothing to do with epistemology, in the sense that what is true does not depend on our epistemological powers. For example, it does not depend on what we are (in principle) able to provide evidence for. On the other hand, Davidson rejects the idea that all our empirical beliefs could be false together, or that we could be fundamentally mistaken about reality. The latter would be a consequence of metaphysical realism: the idea that our beliefs could be exactly what they are, and the world completely different. Davidson tries to formulate a position in between epistemic (or anti-realist) theories of truth and metaphysical realism: Davidson wants a theory that makes our belief veridical but makes truth verification-transcendent.⁸⁹

In earlier work, particularly in *True to the Facts* and *A Coherence Theory of Truth and Knowledge*, Davidson tended to endorse the intuition of correspondence theory: 'truth is correspondence with the way things are'.⁹⁰

Davidson however rejects the idea that it is possible to specify (in a non-trivial way) what a true sentence corresponds with, as fact-based theories have it: the correspondence intuition therefore has little explanatory value. Saying that a belief corresponds to the facts is just a variant on saying that

⁸⁸Cf. Davidson (2001b), introduction, p. xvi and TM, pp. 22-24.

⁸⁹Cf. TP, chapter 2 and cf. 4.2.

⁹⁰CTTK, p. 139

it is true. Davidson rejected fact-based correspondence theory already in *Truth and Meaning* (1967) and *True to the Facts* (1969). Davidson presents an argument that he thinks shows that if true sentences correspond to something, all true sentences correspond to the same thing, the so-called slingshot argument.⁹¹ He thinks this result can be established easily if one is willing to accept the following two assumptions: a) Logically equivalent sentences have the same reference.⁹² b) A sentence does not change its reference if a contained singular term is replaced by another with the same reference.⁹³

Davidson concludes that this argument shows that sentences do not refer (or correspond) to anything, since their reference cannot be individuated (and is explanatory vacuous).⁹⁴ Based on this argument (and, in fact, another argument to the same conclusion)⁹⁵, Davidson rejects the idea that truth can be explained as correspondence to the facts (in as far as this is not trivially understood as just being true). As a consequence of this, Davidson also rejects the idea that sentences represent anything.⁹⁶

⁹¹Cf. TM, p. 19. Davidson uses the argument in *Truth and Meaning* to argue that it makes no sense to identify the meaning of a sentence with its supposed reference, because then all sentences alike in truth value would have the same meaning.

⁹²Where it is assumed that a definite description is a singular term. Cf. following footnote.

⁹³The argument goes as follows (we follow closely Davidson's argument in *Truth and Meaning* here, adding some explanations where we think it is appropriate). Suppose that R and S abbreviate any two sentences alike in truth value. Then the following four sentences have the same reference, where $\hat{x}(x = x)$ stands for 'the unique x such that $x = x$ ':

- (1) R
- (2) $\hat{x}(x = x.R) = \hat{x}(x = x)$
- (3) $\hat{x}(x = x.S) = \hat{x}(x = x)$
- (4) S

Explanation: (1) and (2) are logically equivalent. (3) and (4) are also logically equivalent. By assumption (a) then, (1) and (2) have the same reference, as do (3) and (4). Further, (3) differs from (2) only in containing the singular term $\hat{x}(x = x.S)$ instead of $\hat{x}(x = x.R)$. Since $\hat{x}(x = x.R)$ and $\hat{x}(x = x.S)$ share their reference if R and S are alike in truth value, we have by assumption (b) that $\hat{x}(x = x.R) = \hat{x}(x = x)$ and $\hat{x}(x = x.S) = \hat{x}(x = x)$ have the same reference. Combining the findings we have that (1) and (4) have the same reference. Davidson suggests in TP (pp. 127-130) that the argument as given above is vulnerable to some objections (like its treatment of definite descriptions as singular terms, which goes against Russell's analysis), and he refers to Neale's work (Neale 2001) for a sophisticated discussion of the argument that Davidson thinks definitely settles the case against facts. For further considerations against facts, cf. 4.3.4.

⁹⁴Davidson therefore rejects Frege's idea, consistent with the slingshot argument, that all true sentences refer to The True.

⁹⁵The other argument is simply that facts are not needed to explain truth, cf. 4.3.4.

⁹⁶Cf. TP, p. 41. A cause for Davidson is not a fact, but an event, whence the situations that cause us to hold true sentences (and that are constitutive of their truth conditions) are not facts. Cf. CR and EE pp. 201-202.

However, the early Davidson (e.g. in *True to the Facts*) does not reject the basic idea of correspondence theory as such. To formulate an alternative, non-fact-based correspondence theory, Davidson appeals to Tarski's truth theories. In *True to the Facts* Davidson argues that Tarski has given a respectable notion of correspondence, albeit it not an intuitive one. Tarski has shown how for a sufficiently complex language, a language with quantifiers, formulating a truth theory for that language demands the introduction of an ontology.⁹⁷ According to Tarski's theories, a true sentence is satisfied by all sequences of objects (all assignments of objects to variables). The concept of satisfaction ties sentences to the world. There is no individuation of the world-relata here: all true sentences 'correspond to'/are satisfied by all sequences of objects.

Davidson argues that Tarski gave a material adequacy criterion for a theory of true-in-language (Convention *T*) that corresponds with our intuitive understanding of truth—according to which the truth of a sentence depends only on its meaning and on the way the world is—and then provided a theory entailing all and only all the desired sentences:⁹⁸

The semantic concept of truth as developed by Tarski deserves to be called a correspondence theory because of the part played by the concept of satisfaction; for clearly what has been done is that the property of being true has been explained, and non-trivially, in terms of a relations between language and something else.

If the truth of our beliefs depends on an independent world, the question arises why not all our beliefs may be false. We cannot step outside the realm of beliefs to verify whether our beliefs are true or false. In *A Coherence Theory of Truth and Knowledge* Davidson suggests an answer to this by formulating what he conceived of as a combination of a coherence theory of truth and a correspondence theory of truth. He argued in this essay that belief is by nature veridical, and that the more our beliefs cohere with one another, the more reason we have to suppose that they 'correspond to reality', i.e. that they are true: 'coherence yields correspondence'.⁹⁹ Coherence here consists of the coherence of beliefs in justificatory relationships.

Davidson soon regretted his formulations in *A Coherence Theory of Truth and Knowledge*. He regretted the term 'coherence theory' because of its associations with philosophers that consider reality to be a construct of thought, an idea that Davidson does not want to be associated with.¹⁰⁰

⁹⁷Cf. 1.2.2.

⁹⁸TF, p. 48

⁹⁹Cf. CTTK, p. 137.

¹⁰⁰Another problem is that coherence is to an extent intrinsic to belief, whence the idea is pleonastic: cf. AF, p. 155.

Davidson also regretted his use of the term ‘correspondence’. He decided that the term is explanatory vacuous (for reason we have already discussed) and he renounces the whole idea of correspondence theory henceforth. The only (but important) point of A Coherence Theory of Truth and Knowledge, therefore, is the thesis that belief is veridical, which is crucial to Davidson’s ‘middle position’. The veridicality thesis and Davidson’s ‘middle position’ will be discussed in the third and fourth chapter respectively.

Davidson also came to regret having called Tarski’s theory a correspondence theory. Tarski does not analyse truth, but rather explains the structure truth makes within a language.¹⁰¹ In fact, Davidson argues that we should give up trying to explain what truth consists of.¹⁰² Davidson thinks that Tarski’s truth theories can be used to formalize what is the most important application of the concept of truth: its explanatory role in the theory of interpretation. Based on considerations on the empirical application of truth theories Davidson tries to formulate his alternative view on truth, which we shall discuss in detail in chapter 4.

¹⁰¹Cf. AF, pp. 154-155. Another problem with the idea that satisfaction is of importance to an understanding of truth is that satisfaction seems to have no clear direct application to natural languages, given that natural languages have no clear equivalent of formulas with unbound variables.

¹⁰²Particularly in The Folly of Trying to Define Truth. We shall discuss this in 4.3.

Chapter 2

Radical interpretation: evidence and constraints

2.1 Introduction to chapter 2

In this chapter we are concerned with two aspects of radical interpretation: the nature of the evidence that radical interpretation is based on and the formal and empirical constraints that truth theories constructed on the basis of such evidence need to satisfy in order to be interpretive.

In the first part we shall address Davidson's conception of the primitive evidence that the radical interpreter relies on: the speaker's *holding true* sentences (under certain conditions). (2.2.1) We shall discuss Davidson's claim that this evidence is available to the radical interpreter, and we shall consider some complications that play a role in identifying hold true attitudes and in taking the hold true attitudes to be essential to the understanding of language: the fact that mostly speakers are not concerned with truth, the fact that most utterances are not indicative, Davidson's claim that other than indicative moods can be understood in terms of indicative sentences, and his claim that even if the speaker does not (sincerely) assert something, the interpreter may infer that the speaker holds true certain sentences. (2.2.2)

In the second part we shall consider some constraints that need to be put on truth theories for them to be interpretive. First we shall explain, based on an example given by Lepore and Ludwig, that it is possible for a truth theory to satisfy Davidson's extensional Convention *T* without being interpretive, whence additional constraints are needed. (2.3.1) The causal constraint seems to provide what is missing in the counterexample to the idea that Convention *T* suffices. For the causal constraint to govern interpretation an even more fundamental constraint (or presupposition) is needed: a shared sensitivity between speaker and interpreter as to what is salient and similar. (2.3.2)

2.2 Hold true attitudes as primitive evidence

2.2.1 Primitive evidence: hold true attitudes

It is a constraint on radical interpretation that an interpretive theory must be based on non-semantic evidence, specifically on observable (behavioural) evidence, or on evidence that can ultimately be obtained from the observation of behaviour (and the interaction between speaker and environment and speaker and interpreter).

Quine took the relevant behavioural evidence for the radical translator to consist of *prompted assent*: assent to sentences prompted by specified types of sensory stimuli.¹ Davidson takes the primary evidence of radical interpretation to be an *attitude*: the attitude of *holding true* a sentence, relativized to time. Mostly, this evidence is manifested as sincere assertion of the sentence, but Davidson suggests that hold true attitudes towards a sentence can sometimes be inferred from the utterance of another sentence (we shall come back to this below).

This means that the evidence is not purely behaviouristic: it is of a psychological nature. The evidence consists of an *attitude* that the speaker has towards a sentence, particularly towards a sentence uttered by the speaker, namely that she holds it true.² Quine's evidence (prompted assent) is of a more behavioural nature. Assent need not be sincere, whereas holding true a sentence (or rather, sincere assertion) is: Davidson assumes that sincerity can be found out prior to knowing the speaker's language. He suggests that in case of his evidence a further reduction to behavioural evidence may be possible, but that is clearly not his main concern.³ As opposed to Quine, Davidson rejects the idea that intentional notions like holding true a sentence (which is a belief) can be reduced to non-intentional vocabulary, so in as far as a further behavioural reduction is possible it will not eliminate the intentional vocabulary: a description of behaviour as intentional inherently involves such vocabulary for Davidson (Davidson would presumably say the same of Quine's notion of assent).⁴

By taking the hold true attitude as evidence Davidson seems to meet the most important constraint of the project: not to assume any prior knowledge of the meaning of the speaker's words (i.e. any knowledge of semantic facts). The evidence involves psychological facts, but no semantic facts.⁵

¹Cf. chapter II of Quine (1960) (paragraph 7).

²Although Davidson suggests that hold true attitudes towards related sentences can sometimes be inferred in case of utterances other than assertions, see below. If the evidence is restricted to asserted sentences, it amounts to the identification of sincere assertion (and maybe hold true attitudes that can be inferred from sincere assertions on the basis of logic).

³Cf. BBM, p. 142.

⁴Cf. BBM, p. 154. Cf. TT, p. 162.

⁵It may be objected that *sentence* is a semantic notion (and maybe the same goes for *holding true*), but the evidence does not presuppose any semantic facts specific to the

On the other hand, taking the attitude of holding true as basic evidence may seem to conflict with another of Davidson's tenets: the idea that beliefs cannot be identified independently of what a speaker means, due to the interdependence of meaning and belief. Thinking that a sentence is true, however, is a belief. Davidson thinks of holding true as a special kind of belief, that can be identified prior to having interpreted the speaker's language, because it is non-individuative. The attitude can be described without having to appeal to the meaning of a sentence and can be manifested towards various sentences. Holding true a sentence does not imply anything as to the meaning of the sentence.⁶

Unfortunately, Davidson rarely reflects in more than a superficial way on how the primitive evidence of radical interpretation is available to the radical interpreter. It seems that Davidson holds that identifying a hold true attitude is not much harder than identifying an utterance of a sentence as being intentional and linguistic in the first place. The possibility of radical interpretation as described by Davidson therefore relies on the assumption that the interpreter can find out that the speaker holds true certain sentences prior to knowing what those sentences mean.

2.2.2 The identification of hold true attitudes

Jonathan Bennett, in a review of *Inquiries into Truth and Interpretation* that is otherwise also extremely critical towards most aspect of Davidson's philosophy of language, especially criticizes Davidson for his explanation of the evidence of radical interpretation.⁷ Bennett imputes philosophical negligence to Davidson by taking the complex notion 'intending to express a truth' as an 'unanalysed, unexamined lump'. Bennett writes that it is implausible that at a stage where we do not yet know the meaning of a speaker's sentences we can distinguish attitudes like believing *s* to be true, believing *s* to be plausible, believing *s* to be probable, believing *s* to be desirable, and so on.

In fact, Davidson does suggest that we should take into account these more fine-grained distinctions of attitudes towards sentences. Davidson argues that in a more developed theory probability and desire play a role from the start. In later works he takes the speaker's preference that one sentence be true rather than another as basic. In particular he argues that degrees of holding true (probabilities) need to be taken into account when identifying the meaning of theoretical predicates, since speaker's do not have binary attitudes towards sentences containing such predicates.⁸

speaker's language, except for the individuation of its sentences (or rather: utterances of sentences).

⁶Cf. Glüer, p. 278. Cf. RI, p. 135.

⁷Bennett (1985), p. 611 and p. 617-618.

⁸Although Davidson does so in work published only after Bennett's review. Cf. e.g.

Nevertheless, Davidson is convinced (also while elaborating a more refined theory) that an attitude as to whether a sentence is true or false is the key to understanding for the radical interpreter, and it may be this assumption that Bennett thinks is questionable. Intuitively, most of a speaker's utterances are not concerned with truth, i.e. they are not (sincere) assertions. Moreover, many sentences cannot even be held true: imperatives, interrogatives, etc. have no truth value. (Davidson writes in *Moods and Performances* that non-indicative sentences have no truth value.)⁹

Davidson willingly admits that we are normally not concerned with truth when communicating, but he thinks that truth plays a crucial role in understanding the *meaning* of linguistic expressions (as opposed to their use). Davidson argues that we can only understand an indicative sentence if we understand under what conditions it is true. This is what we have to learn when we learn a language. However, truth-conditional understanding is not restricted to indicative sentences. According to Davidson's account of the logical form of moods other than the indicative mood truth conditions play an essential role in the understanding of all kinds of sentences: the understanding of all other kinds of sentences (interrogatives, imperatives, etc.) presupposes the understanding of related indicatives. For example, interrogatives can only be understood by understanding the indicative sentences that they imply as possible answers (if uttered as a question), or imperatives by understanding under what conditions what they command is obeyed (if uttered as a command).¹⁰

This reasoning inspires Davidson's idea on how a truth theory is able to deal with moods other than the indicative. Utterances of sentences of a non-indicative mood actually consist of two utterances of indicative sentences: one of them is a 'mood setter' (which is true if and only if the other sentence has the mood specified by it) and the other is the related indicative (in the sense suggested above).¹¹ The details of this account need not concern us here, what matters is the idea that for Davidson the meaning of all sentences can be analysed in terms of indicative sentences.

Davidson thinks that this primacy of indicative sentences in the theory of meaning carries over to a primacy of the (related) hold true attitude in the use of language. Davidson suggests in *Thought and Talk* that from many kinds of utterances (other than sincere assertions) we can infer that the speaker holds true certain sentences. For example, if the speaker utters a command, we can infer that he thinks that some related (indicative) sentence is false that he wants to be made true by someone's obeying the command.¹²

TP, p. 65.

⁹Cf. MP, p. 121.

¹⁰Cf. TP, p. 124.

¹¹For open questions there must be more than two related indicative sentences. Davidson does not consider this.

¹²Cf. TT, pp. 161-162. Cf. RI, p. 135.

However, it is not obvious why the primacy of truth conditions in the understanding of all linguistic expressions implies that the ‘correlating’ attitude (holding true) should be the primary evidence for the radical interpreter as to what the speaker means.

One difficulty is in Davidson’s idea that the utterance of other than indicative sentences implies that the speaker holds true (or false) certain related sentences. We grant that this is so, but how could the radical interpreter identify these related sentences at a point where she has no idea yet as to the meaning of any of the sentences? This is only possible if the interpreter already knows the syntax of the sentences, particularly the syntax of the moods.

In Radical Interpretation Davidson indeed suggests that logical structure can be read into the language prior to identifying the meaning of any sentence: the interpreter can identify sentences always held true and study patterns of inferences among those sentences.¹³ This presumably also involves identifying what moods sentences have.¹⁴ In that case the interpreter could for example recognize a sentence as an imperative or an interrogative without yet knowing its meaning. One may think that in that case, if the speaker utters an interrogative sentence, the interpreter can infer from the logical form of the interrogative sentence that the speaker (if he is sincere) is ignorant as to the truth value of related indicative sentences (even though the interpreter knows the meaning of neither of these two sentences).¹⁵

One problem with this idea is that the mood of the sentence is independent of the *force* of its utterance (i.e. asserting, commanding, questioning, etc.). What is formally an interrogative can be used not to ask something, but to assert or imply something (e.g. a rhetorical question). Therefore, inferences can only be made on the basis of additional assumptions. Nevertheless, as long as there is a conventional association between mood and force this need not be a problem for the identification of hold true attitudes towards related sentences.

A more important problem is the circularity in this idea: identifying logical form *also* depends on the recognition of hold true attitudes, so the recognition of the hold true attitude (either towards the sentence uttered or towards a related sentence) cannot be based without circularity on a prior recognition of logical form, including the mood of sentences. Therefore, if interpretation starts by identifying logical form (without yet identifying the truth conditions of any sentence), this must involve classifying sentences into moods and giving a truth conditional analysis of non-indicative sentences.

Would it be possible—theoretically—to identify the compositional structure of *all* sentences without having interpreted any sentence yet? Is it—for

¹³Cf. RI, p. 136. Cf. Lepore and Ludwig (2005), p. 175-178. Cf. TP, pp. 67-75.

¹⁴Cf. Lepore and Ludwig (2005), pp. 178, footnote 152.

¹⁵In case the speaker is a teacher, we assume that educational questioning is not sincere.

example—possible to identify what structure is required to build interrogatives without having interpreted any question yet? In order to do so (it seems) the interpreter at least needs to recognize when the speaker is asking a question (without knowing its meaning yet) in order to identify the structure of the conventionally related interrogative mood.

However, maybe it is more plausible that interpretation starts with utterances of simple sentences (like ‘Gavagai’) and that the logical structure is identified in concert with interpretation of predicates. When the interpreter has sufficient insight into the structure of the sentences, he is able to draw inferences and identify sentences related to the ones uttered that are held true or held false.

The question as to what is possible here may be partly empirical. We shall therefore not pursue it further here.

2.3 From evidence to interpretation

2.3.1 Holistic constraints

Supposing that the interpreter can identify hold true attitudes, how should she proceed to interpret the speaker? What constraints must be observed? We shall see that here the other component of the evidence of the radical interpreter starts playing a role: the environmental conditions, and the way the speaker responds to those conditions.

It has been widely discussed in the literature whether truth theories can function as theories of meaning.¹⁶ Davidson holds that, if a truth theory for the speaker’s language satisfies Convention *T* (more precisely, Davidson’s version of it) and certain empirical constraints, we can replace the ‘if and only if’ in certain entailed *T*-sentences (to be qualified) by ‘means that’ and take them to specify the meaning of the speaker’s sentences.¹⁷

Already in *Truth and Meaning* Davidson anticipates some difficulties

¹⁶For a discussion of the most influential criticisms of Davidson’s truth-conditional approach, see Lepore and Ludwig (2005), pp. 101-118. Pagin (2013), section 4, also provides a clear discussion of what conditions truth theories must satisfy in order to be interpretative.

¹⁷Cf. RF, p. 175. A qualification is needed because not all *T*-sentences *entailed by an interpretive theory* are interpretive. For example, sentences of the form “Schnee ist weiß’ is true in German iff snow is white and *q*’, where *q* is an arbitrary logical tautology, will be provable from any interpretive theory for German (supposing we use classical logic). Obviously such biconditionals are not interpretive. Somehow such nonsensical *T*-sentences must be filtered out. Lepore and Ludwig show how this can be done by defining what constitutes a *canonical proof*, a proof that, roughly, sticks to the terminology in the axioms. Only canonically proven *T*-sentences of an interpretive truth theory are interpretive. Cf. Lepore and Ludwig (2005), pp. 19-21. If an interpreter has found an adequate truth theory for a speaker from which by a canonical procedure the theorem “Schnee ist weiß’ is true if and only if snow is white’ can be derived (and she knows this), she can be said to know that ‘Schnee ist weiß’ *means* that snow is white.

that arise when using a truth theory as a theory of meaning. He observes there that a T -sentence like “Snow is white’ is true if and only if grass is green’ satisfies Davidson’s Convention T (it is a true biconditional), whence we cannot say that a theory entailing it is formally wrong on the basis of this sentence.¹⁸ However, Davidson argues that it is implausible that a theory that entails “Snow is white’ is true if and only if grass is green’ can be such that it satisfies his Convention T for *all* sentences it entails. Davidson thinks it is implausible that a theory that entails “Snow is white’ is true if and only if grass is green’ will also provide correct truth conditions for demonstrative sentences like ‘That is snow.’. Such demonstrative sentences will force the interpreter to interpret the predicate ‘snow’ as being satisfied by snow and not by e.g. grass, which will lead to the exclusion of the undesirable sentence mentioned: a theory entailing the undesirable sentence would give a wrong T -sentence for the sentence ‘that is snow’.¹⁹ In the paper *In Defence of Convention T* Davidson writes that it is exactly because truth conditions need to be relativized to time and speaker that they become closer to being interpretive: ‘The point now is that a theory that makes the right sentences true at the right times and for the right speakers will be much closer to a theory that interprets the sentences correctly than one that can ignore the extra parameters.’²⁰ Davidson’s idea is therefore that the assignment of extensionally correct truth conditions to sentences whose truth value is contingent on the time of utterance and the speaker will guarantee that oddities like “Snow is white’ is true if and only if grass is green’ are not entailed by a theory satisfying his Convention T .

This suggests that Davidson here appeals to the weaker reading of the holistic constraint that we discussed in chapter 1: as long as *all* T -sentences are true, the theory is interpretive. However, it can be shown easily (and Davidson soon realized) that other constraints are needed, as we shall now proceed to demonstrate.²¹

2.3.2 Causal constraints

Let us first consider an example that Lepore and Ludwig give to show that Davidson’s Convention T does not suffice to guarantee that a theory is interpretive.²²

Let LT be a logical truth. Assume that T_1 is an interpretive theory for the language of the speaker. Further assume that T_1 has the following satisfaction axiom for the predicate ‘snore’:

¹⁸Cf. TM, pp. 25-26.

¹⁹Namely something like “That is snow’ is true as spoken by x at t if and only if the substance demonstrated by x at t is grass’.

²⁰RI, pp. 74-75

²¹Cf. Davidson (2001b), introduction, p. xvi.

²²Cf. Lepore and Ludwig (2005), pp. 105-107.

[T_1] ‘snores’ is true of d if and only if d snores.

Now assume that T_2 is the same as T_1 , except that it has a different satisfaction axiom for the predicate ‘snore’:

[T_2] ‘snores’ is true of d if and only if d snores and LT.

Obviously, T_2 is *not* interpretive, since it will for example pair the speaker’s sentence ‘Peter snores’ with ‘Peter snores and LT’.²³ However, it is easy to see that if T_1 satisfies Davidson’s Convention T , T_2 also satisfies it, since the conjunct ‘LT’ in the satisfaction axiom has no influence on the truth value of the T -sentences.²⁴

Now what constraint could eliminate theories like T_2 ? Presumably something like the causal constraints mentioned in chapter 1 needs to be invoked here. These constraints are not clearly formulated by Davidson, but Davidson’s idea seems to be that the interpreter must have evidence as to what typically causes the speaker to hold true a certain sentence, and that the truth conditions must specify this cause. Clearly, it is not (an awareness of the truth of) LT that causes the speaker to hold true a sentence saying that someone does or does not snore, whence such a constraint can rule out T_2 .²⁵

²³We do not think that Davidson would disagree that T_2 is not interpretive by appealing to the inscrutability of reference (even if the difference between the two theories is in reference). The inscrutability of reference, as an empirical phenomenon in interpretation, arises because there may be different ways to assign satisfaction conditions to predicates and to specify the reference of names, all of which generate *empirically* (not merely extensionally) adequate theories. Cf. Davidson’s examples and discussion in IA, pp. 79-81. There is, however, a problem here: the inscrutability of reference is a mathematical fact once we take (Davidson’s) Convention T (extensional adequacy) to be the *only* constraint on truth theories. It is, however, not clear why reference remains inscrutable once we adopt *empirical* constraints (which Davidson suggests are needed next to his Convention T), as we saw (e.g. Davidson thinks the causal constraint discussed is needed). This seems to require an independent argument (unless Davidson does not think empirical constraints are needed, but he seems to think they are needed). Cf. Lepore and Ludwig, chapter 21 for a discussion of this point. Cf. our discussion of this in a footnote to 4.2.2.

²⁴Mind that T_2 can easily be constructed in such a way that the T -sentences described can be proven canonically, by just incorporating the truism in the axioms of the theory (logically any theory that entails the first sentence also entails the second sentence, what matters is canonical entailment.).

²⁵Cf. Pagin (2013), section 4: Pagin argues that causal constraints do not suffice to rule out theories like the one discussed here, but that something like the Principle of Correspondence is needed (p. 240). Lepore and Ludwig point out that theories like T_2 may be thought to be ruled out by their failure to give adequate truth conditions for sentences with intensional contexts (in which adding a conjunct in the *that*-clause may change the truth value), in which case (Davidson’s) Convention T would suffice to weed out the theory. They argue, however, that this solution is not open to Davidson because of his paratactic (non-intensional) treatment of such contexts. (Cf. 4.2.1.) We think this is not the right reason for thinking that the solution is not open to Davidson: in the paratactic account (cf. ST) the truth value of the ‘left side’ sentence (the one referring to the *that*-clause) may change because of adding LT to the ‘right side’ sentence, and

In practice, the interpreter will of course not consider a T -sentence like the odd one (canonically) entailed by T_2 . The reason is that the very way in which the interpreter finds out which T -sentences are true is (according to Davidson) by considering causal connections between the world and the speaker's holding true sentences. The interpreter will therefore satisfy the causal constraint (and presumably any other empirical constraint) in the very attempt to match sentences she holds true with sentences the speaker holds true. The aim of the philosophical project of radical interpretation is therefore to formulate precisely how an interpreter finds out which truth theory is interpretive for a speaker's language. Given Davidson's interpretationism, finding out this is finding out the principles that are constitutive of meaning. The relevance (and the possibility of success) of this project depends on Davidson's assumption that knowing the truth conditions of sentences is essential to understanding the language.

Whether Davidson's Convention T and the causal constraint are jointly sufficient to guarantee that theories are interpretive will not concern us here.²⁶ This discussion would lead us too far afield. We are most interested in the causal constraint, since this is obviously related to Davidson's causal theory of content: if the causes of a belief are constitutive of its content, it is to be expected that they constrain interpretation.

The interpreter must identify the causes in accordance with the Principle of Correspondence. Otherwise she could not identify them uniquely, since, as we discussed in chapter 1, there may be many things that can be taken as cause of the speaker's hold true attitude. She must identify what the speaker takes to be the relevant cause. This requires that she shares a sensitivity to what is salient with the speaker. She must also share an idea of what is similar with the speaker, in order to be able to identify the classes of things the speaker is talking about.

We shall see that the idea of a triangle of a world and two speakers interpreting each other on the basis of hold true attitudes and a shared sensitivity to certain features of the environment is a recurring theme in Davidson's elaboration of the three topics that we shall be concerned with in the following three chapters: his conception of belief (and particularly his thesis that belief is veridical), his conception of truth, and his attitude towards the sceptic.

in that sense wrong predictions are given by T_2 . A better reason why it does not work, also suggested by Lepore and Ludwig, is that the success of Davidson's Convention T would depend on something contingent: whether the language has intensional contexts and whether they are such that they rule out unwanted theories. Cf. Lepore and Ludwig (2005), pp. 105-107.

²⁶This is presumably not so: think of "This is water" is true if and only if this is H_2O : this T -sentence is extensionally and causally correct. Cf. Lepore and Ludwig, p. 232.

Chapter 3

The veridicality of belief

3.1 Introduction to chapter 3

In this chapter we shall focus on Davidson's *veridicality thesis*, the thesis that belief is by nature veridical. According to this thesis our beliefs are mostly true, even though each particular belief may be false. The thesis implies that there are significant constraints on what is conceivable to be believed. In this chapter we shall discuss Davidson's argument for the veridicality thesis.

We shall discuss the veridicality thesis in the context of two important concepts in Davidson's philosophy: radical interpretation (3.2) and triangulation. (3.3) In the context of his work on radical interpretation Davidson explicitly gives an argument that belief is veridical. A weaker thesis that plays a role in this argument is the *agreement thesis*, the idea that successful communication cannot be established without the speaker's and the interpreter's largely having the same beliefs. (3.2.1) The agreement thesis again is a consequence of the idea that interpretation is governed by the principle of charity. (3.2.2) Davidson uses the so-called omniscient interpreter argument to argue that belief is veridical. This argument is controversial, and Davidson rejected it in favour of a simpler way of arguing for the thesis, but it is nevertheless insightful. (3.2.3) The veridicality thesis is reinforced by Davidson's 'triangulation theory', which is supposed to explain why someone's having beliefs depends on there being a world and other minds. (3.3) We shall show how the triangulation theory combines with Davidson's ideas on radical interpretation, despite apparent tensions. (3.4)

3.2 Interpretation and veridicality

3.2.1 The agreement thesis

Davidson holds that communication is only possible if the communicating agents largely share their beliefs: 'I can interpreter your words correctly only

by interpreting so as to put us largely in agreement'.¹ The agreement thesis is used as a premiss in one of Davidson's arguments for the veridicality thesis: the omniscient interpreter argument. The agreement thesis, stated in other words, means that if we ascribe beliefs to someone, we ascribe beliefs that are largely true according to our own standards. We cannot find someone extremely mistaken about the world. Disagreement is possible only against the background of massive agreement. Disagreement therefore is rather the exception than the rule.

The agreement thesis is closely connected to the principle of charity. Charity leads to agreement, both in matters of logic and in matters of the world (the latter both regarding 'analytic/conceptual truths' and 'synthetic truths', although this distinction is vague for Davidson, who follows Quine in this).² If communication has been established through radical interpretation (by applying the principle of charity) this entails that the communicators massively agree. This follows from the fact that the principle of charity leads to the assignment of truth conditions to the speaker's sentences that—in the most cases—obtain when the speaker holds those sentences true. Since the interpreter has no choice but to rely on her own beliefs about what the conditions are under which the speaker holds true sentences, the speaker's sentences will be made true by the interpreter's standards.

If the principle of charity is what the agreement thesis is based on, this suggests that the scope of the agreement thesis should be restricted to understanding that is based on the application of the principle of charity. However, since charity (it seems) is an interpretive principle for *radical* interpretation specifically, a generalization is needed to *all* interpretation to be able to say that *all* speakers largely have the same beliefs. Davidson indeed sometimes writes that all interpretation involves radical interpretation (and hence the application of the principle of charity).³ This assumption is even stronger than what is needed for the agreement thesis to constrain all possible communication: it would suffice to argue that each speaker can in principle be correctly interpreted on the basis of radical interpretation (and hence by applying charity) by another speaker. The problem with this point of view is, however, that Davidson explicitly writes in a later work that he does not *assume* that natural languages are radically interpretable (or at least not in the way he specifies, i.e. based on merely publicly accessible evidence (hold true attitudes), by constructing a justified theory).⁴ Davidson suggests there that he does not have an a priori argument that radical interpretation is possible, but rather wants to argue that this is plausibly so. Davidson writes

¹Cf. MTM, p. 200. Cf. CTTK, p. 150. For what 'largely' means here, cf. our discussion below on the analogous context of the veridicality thesis (3.2.3).

²Cf. CTTK, p. 145.

³Cf. RI, p. 125: 'All understanding of the speech of another involves radical interpretation.'

⁴Davidson writes this in RII. Cf. 1.3.1.

that he thinks of it the other way around: charity is the right principle (for *all* interpretation), and *therefore* radical interpretation (based on the evidence as described) is possible.⁵ It seems therefore that we should take the principle of charity as a more fundamental assumption than the assumption that languages are radically interpretable. In other words: even if radical interpretation is not possible in the way specified by Davidson (for example because hold true attitudes cannot be identified prior to knowing meaning in the way Davidson seems to assume), all interpretation is constrained by charity.⁶

3.2.2 The principle of charity and its justification

As we discussed above, the principle of charity is a principle of how to use the basic evidence for radical interpretation (the speaker's holding true of sentences, and the conditions under which the speaker holds true those sentences) as evidential support for *T*-sentences, and therefore as evidential support for a truth theory. Charity tells us how generalizations over hold true attitudes can be taken to support the *T*-sentences entailed by a truth theory. Charity tells us, roughly, to take the conditions under which speakers hold true sentences to be the truth conditions of those sentences, provided that the speaker is also found consistent to a considerable extent (which is needed to identify logical connectives and quantificational machinery in his language). Charity is the assumption that the speaker—when plausible, taking into account what the speaker can be right about, given what he is able to observe—is not fundamentally mistaken, neither about logic nor about the world. Logically, this is not the only possible principle to treat generalizations over hold true attitudes as evidence for the *T*-sentences entailed by a truth theory, even if it seems the simplest principle. Davidson compares the entanglement of meaning and belief to the entanglement of desire and probability in decision theory, but there is a difference: in decision theory there is an a priori (unique) solution to dissolve this entanglement (on the assumption that the speaker is rational), whereas there is no such an obvious solution for the dissolution of the entanglement of meaning and belief.⁷

The question has arisen in the secondary literature whether and how it can be justified that the principle of charity is the right principle to bridge the gap between evidence and theory. (Apart from discussion on what exactly the principle is, since it is not clearly formulated, but any formulation seems to imply that communication entails a strong form of agreement.)⁸ Davidson's himself does not explicitly address the question of how the prin-

⁵Cf. RII, p. 122.

⁶Cf. Lepore and Ludwig (2005), chapter 11 for a critical discussion of this point.

⁷Cf. BBM, pp. 145-148.

⁸Cf. Lepore and Ludwig (2005), pp. 182-196.

ciple of charity is to be justified. An easy answer would be that the principle of charity is the *simplest* way to connect hold true attitudes and truth conditions. Davidson himself does not seem to think of charity as *just* the simplest way (even though theoretically not the only) of connecting evidence to theory.

Davidson suggests that the correctness of charity follows from a kind of transcendental argument (although Davidson does not call it that way): it would then be a necessary condition for the possibility of a phenomenon whose existence we are certain of, which is interpretation (or communication): ‘Charity is forced on us; whether we like it or not, if we want to understand others, we must count them right in most matters. If we can produce a theory that reconciles charity and the formal conditions for a theory, we have done all that could be done to ensure communication. Nothing more is possible, and nothing more is needed.’⁹ Davidson argues that applying charity (and satisfying the formal constraints on radical interpretation) is necessary for communication to succeed. (Thinking back of our discussion in the previous chapters, Davidson here seems to think of charity as embodying *all* empirical constraints.) For such an argument to succeed he should explain why *only* charity makes interpretation possible.

Davidson argues that belief ascription (or disagreement on particular beliefs between speakers) is not possible without (implicitly) applying charity on a large scale, due to the holism of belief. This would imply that any form of linguistic communication (be it radical or not) cannot exist without charity, at least if we grant that we can only communicate with someone if we can ascribe beliefs to that person (which seems obvious: we have to ascribe beliefs to the speaker as to what his words mean).

Davidson argues that if, for example, John ascribes the belief to Mary that a cloud is passing before the sun, he must implicitly ascribe a range of true beliefs (by his standards) to her concerning the nature of clouds, the sun, etc. (even if he disagrees on the explicitly ascribed belief), which will ensure that he agrees with Mary on many things.¹⁰ Some authors attack this argument by arguing that belief ascription does not require agreement about *particular* beliefs (about particular objects, events, properties, etc. in the environment), but only about conceptual truths concerning the concepts involved in the ascription.¹¹ For example, John could ascribe to Mary the belief that a cloud is passing before the sun without himself believing that the sun is visible at the moment, or that there are clouds in the sky at the moment: he just needs to assume that Mary believes this (otherwise he could not consistently ascribe a belief about a cloud passing before the sun to Mary), but he does not need to agree, as long as their conception of ‘sun’

⁹VICS, p. 197

¹⁰Cf. MTM, p. 200.

¹¹Cf. Lepore and Ludwig, chapter 19 and Pagin (2013), pp. 242.

and ‘cloud’ etc. is shared, and as long as they agree on conceptual truths regarding the meaning of ‘sun’ and ‘cloud’ etc. On empirical matters they may massively disagree. Davidson can defend himself against such criticism by pointing out that he endorses Quine’s rejection of the analytic/synthetic distinction, and that you cannot ascribe the concept of cloud to someone apart from considerations on how that person applies the concept to particular cases in the environment, which determine the content of the ascribed concept. In that case sharing empirical beliefs (and hence beliefs about how the world actually is) would also be required for belief ascription.¹²

This is of course the very idea of radical interpretation: in identifying what the speaker’s predicates mean, the interpreter has to rely on how the speaker applies them in particular cases. Therefore, if John interprets Mary’s sentence ‘a cloud is passing before the sun’, he relies on which objects in his environment Mary applies ‘cloud’ to (or would apply it to) in order to interpret the predicate ‘cloud’, which will ensure that John and Mary largely agree on the environment. This presupposes, of course, that John *radically* interprets Mary’s predicate ‘cloud’, which he need not do if he knows her language beforehand, but for Davidson meaning must in principle (potentially) be determinable on the basis of non-semantic facts.

There is, of course, some circularity in this argument, because—if it is true that belief ascription requires agreement—then assuming the possibility of belief ascription amounts to assuming the need for charity, and the sceptic regarding charity may demand an argument to justify the ascription of beliefs to others. Charity therefore seems a fundamental assumption, almost impossible to justify without circularity. This is presumably the reason that Davidson does not think that it is needed to justify charity on the basis of more basic assumptions: charity is not just the assumption that the speaker is mostly right (that is rather a consequence of charity), but it is the assumption that the speaker’s behaviour can be explained rationally and that the speaker responds to the same features of the environment as the interpreter, in other words: that the speaker is a speaker enough like the interpreter for the interpreter to be able to communicate with her. Charity is therefore for Davidson a *natural* principle, that we always implicitly assume, in all communication.

Therefore, we should not think of the interpreter as first identifying the conditions under which the speaker holds true a sentence and then applying charity in taking those conditions to be the truth conditions: rather, identifying the truth conditions relies on charity, on the idea that what causes the speaker to entertain a belief expressed by that sentence under the given circumstances must be something that would cause the interpreter to entertain a similar belief.

¹²Cf. EE, pp. 194-196.

3.2.3 From agreement to veridicality

Davidson does not merely think that communication entails massive agreement (that interpretation is truth-ascribing), he holds that what is agreed upon is largely true (or, what is believed in general is largely true): ‘beliefs are by nature generally true’.¹³ What exactly is meant by ‘generally true’ (and alike for ‘as much agreement as is needed’ in the agreement thesis) is not made explicit by Davidson.¹⁴ Given that there is no clear upper bound to the number of beliefs we have, terms like ‘as much as is needed’ or ‘generally’ cannot be given a straightforward interpretation. Another way to express the idea would be to say that of an arbitrarily chosen belief, the chance is high that it is true, but that will not make things much clearer if ‘high chance’ is left unspecified. The thesis must rather be understood holistically: our body of beliefs *as a whole* is adequate, and against this background of a body of true beliefs each particular belief may be false. Davidson’s holism is another reason to approach the thesis in this way. In isolation beliefs (and sentences expressing those beliefs) do not have a clear identity: they obtain their identity from the network they are part of. As we shall see in chapter 4, for some utterances (and the beliefs they express) there is not even a fact of the matter as to their truth value except relative to an interpretation of the speaker’s language (and beliefs) as a whole.¹⁵

The veridicality thesis is supposed by Davidson to concern various kinds of beliefs. It embraces beliefs about the world (in the first place perceptual beliefs, but also more theoretical beliefs based on them), but it also concerns beliefs about the contents of minds, both that of one’s own mind and that of other minds.¹⁶ On the other hand, regarding beliefs about the world, the thesis foremost applies to perceptual beliefs. Davidson is not concerned (primarily) with beliefs based on sources or witnesses, beliefs based on induction, etc. (although in a sense Davidson is concerned with the latter: interpretation relies on induction, and we shall see that Davidson thinks that thought depends on the interpretation of others, which is the reason veridicality also pertains to (the content of) other minds for Davidson).¹⁷ Further, he is not concerned with the reliability of our epistemic abilities: the veridicality thesis presupposes for example the reliability of memory, otherwise it would be restricted to occurrent perceptual beliefs and it would not pertain to all our past perceptual beliefs and the beliefs based on them. However, in as far as all such beliefs depend on what can be observed in our

¹³CTTK, p. 153. Davidson defends the idea throughout his work, most prominently CTTK, but for example also in MTM (see below), EE, and TVK.

¹⁴Cf. MTM, p. 201.

¹⁵Cf. 4.4.2.

¹⁶In this chapter we shall concentrate on the first kind of beliefs (belief about the world). In chapter 5 we shall address why the thesis also pertains to beliefs about other minds. (Davidson discusses the latter in particular in TVK.)

¹⁷Cf. SP, p. 11.

environment, the veridicality thesis is important to all kinds of beliefs.

An immediate problem with the thesis that belief is veridical, which Davidson unfortunately does not discuss in the context of his argument for it, is that for the thesis to make sense a kind of consistency constraint on belief must be presupposed. If it were possible for our beliefs to be largely inconsistent, claiming that they are veridical would make no sense. In fact, if beliefs are closed under implication and we assume the principle *ex falso sequitur quodlibet*, we trivially have all true beliefs if we have one pair of contradictory beliefs. But that is of course an absurd idea, that can hardly be used as an argument against the intelligibility of the veridicality thesis, whence one of its premisses must be false. Nevertheless, even if our beliefs are not closed under implication (or *ex falso sequitur quodlibet* does not hold), we do probably have many inconsistent beliefs. If this is so on a large scale, this may still mean that in a trivial way many of our beliefs are true (but also many are false).

For Davidson, however, consistency must be the rule, and inconsistency is only possible against a background of global consistency, otherwise our beliefs could not be logically structured: ‘beliefs are individuated in part by their logical properties; what is not largely consistent with many other beliefs cannot be identified as a belief’.¹⁸ Further, to understand Davidson’s point, we should probably think of the thesis in a more qualitative way: it cannot be so that our conception of the world is fundamentally mistaken. This is particularly so because our perceptual beliefs are veridical, beliefs concerning macroscopic features of our environment.¹⁹

These considerations make it clear that Davidson leaves the understanding of the thesis a somewhat intuitive notion. The best way to understand it therefore may be to consider Davidson’s argument for it, which we shall discuss below. Another way to understand it is in terms of what Davidson thinks it implies for scepticism, which it seems to bear on immediately, and in terms of Davidson’s view on the possibility of knowledge (both of which we shall discuss in chapter 5).

One way Davidson has tried to extend the agreement thesis to the veridicality thesis is by the omniscient interpreter argument.²⁰ The argument has elicited a lot of criticism and Davidson later regretted having invoked a non-existent creature for his argument.²¹ Even though Davidson has regretted the invocation of the omniscient interpreter in his argument for the veridicality thesis, it is fruitful to start our discussion of the veridicality thesis by attending to this argument. It identifies all the premisses that are crucial to it.

The argument relies on the idea that there could be an omniscient in-

¹⁸AF, p. 155

¹⁹Cf. EE, p. 201.

²⁰Cf. MTM, pp. 200-201 and CTTK, pp. 150-151.

²¹Cf. RTG, p. 192.

terpreter. The omniscience particularly consists of the fact that such an interpreter knows all about what does and would cause a speaker to assent to sentences. This interpreter, Davidson argues, when radically interpreting us, could not but find us largely right about the world. Otherwise she would interpret us as being massively mistaken, which is impossible, given the agreement constraint on interpretation. But if an *omniscient* interpreter finds us largely right, we *are* largely right.

It may seem strange that an *omniscient* being would have to interpret the speaker, and does not immediately know what he means. Therefore it is more suitable to suppose that the interpreter is omniscient with regard to the observable data needed to interpret the speaker (the data a radical interpreter uses): the omniscience is limited, not including facts about the content of the speaker's mind and semantic facts about his language.²² However, as we discussed in chapter one, what a speaker means for Davidson can ultimately be found out on the basis of observable data, and therefore there is nothing as to what a speaker means that could not be reconstructed on the basis of all evidence. (So we could also say that an omniscient interpreter (infallibly applying the principles of interpretation) ipso facto knows what a speaker means by knowing all relevant observable data on which interpretation is based.)

In a detailed discussion of the argument, Genova presents it as a *reductio* (following the way it is presented in *The Method of Truth in Metaphysics*). Instead of invoking an omniscient interpreter, he presents it in terms of a veridical interpreter (VI), an interpreter that is supposed to have only true beliefs (but who has no beliefs yet as to what the speaker means), and that is in particular fully informed with regard to the data relevant to interpretation.²³ The argument is supposed to show that a massively false belief system (MF), i.e. a speaker with massively false beliefs, is inconceivable. Simplified, Genova's reconstruction has the following form.²⁴

Suppose there is an (MF) (i.e. a speaker whose beliefs are massively mistaken). Suppose a VI interprets the MF. Then (exhaustively) we have one of the following two. (a) The VI interprets the MF as an MF, which means that she massively disagrees with the MF (since she has only true beliefs). (b) the VI does not interpret the MF as an MF, which means

²²Cf. Genova (1992), p. 179.

²³Genova is interested in whether the argument can have implications for scepticism, and he thinks that assuming that the interpreter *knows* the relevant things begs the question against the sceptic who denies the possibility of knowledge (but who does not deny that all our beliefs may be true). However, it is questionable whether a correct interpretation can be based merely on true beliefs, and that interpretation does not require knowledge. If the argument is therefore taken to be concerned with a sceptic regarding knowledge claims, the sceptic could insist that for the VI to interpret that speaker the VI must *know* an interpretive theory for the speaker, but knowledge (the sceptic holds) is impossible. Cf. Genova (1992), pp. 179-180.

²⁴Cf. Genova (1992), section II, particularly pp. 181-182.

that she interprets it as not being massively mistaken and hence as to a considerable extent agreeing with her own beliefs. We then have that (a) is impossible by the agreement thesis. We further have that (b) is impossible, because if a fully informed interpreter interpreted an MF not as an MF, this means that she agrees to a considerable extent with false beliefs (by the agreement thesis), which contradicts the assumption that she is fully informed. Therefore, an MF cannot exist.

Genova points out that there is an important implicit assumption in this argument: the assumption that each conceivable belief system is interpretable. Without this assumption the argument does not go through because the MF may simply be uninterpretable for the veridical interpreter.²⁵

In fact, if a fallible interpreter radically interprets someone's observation sentences, she is implicitly committed to the assumption that she is mostly omniscient with regard to the data relevant for interpretation (i.e. that her beliefs that matter to interpretation constitute knowledge and are veridical) in quite a strong sense: if she does not assume this, she cannot justifiably believe that she is right in identifying the typical conditions that prompt a speaker to hold true sentences. The only alternative for the interpreter is to merely assume that she mostly shares beliefs about those conditions with the speaker, and not to exclude the possibility that communication is established on the basis of massive sharing of false beliefs. In the remainder of this chapter and in chapter 5 it will be clear that the latter is an untenable idea for Davidson, which means that for Davidson one is committed to assume that one's beliefs are veridical (at least if one is an interpreter).²⁶

Davidson himself later regretted the omniscient interpreter, and he sug-

²⁵Ludwig and Lepore argue that the omniscient interpreter argument fails since we could just as well assume the conceivability of an omni-ignorant interpreter: an interpreter most of whose beliefs are false. They argue that an omni-ignorant interpreter would conclude (by the agreement constraint) that he largely agrees with our beliefs, whence it would follow that our beliefs are largely false. However, it seems an inconsistent idea that an omni-ignorant interpreter could arrive at a correct interpretation, since (by assumption) all (or most of) his beliefs regarding the data relevant to interpretation are false. Therefore we would rather say that the interpretation of such an interpreter is wrong (or rather: it is inconceivable that such an interpreter could interpret us). We do, however, agree that assuming the possibility of successful radical interpretation implicitly means assuming that we are mostly right about our environment, otherwise we could not identify what a speaker is talking about (see below). The very methodology of radical interpretation (like the causal theory of content that it is closely related to) presupposes veridicality. Cf. Lepore and Ludwig (2005), pp. 326-328.

²⁶In fact, if Davidson is right that beliefs are about what typically causes them, the interpreter *must* assume she is largely omniscient with regard to what causes the speaker to hold true sentences, otherwise she could not justifiably ascribe beliefs to the speaker. (She could not, for example, leave open the question whether what she observes *really* causes the speaker to hold true sentences, not excluding the possibility that communication arises on the basis of a systematic shared misconception (between speaker and interpreter) of what is going on in the environment (particularly on the basis of shared mistaken beliefs on what the causes of hold true attitudes are).) Cf. 5.5.

gests that the thesis simply follows from considerations on what it means to have beliefs. Davidson invokes his causal theory of content to argue for this. One just needs to think, Davidson suggests, of the idea that in basic cases causal connections give content to our beliefs.²⁷ (Davidson assumes that it is plausible to think this, although we shall see below that the causal theory is supported by further ideas.) Our perceptual sentences obtain their meaning from what causes us to hold them true: they are typically caused to be held true under certain conditions and these conditions constitute their truth conditions. Holding true sentences does not make them true in any particular case, but when we hold them true is not independent of when they are true: for them to be about anything, they must typically be held true when they are in fact true. This is for Davidson what connects our beliefs to reality. This means that our perceptual sentences and the beliefs they express are typically true, and the same goes for all our other beliefs, in as far as they cohere with (and can be justified in terms of) such beliefs (analogous to how an interpreter would interpret our more theoretical sentences on the basis of perceptual sentences, the principle of charity, and other formal and empirical constraints).

In this sense, then, it seems that the veridicality thesis has little to do with radical interpretation or charity. Rather, the causal origins of our beliefs and the logical, analytical, and inductive relations between them guarantee that they are veridical. It seems then that Davidson wants to make this idea plausible by reflecting on how a radical interpreter goes about interpreting someone, a process in which she identifies meaning on the basis of causal connections and in which charity is forced on her. However, reflecting on interpretation for Davidson is not merely a heuristic means to make us aware of how our beliefs come about: for Davidson the principles that govern interpretation are constitutive of what we mean and hence of what we believe. According to Davidson's triangulation theory, which we shall discuss in the next part of this chapter, there is a fact of the matter as to what cause is constitutive of the contents of our beliefs only in as far as this can be found out potentially by an interpreter in accordance with the principle of charity.

Davidson's argument for veridicality therefore assumes a third person perspective. One may think that assuming that each belief system is interpretable, particularly if it is supposed to convince a sceptic, is question begging: it presupposes what is to be established by an argument, namely that there is a world and that we are connected to it in such a way as to be interpretable. What therefore needs to be established is that it follows from the very fact that we have beliefs that we are interpretable, and that our beliefs *must* be intelligible from a third person perspective. Davidson tries to establish this by his triangulation argument, which we shall now go into.

²⁷Cf. CTTK, p. 153 and EE, pp. 200-201.

3.3 Triangulation, objectivity, and veridicality

3.3.1 Individuation of the contents of thought

In the first chapter we have already briefly discussed Davidson's causal theory of content and we mentioned Davidson's idea that 'triangulation', a process in which two persons ('triangulators') interact with each other and a shared world, fixes what their beliefs are about by establishing a common cause of their beliefs: the objects of thought are individuated socially.²⁸ Davidson's triangulation account is an essential elaboration of his externalism. Below we shall further elaborate on how it relates to Davidson's ideas on radical interpretation. Here we shall analyse Davidson's idea in more detail. Rather than in fully evaluating all the details of Davidson's argument, which essentially is an argument for the interdependence of thought and language and the impossibility of a private language (its adequacy is a highly disputed issue, and Davidson never gave more than a coarse-grained outline of his arguments regarding these topics), we are interested in the consequences that the idea has for the nature of belief.²⁹ We shall, however, go into Davidson's idea that a solitary (someone who never communicated) cannot have thought, since exactly this point is essential for Davidson's position on scepticism (to be discussed in chapter 5). We shall also try to defend the plausibility of this idea.

Davidson's idea of triangulation is essentially an idea of how the contents of our beliefs and the meaning of the sentences expressing those beliefs are individuated. Answering this question, Davidson thinks, will illuminate the nature of thought, since having thought is impossible without having beliefs. Our thoughts can only have content if we have a body of beliefs that give content to the thought, although a thought is not itself a belief.³⁰

3.3.2 No belief without language

Davidson argues that we cannot attribute beliefs to a creature if it does not have a language. The reason is that for such a creature we cannot find

²⁸Triangulation is discussed in RA (p. 105, the idea pops up first here), TVK (This essay concerns the three kinds of knowledge related to triangulation, cf. 5.4.), SP (pp. 117-121), ET (pp. 128-130), and STL (138-141). Davidson withdraws some of the claims made in the passages we discuss here in Davidson (2003), pp. 693-694. However, this changes nothing to our analysis below. The discussion in this section is not restricted to these passages, but concerns Davidson's argument for the interdependency of thought, language, and communication in general, as for example discussed in *Thought and Talk*.

²⁹Cf. Verheggen (2013) for a critical but sympathetic discussion of the triangulation argument.

³⁰Davidson gives the following example to illustrate this: if someone thinks about going to a performance of Beethoven, this is as such not a belief, but she needs to have many true beliefs, e.g. about who Beethoven is, what a performance is, etc. in order to have this thought. Cf. TT, pp. 156-157.

out if it merely reacts to its environment in the way it is programmed by nature to do, or whether its reactions are informed by a conception of its environment. Creatures without language may interact in fine-grained ways with their environment, but for them to have beliefs they must have *concepts* to subsume under the things they observe in their environment. Merely having fine-grained discriminatory dispositions with regard to what is in the environment does not suffice to have concepts. Having concepts means being able to *classify* things in one's environment. To be able to classify, one must be aware that one's classification may be mistaken. Davidson describes this as an awareness that there is a difference between the way the world is and the way one conceives of it. In other words, one must have the concept of objectivity, which involves having the concepts of belief and the concept of truth: one must have the idea of the difference between what one believes and what is true. Only if the concept of objectivity can be ascribed to a creature does it make sense to ascribe mistakes to the creature. Having the concept of objectivity (which involves having the concept of belief and truth) is therefore a prerequisite for having any other concept.³¹

For Davidson the concept of objectivity can only be ascribed to creatures with language. The reason is that for Davidson it only makes sense to ascribe the concept of belief (or any concept) to a creature, and hence to ascribe beliefs to a creature, if a norm can be ascribed to it as to how those concepts must be applied. And this norm, Davidson thinks, must be external to the creature.³² This norm can only exist if there is another creature that is sensitive to the same features of the environment: another creature can provide the norm if it communicates linguistically with the creature and hence ascribes concepts (and therewith norms of application) to it. Absent the norm provided by language, it makes no sense to say that the creature makes a mistake by responding in one way rather than another to its environment, since it is not committed to a standard as to what is right.³³ A language provides an interpersonal standard.

If we ascribe a language to a creature, we ascribe norms to the creature as to how to employ concepts—the concepts we interpret the speaker to be using—and as to how concepts are related. Ascription of concepts must come (to a considerable extent) all at once for Davidson, given the holistic nature of thought: concepts partly derive their meaning from their relation to other concepts.

There is another, related reason for Davidson not to ascribe concepts and beliefs to creatures without language: concepts and beliefs are too fine-grained to be justifiably ascribable to them: there is a vast underdetermination in ways to explain the behaviour of such creatures.³⁴

³¹Cf. STL, p. 138-141 and ET, p. 124.

³²Cf. TVK, p. 212.

³³Cf. TT, p. 170.

³⁴Cf. PO, pp. 12-13.

For example, it seems hard if not impossible to find out (if we want to explain its behaviour normatively) whether a dog (which we suppose lacks a language, or at least one that we can interpret) that barks against a tree thinks of the tree as the same tree the cat went up the last time he chased it, the oldest tree in sight, etc.: on the basis of the dog's behaviour we cannot, it seems, infer such fine-grained distinctions. For Davidson this shows we cannot ascribe any conception of a tree at all to a dog: all conceptions are too fine-grained to be supported by the evidence.³⁵ (Of course, in case we observe a human being looking at a tree we can also not infer from his behaviour under which specific description he thinks of the tree, but we have a reason to think that he thinks of the tree under *a* description, exactly because he has a language and we expect that he could tell us.) Davidson admits that we are nevertheless sometimes tempted to ascribe mistakes to a dog. For example, if we noticed that a cat climbed a tree and we then see a dog barking at another tree very close to it, we tend to think that the dog mistakenly believes the cat climbed that tree. But for the reasons discussed such an explanation is just an anthropomorphism for Davidson: any concept or notion we use in our characterization of the dog's 'mistake' is underdetermined by the evidence: the dog cannot tell us whether we are right or wrong in explaining his behaviour in this or that way.³⁶ As a consequence, Davidson holds the somewhat controversial thesis that (non-human) animals do not have beliefs, nor concepts.³⁷

Ascribing a language to someone is therefore for Davidson the key to a teleological, normative explanation of her behaviour (in terms of beliefs and desires). Ascribing a language starts with ascribing the concept of truth. As we have seen, the evidence on the basis of which this is done is the speaker's *holding true* sentences, and the rich compositional pattern discernible among those hold true attitudes. What is therefore distinctive qua behaviour of human beings is that it is open to radical interpretation, precisely by featuring speech behaviour that manifests possession of the concept of truth: on the basis of this relatively poor behavioural evidence an interpretive theory can be constructed that is extremely rich in both content and structure and that enables normative explanation and hence the 'leap into the intensional realm'.³⁸ But of course, *holding true* itself is an intentional notion, identifying which therefore ipso facto means identifying behaviour as intentional.

³⁵Cf. RA, p. 97.

³⁶Cf. RA, pp. 96-99.

³⁷That such creatures may have a language among themselves is a possibility that Davidson is not interested in. As long as we cannot interpret that language, we have no reason to ascribe beliefs to them. Davidson thinks it is quite a mystery *that* we have thought (it is not immediately clear what the evolutionary advantage is above merely having fine-grained innate discriminatory abilities). Davidson does not want to explain this, but rather what the conditions for having thought are. Cf. PO, p. 7.

³⁸Cf. Verheggen (2013), p. 467.

One may object to Davidson's argument that there is in fact a reason to ascribe beliefs (and desires) to, for example, a dog: a teleological explanation of its behaviour (in terms of beliefs and desires) may be much simpler than a causal explanation.³⁹ (For example, in case of a dog barking against a tree, we may say that the easiest way to explain its behaviour is that it desires to frighten the cat.) However, Davidson insists that this is an explanation by analogy, in a vocabulary not really suited for it: we lack an adequate vocabulary to describe the behaviour of non-human animals, that only has some (but not all) features that must be in place for linguistic communication, and hence for thought (see below).⁴⁰

From the above considerations it will be clear that for Davidson the language required to ascribe to a creature the concept of objectivity is not a private language: a private language does not constitute an external norm. It must be a public language: a language that is interpreted by other creatures.

3.3.3 Triangulation and the common cause

There is another, related problem for Davidson in the idea that a creature without a language (or with merely a private language) could have thought. This problem concerns the identification of what features of the environment such a creature is reacting to (and that his beliefs or utterances could be about). We would have to infer what the creature is reacting to on the basis of its behaviour, but the evidence is underdetermined in two dimensions: both in terms of how distal the cause of the creature's response is and in terms of what its width is. In terms of how distal the stimulus is, the creature could be responding to anything, ranging from the sensory stimuli to objects and events (as we identify them) in its environment (and anything beyond in space and time). In terms of width, it is unclear, even if we could settle how distal the stimulus of the creature's response is, how wide the cause is; in other words, even if we could settle what the beginning of the causal chain is that elicited the creature's response (and hence how far away the object of the creature's response is), the question remains how wide the causal chain is. A further problem is that in case of a solitary creature there is no answer to the question which responses of the creature count as being of the same type, and hence which responses are responses to the same thing for the creature. Davidson thinks that the question as to what the stimulus is (and what counts as a reply of the same type) can only be answered if the creature interacts with another creature, and if a *common cause* of their responses can be identified. He uses the metaphor of triangulation to illustrate this.⁴¹

Davidson conceives of triangulation in the following way. Two (or more)

³⁹Evolutionary phenomena are often also explained in teleological terms.

⁴⁰Cf. RA, pp. 101-102 and ET, p. 128.

⁴¹Cf. RA, p. 105, EE, p. 201, and ET, p. 128.

triangulating creatures observe each other's responses to objects and events (and whatever else there is to ontology) in their environment. Let the triangulating creatures be A and B. Creature A observes a causal connection between the presence of similar objects and events in the environment and similar responses given by creature B when such objects and events are present. Creature B, in turn, is able to correlate similar responses of creature A with the presence of similar objects and events. Further, creature A and B respond to each other's responses. If creature A produces the type of response previously associated by creature B with certain conditions, creature B may react to creature A's response as if those conditions obtain (even if it did not itself observe them) (and vice versa).⁴²

If there were just one creature responding to its environment, Davidson argues, there is no way to settle what it is responding to. Once we have the triangular interaction described above, we can identify where the causal chains from the triangulators' skins into the world intersect, and we can identify a *common cause*, something they are both responding to.⁴³ In the triangular setting there is also an intersubjective criterion as to which responses of a creature are of the same type, and are therefore to be considered as responses to the same type of stimulus: this is determined by the other creature's recognizing them as being of the same type and associating them with conditions in the environment (which could be manifested by the other creature's responding to those responses in the way it would respond if it observed itself that those conditions obtained).

For the kind of triangulation described above to amount to linguistic communication more is needed than mere coordination of responses: propositional content must be conveyed between the triangulators. In fact, in our description we already assumed that this happens by using intentional notions, like 'observe' and 'associate'. Triangulation *without* the communication of content (creatures reacting to stimuli and to each other's reactions to stimuli in a coordinated way) is a process that Davidson thinks occurs widely among animals.⁴⁴ In case of such 'primitive' triangulation, we can at least—to some extent—identify common causes, but the process is not sophisticated enough to ascribe thought or language to the triangulators. For Davidson it is no more than a sophisticated mechanistic process, in which the creatures involved manifest fine-grained dispositions not merely in their reactions to the environment, but also in their reactions to each other's reactions.

Davidson's suggestion is, however, somewhat speculatively, that even though thought and language play no role in such a process, such a process is the behavioural framework needed for thought and language to arise. David-

⁴²Cf. STL, pp. 138-141.

⁴³Cf. TVK, p. 212.

⁴⁴Cf. STL, p. 140.

son thinks of primitive triangulation as creating space for (and being part of) a more sophisticated triangulation: linguistic triangulation. For the creatures to triangulate linguistically, they must have the concepts related to objectivity: truth, belief, error, etc. Primitive triangulation ‘provides the framework’ for the idea of the concept of error.⁴⁵ In a tentative way Davidson gives the following scenario to illustrate this. If a creature has associated another creature’s response to certain conditions (association, of course, presupposes intentionality), and at some point the other creature produces the response without the expected conditions’ materializing, the notion of error may occur: the former creature may realize that the latter is mistaken. If it does so, the creatures succeeded to communicate content, and this requires language. Davidson finds it a natural assumption, on the basis of these considerations, that grasping the concept of truth (and the other concepts related to objectivity) and communicating content can only arise simultaneously. These notions are mutually dependent. Therefore, thought and language are mutually dependent.⁴⁶

Davidson sometimes describes the triangular setting as a language learning situation, which suggests that he thinks it points at some basic truth about what is needed to acquire a language.⁴⁷ Triangulation so conceived of involves a teacher, a child, and a shared world inhabited by objects and events. While triangulating, the child observes the teacher reacting to classes of objects and events it deems similar. The child is able to correlate those classes of objects and events with similarity classes of the teacher’s responses. The child will, following the teacher in this, use similar responses to reply to similar observations. The child need not use the same responses as the teacher, but for learning to be successful the teacher must be able to classify the child’s responses in types. The teacher will reward or punish the child, depending on whether or not the child responds in an interpretable way to the objects and events, i.e. according to how the teacher would classify those objects and events. That the teacher and child are responding to *those* distal objects and events (and not for example to retinal images) is so because those objects and events are the only common cause for teacher and child: here the causal chains (or, as Davidson calls them, ‘the lines of thought’)⁴⁸ from the teacher’s and the child’s skin into the world cross, and this is what teacher and child take to be the causes of their responses. A precondition for triangulation is that child and teacher (roughly) share their sensitivity to some rather than other environmental conditions and have the same idea of which stimuli and responses are of the same type.

⁴⁵Cf. STL, pp. 140-141.

⁴⁶Cf. STL, p. 141.

⁴⁷Cf. SP, p. 120.

⁴⁸Cf. SP, pp. 120-121.

3.3.4 The need for an external criterion

At first sight there is a blatant circularity in Davidson's idea of triangulation.⁴⁹ Davidson seems to tell us no more than that in order to have language and thought we must have language and thought: Davidson cannot explain language and thought in terms of triangulation and presupposes rather than explains that linguistic triangulators convey content. Davidson himself concedes that 'we move in a circle' at the point where we say that for the process to upgrade from primitive to linguistic triangulation, the creatures must convey content and grasp the concept of truth.⁵⁰

There seems to be a problem with Davidson's account indeed. The argument starts with the idea that in order to settle what a creature is reacting to another creature is needed to determine what the common cause of their reactions is. Davidson then suggests that triangulation (i.e. establishing a common cause) is a necessary condition for having a language. The motivation for this seems to be the idea that we could only make sense of someone's utterances' being about something if it can be determined, from a third person perspective, what they are about, and this requires that there be another person reacting to the environment, which enables locating a common cause at the intersection of the causal chains eliciting their responses. But then Davidson writes that such a primitive interaction is not sufficient for the creatures involved to speak a language: they must convey content, which means that they must interpret each other as speaking about the same features in their environment. That they are able to do is because *those* features are the ones that they naturally classify: 'we find the child's mouthings of 'table' similar, and the objects in the world we class together that accompany those mouthings is a class of tables. The acoustical and visual patterns that speed at their various rates between bell and dog ears, and tables and child eyes, we cannot easily observe'.⁵¹ But that means, it seems, that what we called primitive triangulation is explanatory vacuous; in fact, it seems that it is only *because* we think that there is a common

⁴⁹Cf. Verheggen (2013) for a discussion of the reception of Davidson's triangulation argument in the literature. Verheggen herself is sympathetic to the idea of triangulation. She does think there is a circularity in the argument, but she thinks the circularity in it is not to be vituperated, because of the very nature of the topic: the borderline between the non-intentional and the intentional. Verheggen also considers (and tries to defend Davidson against) criticism on empirical grounds: the idea that having language depends on having the concept of truth would be defeated by children with certain kinds of autism (who speak but cannot successfully do a false belief task), cf. Verheggen (2013), 3.3. For a highly critical and detailed discussion of all aspects of the triangulation idea (and all the dependency claims (with respect to language, thought, and communication) related to it), see Lepore and Ludwig (2005), chapter 22. Bernecker (2013) places triangulation in a broader discussion of Davidson's externalism. Carpenter (2003) discusses triangulation in the context of Davidson's argument for the veridicality thesis.

⁵⁰Cf. STL, p. 141.

⁵¹SP, p. 118

cause: we thinkers take ourselves to be communicating about tables and not about what is on the surface of our skin, and the causal chains as such do not determine this. Therefore, if we see animals triangulating in a primitive way (reacting to the environment and to each other in a coordinated way) *we* find it natural to explain their reactions in terms of certain causes rather than other features. But the fact that they triangulate seems not essential to our finding those causes rather than others natural. Indeed, Davidson writes about a dog's salivating when a bell rings, that we find it natural to take the dog to be reacting to the ringing of the bell rather than to a type of acoustical patterns produced by the bell (even though the latter, not the former, is necessary for the dog to salivate): this is not because the dog triangulates with other dogs, but—as Davidson writes—we find that an easy way to describe things (it is very hard (except by referring to the bell) to describe for us what the acoustical patterns produced by the ringing have in common). But even if one thinks there is a fact of the matter as to what the dog is responding to, it seems that primitive triangulation (involving more than one creature) is not needed to find the common cause. One may argue that the dog is *really* responding to the bell and that this can be tested by experimenting: one may increase the distance between the bell and the dog (so as to decrease the intensity of the acoustical patterns reaching the dog's ears), and if the dog still salivates (with the same intensity even) it seems that it reacts to the bell rather than to auditory stimuli (at least, such kind of experimenting may give as much certainty as observing the dog triangulating with other dogs). Nevertheless, we could grant that the idea of *the* cause is corroborated once we observe multiple dogs involved in a kind of primitive communicative process, responding to each others' responses (which adds an additional element: identifying what counts as a response of the same type).⁵² But even if primitive triangulation establishes a common cause, it is hard to say how it plays any explanatory role for the way we are able to communicate, since for that to be possible the communicators must already be sensitive to the same features of the environment (cf. the Principle of Correspondence).

However, we mistake Davidson's point in this way. Davidson's interest is not in analysing intentionality as such (which he thinks is irreducible to non-intentional phenomena like primitive triangulation), but rather in identifying the (non-intentional) conditions required for intentional phenomena like belief and language to occur.⁵³ What Davidson is concerned with is to

⁵²As we wrote above, Davidson does suggest that primitive triangulation widely occurs among non-thinking animals, by which he seems to mean that there is a fact of the matter what such animals are responding to (independent of how we find it natural to describe their behaviour). Cf. STL, p. 140. Whether this is so is however not crucial to what we think is the essence of the argument: that a sharing of a common cause underlies linguistic communication. Cf. below.

⁵³Cf. STL, p. 139.

establish that for our beliefs to be about anything there must be a criterion external to us as to what our beliefs are about: we cannot ourselves give this criterion. An external criterion is needed both for what the causes are of our hold true attitudes (both to locate the objects and events that our beliefs are about in space and time and to determine which objects and events must be classed together) and to determine which of our utterances count as of the same type (i.e. as expressing the same belief, in case they are sincere assertions). Davidson's point is that there can only be such a criterion if we communicate with others. And communicating with others, again, relies on there being common causes of our beliefs that we are able to identify on the basis of what we naturally find salient and similar (and can expect others to find so).⁵⁴

For Davidson a creature cannot give these criteria herself. Even if a solitary creature (a creature that never triangulated) could potentially have thought (like a child that was just born), there would be no way for her to go about in actually acquiring beliefs. The reason is that she cannot individuate objects of thought in such a way that we could sensibly say that her beliefs are true or false. She may have a tendency to be responsive to some features of her environment rather than others, but that as such would not suffice to determine what her beliefs are about. She may point to things in the world as being what she reacts to, but the identification of those things would rely on herself only: there could be nothing in the world that determines what she is thinking about. For that to be determined, another mind is needed. Even if, per impossibile, she could think, the criteria for when her beliefs are true could only exist in her head: she cannot objectify her beliefs by determining the conditions under which her beliefs are true with reference to another creature, which would allow her to locate the criterion in the external world (in between her and the other creature). Determining those conditions with reference to the world alone (without another creature) is impossible, since (as we saw) there is no fact of the matter what the creature is responding to (an observer not interacting with the solitary could not find this out, except maybe by appealing to what is natural to him). A solitary would therefore have to specify the links of her beliefs (or sentences) to the world herself, in terms of features she finds salient. This means that the connections between her beliefs and the world are internal, which for an externalist like Davidson is unintelligible. For a solitary, truth would therefore collapse to coherence.⁵⁵ Only interpretation can establish 'lines of thought' and connect beliefs to the world (and hence enables belief), on the

⁵⁴Cf. TVK, p. 212: 'The criterion on the basis of which a creature can be said to be treating stimuli as similar, as belonging to a class, is the similarity of the creature's responses to those stimuli; but what is the criterion of similarity of responses? *This* criterion cannot be derived from the creature's responses; it can only come from the responses of an observer to the responses of the creature.'

⁵⁵Cf. our discussion of the brain in the vat in 5.4.

basis of the intersubjective criterion that it generates.

One may think that if we consider a solitary creature that behaves (and speaks) sufficiently like us, it is natural to interpret that creature as speaking and having beliefs. However, Davidson holds that as long as such a creature has not *actually* been interpreted by others and has interpreted others, there is no fact of the matter what her words are about, and there is no way to tell whether she is mistaken or not. (We shall discuss Davidson's own example of such a creature below: Swampman.)⁵⁶

Davidson writes that the idea of having a thought (or belief) implies that the (content of the) thought is true or false independent of the existence of the thinker.⁵⁷ Davidson's argument seems to be that for that to be possible the thought must be 'public', in the sense that the same thought can be entertained by various people: only in that case is the thought independent of the thinker. That again is only possible if the evidence of the thinker's having the thought she has is (potentially) public: other thinkers can find out, by means of interpretation, that a thought they have is the same as the thought of the interpretee. But if there are no others to interpret the thinker (nor have there ever been), the content of the thought cannot exist independently: the thought does not exist (not even potentially) outside of the speaker. Thought for a solitary would be an illusion without content.

There seem to be two closely related considerations behind Davidson's argument against the idea of a thinker who never communicated. First, a solitary cannot establish a mind-independent criterion for the conditions under which her beliefs are true or false, whence it cannot be explained how her beliefs are connected to the world. Second, the content of her beliefs (assuming by contradiction that she could have beliefs) is merely psychological and not objective, since the only way in which thought could be objective (namely that the same thought content can be shared by people) does not apply to her 'thought'.⁵⁸

Somewhat tendentiously we could say that the only alternative to this is to think that thought is some mysterious entity existing independently of any thinker, that is connected to the world in a way we do not understand (to account for the first problem) and that can be 'grasped' by various people (to account for the second problem). Indeed, Davidson is fiercely opposed to the idea that thinking consists of 'grasping' an (abstract) (non-psychological) entity (like a proposition), as we shall see in the next chapter: the only way to explain the objectivity of thought for Davidson is by relating the mental process to the process of interpretation: interpretation can actualize the

⁵⁶Cf. 3.4. Cf. SP, p. 119: 'If we consider a single creature by itself, its responses, no matter how complex, cannot show that it is reacting to, or thinking about, events a certain distance away rather, say, on its skin. The solipsist's world can be any size; which is to say, from the solipsist's point of view it has no size, it is not a world.'

⁵⁷Cf. ET, p. 129.

⁵⁸Cf. 4.3.2.

human potentiality for having thought.⁵⁹

For the triangulator, on the contrary, there is an external criterion as to whether her beliefs are true or false: this is the criterion shared with those she communicates with. Her beliefs derive their truth conditions from the fact that she expressed those beliefs (or at least some of them) by means of sentences that she held true and that were interpreted as such by another triangulator: being interpreted anchors her beliefs to the world. For a creature to try to be interpreted (or to learn) requires that she is interpretable, and hence that she can identify conditions in the environment that the interpreter can also identify, whence intentionality must be in place already for interpretation to arise; but, if Davidson is right, interpretation is needed for there to be an external criterion and to objectify what is merely potentially present in a creature capable of thought. Davidson writes:⁶⁰

We have the idea of belief only from the role of belief in the interpretation of language, for as a private attitude it is not intelligible except as an adjustment to the public norm provided by language. [...] Can a creature have a belief without having the concept of belief? It seems to me it cannot, and for this reason. [...] [T]his requires grasping the contrast between truth and error—true belief and false belief. But this contrast, I have argued, can emerge only in the context of interpretation, which alone forces us to the idea of an objective, public truth.

Interpretation itself relies on beliefs: beliefs about the world, beliefs about one's own mind, and beliefs about other minds. This may seem like another circularity: interpreting and being interpreted presupposes having beliefs. Davidson admits that this question, in as far as it is a question of how for example a child can transcend the state of having no thought and enter the state of having thought, is enigmatic to him: Davidson thinks that this is an enigma because we have no proper way yet of understanding the semi-developed mind.⁶¹ Davidson's story is therefore rather a story of how we understand and explain the notion of belief as it applies to developed minds.

⁵⁹Cf. KOOM, p. 35 and 3.3.2.

⁶⁰TT, p. 170. The claim that we can only have the concept of belief if we communicate may seem empirical speculation (cf. Verheggen (2013), 3.2), but for Davidson it seems to be a conceptual truth following from his assumptions: if we agree with Davidson that having a concept requires that there be an external criterion (in the form of another possessor of a mind that we interpret and that interprets us), then we can only have the concept of belief (or any other concept) if we communicate. As long as we do not yet communicate, we cannot have the concept of belief (nor any other concept). A non-communicating creature could at most be under the illusion that it has the concept of belief, cf. 5.4.

⁶¹Cf. ET, p. 128.

The question arises here why it would not suffice for the triangulator to intend to be interpretable by another creature, enough like herself, for there to be a fact of the matter what her beliefs are about (without actually being interpreted). Davidson indeed thinks that the intention to be interpretable is the crucial commitment that someone should adhere to in order to belong to a linguistic community (as we shall see below), but for Davidson the objectivity of belief cannot be understood without there having been actual interpretation: only then can there be an actual criterion that the speaker can refer to.

Whether a solipsist will be convinced by this argument, is questionable. Lepore and Ludwig find Davidson's argument, after discussing it in much detail, in the end not convincing enough, and they suggest that Davison presupposes rather than proves the adequacy of a third person perspective on epistemology.⁶² Nevertheless, we think that it establishes at the very least that one cannot be both an externalist and believe that a private language (or a solitary thinker) is possible: externalism collapses to internalism in case of a solitary. And for Davidson internalism is unintelligible, exactly because of the considerations that give rise to his idea of triangulation.

What is therefore distinctive of linguistic triangulation for Davidson is *intention recognition*. This is for Davidson the source of normativity in linguistic communication: the fact that speakers try to recognize each other's *linguistic* intentions and that they intend to be interpretable in order to enable the other to recognize what one means.⁶³ For the speaker to be interpretable, an interpretive theory for her must ultimately be justifiable on the basis of non-semantic facts, the facts manifested by the human equivalent of primitive triangulation. In this sense the primitive triangle 'provides a framework' for linguistic triangulation (but for it to be the basis of interpretation presupposes that the triangulators conceive of each other's behaviour as intentional and explicable in terms of beliefs and desires, otherwise speech acts like assertion could not be identified).⁶⁴ Davidson presents his argument for intention recognition as the basis of normativity after discussing (Kripke's interpretation of) Wittgenstein's rule following considerations.⁶⁵ Davidson agrees with Wittgenstein that for a speaker to speak a language, she must be part of a community of speakers: a private language is impossible. However, Davidson disagrees that the norm for speaking a language consists of the speaker's complying with conventions, customs, or institutions: language is not essentially governed by shared rules or practices.⁶⁶ For Davidson complying with conventions is neither necessary nor sufficient

⁶²Cf. Lepore and Ludwig (2005), chapter 22, particularly Conclusions (pp. 418-419). Cf. 5.4 for a discussion of Davidson's third person perspective on epistemology.

⁶³Cf. SP, p. 116.

⁶⁴Cf. STL, pp. 140-141; 'provides a framework': STL, p. 141.

⁶⁵In SP.

⁶⁶Cf. SP, p. 114.

for someone to speak a language. What Davidson thinks is essential is that the speaker intends to be interpretable (on the basis of reasonable expectations and publicly available clues) and is in fact interpreted in a successful way. Speaking like others is a great help to be interpretable, but it is not necessary nor sufficient for being a speaker. For Davidson that speakers interpret each other's words in the way they intend them to be interpreted is what makes it possible to say that those words are about anything.⁶⁷ The intention to be interpreted amounts to subscribing to the norm provided by a linguistic community: it amounts to intending to be interpreted in terms of concepts shared with others.

The contents of thought are therefore determined socially for Davidson. In order to have beliefs, we must have communicated with others, and communication is constitutive of the contents of those beliefs, it is what anchors our beliefs to the world and determines the conditions under which they are true. Objectivity is therefore construed as intersubjectivity: our beliefs are objective in as far as they are interpretable by other creatures, on the basis of the same interpretive principles as we use.

Even though this argument is compelling, various questions remain (and Davidson is willing to admit that he just gave a crude outline of what he thinks is needed to have beliefs and language).⁶⁸ The most important question is maybe how it is possible that most of the times we have beliefs without being involved in communication. If Davidson is right, then—even though it may not be possible to see (unuttered) thought merely as a linguistic activity—at least the content of all thought must somehow depend on occasions of linguistic communication.⁶⁹ Another question is how often we must communicate in order for it to make sense to say that we still speak a language (and hence have beliefs). We shall however not go into these issues.

In chapter 4, when discussing Davidson's conception of truth, we shall see from another perspective why the idea of a solitary's having beliefs is unintelligible for Davidson: the truth of a belief for Davidson can only be explained in terms of the truth of a sentence expressing that belief, and the truth conditions of a sentence are determined by the principles of interpretation.⁷⁰

⁶⁷Cf. 1.3.1 on *interpretationism* and our discussion of Swampman in 3.4.

⁶⁸Cf. SP, p. 117.

⁶⁹Cf. 4.3.2.

⁷⁰Cf. 4.3.2.

3.4 Radical interpretation and triangulation

3.4.1 A tension?

The idea of radical interpretation and that of triangulation have been developed by Davidson quite in isolation from each other. Both seem essential to his account of meaning. There are, however, some apparent differences and even tensions between the two approaches to meaning. To illustrate this, we shall focus on Nathaniel Goldberg's work on this.⁷¹

One apparent difference is that triangulation is interactive, whereas radical interpretation seems not necessarily so. Pagin writes that for Davidson the radical interpreter is just a dramatic device.⁷² What determines meaning is the totality of facts that the interpreter, if he did his job well, would make use of, and meaning supervenes on these facts in accordance with the correct principles of interpretation. The interpreter himself could be left out, and there would still be meaning. However, if meaning supervenes on facts about (linguistic) behaviour, and no interpreter is needed for this to be so, it seems there could be a private language, contrary to Davidson's triangulation argument.⁷³

Moreover, the radical interpretation account and the triangulation account could be seen as given different criteria for what a speaker's words mean. Goldberg elaborates on this tension, and he illustrates it by Davidson's Swampman thought experiment. Swampman is an imaginative creature that seems to speak English (and therefore seems to be radically interpretable), but who never learnt what his words mean, and whose words therefore—according to Davidson—lack meaning:⁷⁴

Suppose lightning strikes a dead tree in a swamp; I am standing nearby. My body is reduced to its elements, while entirely by coincidence (and out of different molecules) the tree is turned into my physical replica. My replica, Swampman, moves exactly as I did; according to its nature it departs the swamp, encounters and seems to recognize my friends, and appears to return their greetings in English. It moves into my house and seems to write articles on radical interpretation. No one can tell the difference.

But there is a difference. [...] It can't mean what I do by the word "house", for example, since the sound "house" Swampman makes was not learned in a context that would give it the right

⁷¹Goldberg (2008) and (2012)

⁷²Pagin (2013), pp. 230-231. Pagin thinks this idea is confirmed by what Davidson writes in RQL, replying to a question by David Lewis about this.

⁷³Cf. Pagin (2013), p. 235.

⁷⁴KOOM, pp. 18-19. Davidson does not explicitly mention triangulation in this paper. However, triangulation and language learning are closely related notions for Davidson, as we saw above.

meaning –or any meaning at all. Indeed, I don't see how my replica can be said to mean anything by the sounds it makes, nor to have any thoughts.

Davidson claims that Swampman's words do not have meaning because he did not learn their meaning in the right context. Goldberg thinks this conflicts with the intuition that Swampman is (radically) interpretable (as being a speaker of English), and that therefore his words *do* have meaning. Of course, Swampman is just an unlikely creature and we may expect that intuitions about him are not clear. However, Goldberg thinks Swampman's case points to a deeper tension: a synchronic and diachronic meaning externalism. According to the synchronic externalism meaning depends only on relations that the speaker has to her present environment (as accessible to a radical interpreter). According to the diachronic account meaning (also) depends on her relationship to her past environment, which according to Goldberg's interpretation would in Davidson's case be a learning situation.

Goldberg's thinks the two accounts give rise to the following two meaning principles:⁷⁵

- (RI) For any term in L [in the speaker's language], the term means what it does just in case a radical interpreter under conditions of radical interpretation could interpret it in L to have that meaning.
- (LL) For any term in L [in the speaker's language], the term means what it does just in case a language learner of L under conditions of language learning learned that the term has that meaning.

According to (RI) the speaker's words have meaning only if a radical interpreter by radically interpreting the speaker can find out that they have that meaning. According to (LL), the speaker's words only have meaning if she acquired them as having that meaning in a triangular learning situation.

Goldberg thinks (RI) is the right condition, and (LL) should be abandoned. This does not mean that Goldberg thinks that Davidson's ideas on triangulation are irrelevant. Goldberg distinguishes (and claims that Davidson distinguishes) historical triangulation (a situation in which the meaning of terms is acquired) and ahistorical triangulation (triangulation that takes place between speakers while communicating (or trying to), for example between a radical interpreter and a speaker). He proposes that it is a condition for being interpretable that one acts *as if* one has been in the relevant learning situation, which may not actually have been the case. Goldberg thinks this condition is equivalent to (RI). According to this analysis, Swampman acts as if he learned what his words mean: he is able to triangulate with radical interpreters, and therefore his words have meaning.⁷⁶

⁷⁵Cf. Goldberg (2012), p. 153 and p. 155. Hahn (2003), pp. 44-45 and Bernecker (2013), pp. 446-447 also think that Davidson proposes two different meaning principles.

⁷⁶Cf. Goldberg (2008), section 4.

3.4.2 A unified account

In this last section we shall propose that radical interpretation and triangulation are part of the same account of meaning. We think Goldberg is right that radical interpretation involves triangulation and is hence essentially social, but we think Goldberg's distinction between ahistorical and historical triangulation is untenable.

Pace Goldberg, it is improbable that Davidson makes a distinction between historical and ahistorical triangulation. It would be strange to ascribe the view to Davidson that a learning situation far away in the past (unobservable by the interpreter) is essential to what a speaker currently means.⁷⁷ As we have discussed in the previous section, a speaker must be understood according to his own understanding of his words, and this understanding is determined by how he himself uses his words (not by how he happened to use his words when he learned them, though the two may very well be the same). Davidson does sometimes suggest that his externalism is historical, but he seems to mean nothing more by this than that what we mean is not determined by our present (physical) state alone, but by how we interacted with the environment prior to uttering our words, and those interactions are situated in the (nearby) past.⁷⁸ Goldberg thinks Swampman's words have meaning (they are English words) because he is radically interpretable. We think, however, that the problem is that Swampman has not (yet) *actually* been interpreted on the basis of his responses to objects and events in his environment to give meaning to his words, which is needed to turn his speech from a simulation of English into real English. The problem therefore is not that he is not radically interpretable, but that he has not *actually* been radically interpreted. Being involved in interactive radical interpretation would amount to a meaning baptism for Swampman, in which his words, through triangulation, are connected to the world.⁷⁹ As long as no triangulation has taken place, his words are not English, but just a phonetic simulation of English, lacking meaning. (Which, again, does not make Swampman less strange, it rather shows the consistency of Davidson's views.)

Hacker objects to Davidson's Swampman example that causal history does not matter to meaning, but only the speaker's ability to explain what his words mean and to respond appropriately to other speakers (and that hence a creature like Swampman is a speaker, because it apparently knows the meaning of all the words of English and it replies to other speakers in

⁷⁷Cf. NDE.

⁷⁸Cf. RTS, p. 165, where Davidson calls this historical externalism.

⁷⁹Cf. STL, pp. 139-140. These considerations do not make Swampman a less strange creature. Goldberg suggests that Davidson's account of meaning commits him to holding that Swampman is a zombie, without intentional states (and hence without intentions to be interpreted in certain ways). Cf. Goldberg (2012), section 6. One may explain this by saying that Swampman is like a child with intentional capacities, but who has not yet learned a first language.

an appropriate way).⁸⁰ This seems likely indeed, but Davidson's intuition behind Swampman seems based on the idea that a justified ascription of meaning to a speaker's words must ultimately be based on non-semantic facts. To interpret Swampman, we must be justified to ascribe a linguistic intention to Swampman to be interpreted in a certain way. In practice we often base the ascription of such an intention on a prior expectation of what the speaker's words mean, which also seems possible for Swampman (since he seemingly speaks English). However, for such an intention to be justifiably ascribable there must (potentially) be non-semantic evidence that the speaker has these linguistic intentions: it must be possible to justify the ascription on the basis of radical interpretation. For Swampman there is no such evidence: he immediately started uttering sounds after coming into existence, but as long as he has not been radically interpreted, we cannot justify the assumption that he speaks English. It seems, however, that once Swampman has been in touch with his environment for a while, it must be possible to justifiably ascribe linguistic intentions to Swampman as well on the basis of radical interpretation.

Radical interpretation and triangulation are therefore tightly related notions: the radical interpreter triangulates with the speaker, trying to detect the speaker's linguistic intentions on the basis of non-semantic facts.

⁸⁰Cf. Hacker (1998), pp. 551-552.

Chapter 4

The concept of truth: pattern and content

4.1 Introduction to chapter 4

In this chapter we shall be concerned with Davidson's reflections on the concept of truth. We shall start by considering Davidson's rejection of some influential theories of truth, particularly correspondence theory, coherence theory, the anti-realist conception of truth, pragmatic theories, and deflationary theories. We shall also see that Davidson is opposed to some dominant views in denying the existence of propositions and in taking sentences to be the primary truth bearers. (4.2) Instead of trying to define truth or providing some general hints as to what truth consists of, Davidson wants to clarify the concept by reflecting on the central role he thinks it plays in the interpretation of speech and action, thereby using Tarski's work as a formal tool. Tarski showed how to define true-in- L , for L a formal language. Some philosophers have suggested that Tarski's work is either philosophically irrelevant for explicating the concept of truth or supports a deflationary view on truth. Davidson argues that both views are wrong, and that the importance of Tarski's work can be appreciated by considering the role truth theories can play as empirical theories of verbal behaviour. (4.3.1) For Davidson this shows something important on the nature of truth and of truth bearers. (4.3.2) Constructing truth theories relies on a pre-theoretical grasp of the concept of truth, which Davidson thinks is not open to analysis or definition: Davidson takes truth to be a primitive concept. (4.3.3) Davidson thinks these considerations have consequences for what a semantic theory should achieve, and specifically for the role ontology plays in explaining what the truth of a sentence consists of. (4.3.4) We shall discuss how Davidson's conception of truth fits in with our idea that Davidson is concerned with vindicating a common sense realism. In this respect we are particularly interested in Davidson's conception of truth as verification-

transcendent and his idea that we can (nevertheless) not be wrong in most cases. (4.4.1) Lastly, we shall consider the compatibility of the indeterminacy of interpretation with realism. (4.4.2)

4.2 Rejection of existing conceptions of truth

4.2.1 Sentences as truth bearers

One important aspect in which Davidson deviates from some theories of truth is in that he takes sentences rather than propositions to be the primary vehicles of truth.¹ Davidson follows Quine in this. Propositions are often conceived of as truth bearers expressed by sentences (and sentences are then only truth bearers in a derivative way). They are further supposed to be the objects of propositional attitudes like beliefs. What is supposed to be a crucial feature of propositions is that they are shared among speakers of different languages: the same proposition can be expressed by sentences belonging to different languages and speakers of different languages can share a proposition as object of belief.² Davidson thinks such entities (and in general meanings as entities) lack explanatory value and he therefore wants to do without them and take sentences to be truth bearers. One of the motivations to introduce propositions (and in general meanings as entities) in semantics has been to account for contexts where an extensional treatment of meaning apparently fails, particularly for the content clauses ('that-clauses') of propositional attitude reports and indirect discourse. Davidson, however, has proposed a paratactic analysis of such clauses, in which (he claims) there is no need to introduce meanings or propositions. Davidson's paratactic analysis treats indirect discourse purely extensionally, by analysing an occurrence of a content clause in terms of a separate sentence (i.e. paratactically) that is referred to in a certain way by the sentence it is embedded in. The details of this will not concern us here.³

To be precise, not sentences but utterances and inscriptions are the most primary truth bearers for Davidson. Davidson conceives of sentences as useful abstractions: they are *types* of utterances or inscriptions. Due to the indexical features of natural language (as discussed above), sentences only have truth values if appropriately parametrized (to (at least) speaker and time).⁴ In *Truth and Meaning* Davidson proposes to see truth as a relation

¹Cf. TP, pp. 49-50 and pp. 123-124.

²Cf. McGrath (2014).

³Cf. ST.

⁴Cf. beginning of chapter 3 of TP. Technically, the language should also be specified (two languages may share a sentence type), although this is not needed if we think of sentences as entailed by a particular truth theory (and hence as part of a language). Davidson (implicitly) assumes there is a finite upper bound to the number of parameters needed. If this is not possible, it would hardly make sense to say that sentences, qua abstractions, qualify for a truth value.

between a sentence, a person, and a time.⁵ In the following we shall for convenience simply speak of true or false sentences, leaving out the required parametrization.

4.2.2 Rejection of realist, anti-realist, and pragmatist conceptions of truth

Davidson not only rejects the conception of truth bearer of some existing theories of truth, but existing theories or conceptions of truth in general. In this section we shall discuss Davidson's rejection of realist, anti-realist, and pragmatist conceptions of truth. In the next section we shall go into Davidson's rejection of theories according to which truth is a trivial concept, as opposed to what is (or seems to be) presupposed by the positions discussed in this section.

In Epistemology and Truth Davidson considers two opposed conceptions of truth: realist and epistemic (or anti-realist) views (which Davidson somewhat tendentiously classifies as 'objective' and 'subjective' views respectively).⁶ Davidson endorses some of the intuitions behind both conceptions, but rejects most of the things philosophers have said to flesh out these intuitions in an attempt to formulate a 'theory of truth'. Another tradition that Davidson thinks is on the wrong track is the pragmatist approach to truth (which, Davidson suggests, was born out of frustration over the traditional debates on truth).

Davidson conceives of the realist view as the intuitive idea that truth (or the real) is not dependent on our beliefs.⁷ Davidson agrees with the intuition that what is true or what is real does not depend on what we believe or on our mind, but he is dissatisfied with the way philosophers have tried to capture this intuition in accounts of truth. What Davidson has in mind when criticizing the realist conception of truth is (what he also calls) *metaphysical realism*: the idea that the world (construed as what makes beliefs true or false) and beliefs are independent of each other in such a way that the world could be the same and we could have very different beliefs (or vice versa). Davidson thinks that the metaphysical realist conception of truth is based on the flawed intuition that truth consists of correspondence of truth bearers to reality, which (as we saw in chapter 1) Davidson rejects: to say that a true sentence corresponds to reality is either trivial (i.e. it is just to say that the sentence is true), or, if it is not meant trivially, there is no way to make sense of it, since we cannot point out anything that a true sentence corresponds to.⁸ Davidson endorses the intuition that the real does not depend on what we believe, and that truth and falsity are verification-

⁵Cf. TM, p. 34.

⁶Cf. EPT, p. 178.

⁷Cf. TP, p. 33 and p. 45.

⁸Cf. 1.4.

transcendent, but he rejects the reverse: what we believe is not independent of what is true; belief is by nature veridical.

Epistemic theories for Davidson are theories that say (to put it very generally) that the scope of truth is in some way epistemically constrained.⁹ Davidson has in mind here a wide range of conceptions of truth, all of which have in common that they deny—to a significant extent—that there is a mind-independent reality that is relevant to whether beliefs are true or false. For Davidson the intuition behind such theories is essentially the idea that metaphysical realism is unintelligible: if reality is independent of the mind, there is no way to explain how our beliefs are connected to it, since we do not have access to a mind-independent reality. This is first of all a metaphysical thesis, but it has implications for the concept of truth.

A typical extreme form of an epistemic theory is a coherence theory of truth, the idea that truth consists of the coherence (or consistency) of our beliefs: there is not something outside the realm of beliefs that determines whether our beliefs are true or false.¹⁰

Another (less extreme) specific epistemic conception that Davidson discusses is Dummett's anti-realist view, according to which truth is restricted to what is verifiable.¹¹ We shall discuss Davidson's attitude towards this theory (and how his idea of truth contrasts with it) in some more detail, since it sheds light on Davidson's conception of truth. The idea of anti-realism is that there can be no fact of the matter as to the truth value of a statement unless there is (at least potentially) a way to prove that it is true or false: verification-transcendent statements, statements for which we (finite human beings) could a priori not have sufficient evidence to find out whether they are true or false, are neither true nor false.

A particular example that Dummett uses to illustrate this and that Davidson discusses is the sentence 'A city will never be built on this spot.'¹² For Dummett, this sentence has no truth value, because it is verification-transcendent: we cannot imagine having sufficient evidence to justifiably believe that this sentence is either true or false.

Davidson finds it unintuitive that such a sentence has no truth value, for which he gives two reasons: (1) We would have no idea what it means for our verification-transcendent beliefs to be true (since they have no truth value). (2) One could believe something (namely something that cannot be verified) without believing that it is true.¹³ These formulations beg the question in as far as they presuppose that we can believe something *and* believe that it is verification-transcendent (which Dummett may deny). However, there

⁹Cf. TP, pp. 45-47.

¹⁰Davidson regretted having used the notion of coherence in CTTK exactly because of its associations with a theory like this. Cf. AF.

¹¹Cf. Dummett (1959), p. 98.

¹²Cf. Dummett (1959), pp. 94-98.

¹³Cf. TP, pp. 46-47.

is a more fundamental difference of view behind Davidson's intuitions: for Davidson understanding an indicative sentence consists of understanding its truth conditions, whereas for Dummett it consists of knowing the conditions under which one can justifiably assert it. For Dummett, a speaker can justifiably assert a sentence if and only if the speaker has something that counts as 'conclusive evidence' for the sentence.¹⁴

Davidson adopts the reverse explanation: we learn an indicative sentence by learning under which conditions it is true, and we are justified to assert it if and only if we have evidence that it is in fact true. That we are globally right about when the sentence is true, is, in basic cases, simply so because the condition under which we hold the sentence true (or under which we learn the sentence) is the truth condition of the sentence. This is of course not so for sentences like 'A city will never be built on this spot.', the truth condition of which we could never be exposed to. The reason that we understand the truth conditions of verification-transcendent sentences (and the reason we can learn them at all) is therefore because of the way they are conceptually related to sentences that we can directly learn the truth conditions of, by observing that they obtain.

Davidson finds the idea of limiting truth to statements that we can have conclusive evidence for (or even defining truth as what is asserted on the basis of conclusive evidence) unintelligible.¹⁵ What is verifiable, Davidson argues, depends on human abilities, and human abilities change through time. A dilemma therefore occurs: either the standard of verifiability is defined in terms of our actual epistemic possibilities or it is defined in terms of the possibilities of an idealized agent.¹⁶ In the first case, the truth value of a sentence can change depending on the time it is uttered *merely* because of what we are able to verify at the relevant times, which Davidson thinks is counterintuitive. In the second case, the standard of verification must be independent of our abilities at specific times and hence be idealized, but it is unintelligible how such a standard can be defined (because we do not know what we will be able to verify in the future, and hence also not under ideal circumstances).

Davidson argues that both realist and epistemic views of truth give rise to scepticism (apart from the fact that he finds the former unintelligible in as far as it relies on a correspondence conception). The former because they give rise to the idea that we may be completely mistaken about the world. The latter because they 'make reality to less than we believe there is' (as Davidson puts it) by restricting truth to the domain of the verifiable.¹⁷

¹⁴Cf. TP, p. 46.

¹⁵Cf. TP, p. 46.

¹⁶Cf. TP, p. 46.

¹⁷Cf. EPT, p. 178. There is a curious footnote in *Truth and Predication* (p. 75, footnote 18), inserted by the editor Charles Parsons (who helped finalizing the manuscript after Davidson's death), discussing that Davidson wanted certain notes to be incorporated

As we saw in chapter one, in *A Coherence Theory of Truth and Knowledge* Davidson tried to define a theory of truth that captures on the one hand the correspondence intuition (the idea that the world, independent of the mind, is relevant to truth) and on the other hand the intuition behind the epistemic theories that there are essential connections between what is true and what we believe (and that what we believe is dependent on what is true). Davidson later wrote that as a *theory of truth* it failed because truth has nothing to do with coherence and the notion of correspondence is unintelligible.¹⁸ What is important about the essay is its argument for the veridicality thesis.

Another conception that Davidson rejects is the pragmatist conception of truth. Davidson has few kind words to say about pragmatism, even though (ironically) he has sometimes himself been associated with this tradition (and Rorty at some point tried to characterize Davidson's view as pragmatic).¹⁹ Davidson construes the pragmatic position as the idea that truth is a social construct: what we normally call true is true. Truth is then defined in terms of notions like optimal justification or scientific success.²⁰ What Davidson particularly rejects about this position is that it turns truth into a goal: if truth is defined in terms of what is optimally justified or what is predicted by successful science, it seems that we can strive for truth, which Davidson thinks is impossible, because truth has no criterion and justification (however optimal) does not guarantee truth. Truth cannot in any way be defined in terms of justification or any other cognitive or social process.²¹ What Davidson nevertheless appreciates about the pragmatist position (and epistemic conceptions) is the idea that 'human concerns, like language, belief, thought and intentional action' are important for understanding truth.²²

whose content suggests that Davidson thought an epistemic theory might be right after all. However, it is hard to interpret Davidson's remarks in those notes, and we shall ignore them. The notes are too hard to make sense and to connect to the general line of Davidson's argument. (See Parson's discussion; Parson argues that Davidson otherwise rejects epistemic theories, but does think that there is an essential connection between truth and belief, giving rise to the veridicality of belief.)

¹⁸Cf. AF. In AF (p. 155) Davidson writes that coherence is the same as consistency, and that beliefs are true cannot be explained in terms of their consistency. In fact, beliefs are by nature consistent to a great extent, Davidson writes.

¹⁹Cf. Rorty (1986).

²⁰Cf. TR, p. 6 and 7.

²¹Cf. TR, p. 7. This position is related to epistemic theories, but not the same: epistemic theories may also define truth in terms of what is optimally verified or justified, but epistemic theories do not think of the standard of what counts as justification as a social construct. For Dummett what counts as conclusive evidence for a statement is included in the meaning of a statement (it is how we learn what the statement means). Cf. Dummett (1959). Cf. TP, pp. 45-46.

²²Cf. TR, p. 17.

4.2.3 Rejection of deflationary conceptions of truth

Conceptions of truth not among the ones discussed above are deflationary theories of truth (there are several of them, which we shall not distinguish in our discussion). According to such theories, truth is not the interesting concept it has often been thought to be.²³ Deflationists think that truth is not a property: everything there is to say about truth can be found out by considering how the truth predicate is used, and employing the truth predicate seems to either serve no purpose at all or to enable us to talk about the truth or falsity of infinitely many sentences. As an example of the former, to say that snow is white is true (or to say that it is true that snow is white) seems the same as saying that snow is white. Predicating truth does not add anything to the statement. The only thing useful about the truth predicate is that it allows you to say things like ‘everything the pope says is true’: you can talk about the truth (or falsity) of a potential infinity of sentences, which you cannot do without having a truth predicate. Deflationists think there is nothing more to find out about truth than what is revealed by the quite trivial way in which the truth predicate is used. The concept of truth therefore has no great metaphysical or epistemological (or philosophical, for that matter) import, it is rather a linguistic concept.²⁴

Deflationists may be thought to find support in the work of Tarski, because Tarski showed how to define a theory that proves all sentences of the form ‘Snow is white’ is true if and only if snow is white’. This could be interpreted as that Tarski has proven that the truth predicate can be eliminated from all discourse.²⁵ In very general terms, the reasoning behind this could be explained as follows. Supposing that truth theories can be defined for natural languages (disregarding matters of consistency for the moment), Tarski has (implicitly) given an analysis of the use of the truth predicate even in non-trivial sentences like ‘everything the pope says is true’ (analysis of which seems to depend on a reference to the set of true sentences in a language), by showing how the set of true sentences of English can be defined in a metalanguage. The only thing that a truth theory needs to do to achieve this, is to entail all the *T*-sentences, and each *T*-sentence is trivial (at least in case English is included in the metalanguage and the sentence used at the right side of the biconditionals is the sentence mentioned at the left side).

Davidson rejects the idea that Tarski’s work supports the deflationist intuition. In general Davidson thinks deflationists are fundamentally wrong. We shall discuss Davidson’s reasons for this below, when discussing David-

²³There are several names for and variants of such theories, but we will restrict here to what they all have in common. The theory can be traced back to some observations in a paper by Frank Ramsey (Ramsey (1950)).

²⁴Cf. Stoljar, D. and Damjanovic, N. (2014).

²⁵Cf. TP, chapter 1, particularly pp. 13-16.

son's interpretation of the philosophical significance of Tarski's work.

What Davidson does share with deflationists is the dissatisfaction with existing attempts to give a substantive account of truth. Nevertheless, Davidson does think that truth is philosophically highly important and interesting. This makes Davidson's philosophy remarkable. On the one hand Davidson has an extensive negative programme consisting of an attack on traditional views on metaphysics and epistemology, specifically on the idea that we can somehow find an ultimate foundation for knowledge claims.²⁶ Whereas pragmatists, postmodern philosophers, and philosophers like Rorty, based on similar dissatisfaction, have tried to downplay what seems to be the cornerstone of traditional metaphysics, namely the concept of truth, Davidson's programme is to rehabilitate truth, and to assign a renovated, central role to it.²⁷

There is a long tradition according to which the concept of truth is one of the most important subjects for philosophical discussion, but in this century the tradition has come to be seriously questioned [...] because of various tempting errors and confusions. [...] The concept of truth should be restored to its key role in our understanding of the world and of the minds of agents.

One might wonder why truth is so important for Davidson. Is it for his programme in truth conditional semantics? There seems to be more to it. Truth for Davidson is the central concept in a theory of rationality, and hence for an understanding of what makes humans different from non-human animals (to which, as we have seen, rationality does not apply in any way for Davidson). More even than possessing language, it is our ability to distinguish between how we think the world is and how the world really is that Davidson thinks is distinctive of rationality. All the other notions of rationality (interpretation, a normative explanation of behaviour in terms of beliefs and desires, irrationality, etc.) apply to humans because they know what truth is. Moreover, for Davidson the concept of truth is the most elementary of all concepts: to have any concept at all, we must know the difference between what we believe is the case and what really is the case.

4.3 Tarski's truth theories and the concept of truth

4.3.1 The philosophical significance of Tarski's work

In many of his essays and in his only book, *Truth and Predication*, Davidson defends the importance of Tarski's work on truth. For Davidson Tarski's

²⁶Which will be discussed in the next chapter.

²⁷Cf. TR, p. 3.

formal work is not just important as a source of inspiration for truth conditional semantics, but it also helps us to understand the concept of truth. Tarski himself defended the philosophical significance of his work in *The Semantic Conception of Truth and the Foundations of Semantics*. Davidson, however, does not follow Tarski's interpretation in all details.

Tarski himself thought of his material adequacy criterion (the criterion that a truth theory must entail all *T*-sentences) as supporting what he deemed to be the intuitive, correspondence conception of truth.²⁸ Davidson, in an early work (*True to the Facts*), also thought that Tarski's truth definitions supported a correspondence theory of truth, albeit not the standard idea of correspondence, according to which sentences correspond to facts.

What then does Tarski's work tell us on the concept of truth for Davidson? Here it is fruitful to first consider two views on Tarski's work that are incorrect according to Davidson. The first of these is the idea that Tarski's work somehow proves that the intuition of the deflationists is right. The second is the idea that Tarski's work (as an account of the concept of truth) is philosophically insignificant.

In *Truth and Predication* Davidson discusses at some length why Tarski's work does not show that a deflationary theory of truth is right, as some have thought.²⁹ According to such an interpretation, Tarski has given theoretical support to the idea that the truth predicate has no interesting meaning: the extension of the truth predicate is captured by the trivial *T*-sentences, and Tarski showed how such sentences can be finitely axiomatized, and there is nothing more to find out about truth.

Davidson thinks this idea is flawed: this can be appreciated if we take the object language not to be part of the metalanguage. Think for example of the *T*-sentence 'Schnee ist weiss' is true-in-German if and only if snow is white'. This sentence is informative (and hence non-trivial) for a speaker of English who does not speak German.³⁰ This points to something deeper for Davidson, that he thinks is overlooked by deflationists: Tarski's definitions can be used to define truth for *a* specific language in *another* language, but he does thereby not define (nor provide an analysis of) *the* concept of truth. In fact, the idea that there is something about truth that truth theories cannot capture seems to follow already from Tarski's undefinability

²⁸Cf. Tarski (1944), section 3.

²⁹Horwich thought so (cf. TP, p. 12), cf. below. See particularly chapter 1 of TP on this issue.

³⁰In fact, once we drop the assumption that the language for which we define the truth predicate is (included in) the language of the theory (i.e. once the radical interpreter drops the assumption that the speaker speaks the same language as she does), even apparently trivial *T*-sentences are not trivial. I.e. if the radical interpreter (who speaks English) finds out that all instances of 'Snow is white' is true-in-the-language-of-the-speaker if and only if snow is white' hold, this is informative: she found out that she speaks the same language as the speaker.

theorem, according to which true-in-a-language cannot be defined within that language itself, which, *if* formal results have philosophical significance, seems to imply that there are limitations to what can be said on truth.³¹ Related to this, even if the use of the truth predicate is trivial in many cases, what seems not trivial is the fact that a language cannot consistently have its own truth predicate. The fact that natural languages have their own truth predicate, which is used in an apparently trivial way, is at the cost of consistency: they contain a liar sentence. This is the very reason that a Tarski-style definition can only be stated in a richer metalanguage. Tarski's work does not reveal, Davidson argues, what the truth predicates for all those languages for which it can be defined (within a metalanguage) have in common. His analysis is relative to a language (or relative to an object language and a metalanguage).

Deflationists have tried to meet the criticism that the truth predicate seems only trivial if we consider its use within a specific language by arguing that the triviality intuition applies to propositions as well, and if we take propositions (construed as what is expressed by a sentence) to be the primary truth bearers, the triviality intuition applies to the language-transcendent, general concept of truth. An example that Davidson discusses is Paul Horwich's idea that all instances of the form 'The proposition that p is true if and only if p ' exhaust our understanding of the *unrelativized* concept of truth. Asserting that a proposition is true is then the same as simply asserting the proposition. The truth predicate as applied to propositions would then be trivial, and therefore the concept of truth (captured by the instances of the schema) would be trivial.³² Davidson, however, finds Horwich's schema unintelligible, for which he gives various reasons the details of which need not concern us here (apart from the question what a proposition is, they have to do with the way a sentence is related (i.e. how it expresses or refers to) a proposition). More in general, Davidson argues that for Horwich's idea to work, Horwich must assume that we can understand meaning without understanding truth. Horwich would have to explain how we can attribute truth to sentences and utterances (in a derivative way) in terms of how we attribute truth to propositions. Such an explanation would presumably appeal to meaning, since what proposition a sentence expresses presumably depends on its semantic features. To avoid circularity, Horwich would then have to explain meaning without appealing to the concept of truth. For Davidson this is, of course, impossible, given his truth-conditional understanding of meaning. For Davidson it is exactly in our grasping meaning or content ('propositions') by understanding a sentence that a pre-theoretical and non-trivial understanding of truth is manifested.

The other interpretation of Tarski's work that Davidson attacks is the

³¹Cf. 4.3.3. Cf. FTDT, p. 21.

³²Cf. FTDT, pp. 31-34.

idea that it says nothing philosophically significant on the concept of truth. Tarski seems to only reveal a structure that truth makes in language, without telling us anything on what role the world (or anything outside of language) plays in an explication of the concept of truth. It is not easy to connect Tarski's definition with our intuitions about what truth is.³³

Davidson agrees that Tarski has not revealed all there is to know about truth. For Davidson the importance of truth theories—as instruments of analysis for natural languages (which Tarski himself thought his theories cannot be used for)—is in the fact that we can use them to describe a pattern observed in empirical phenomena. Truth theories define a structure that can be found in human behaviour, manifested in the use of natural languages. As Davidson puts it:³⁴

What Tarski has done for us is show in detail how to describe the kind of pattern truth must make, whether in language or in thought. What we need to do now is say how to identify the presence of such a pattern or structure in the behaviour of people.

Davidson thinks of a truth theory, used as an instrument for semantical analysis, as revealing constraints on the application of the notion of truth without defining truth.³⁵ In other words: Tarski showed how to define a pattern (not any of his particular truth definitions do, but the form that all his truth theories share) that truth makes in language, by showing how the truth conditions of infinitely many sentences follow from finitely many axioms (and hence how the truth conditions of sentences are related according to a recursive structure, determining patterns of truth preservation (logical form)). For such a theory to have an empirical application, an empirical touchstone has to be specified. This is the phenomenon of the speaker's holding true sentences (under certain environmental conditions). Finding out which truth theory applies to the language of a speaker depends on interpretation, and interpretation depends on an understanding and application of the *unrelativized* concept of truth, and this is what gives a truth theory for a language empirical content. That truth makes a certain recursive pattern in language is an a priori assumption for Davidson (charity demands this), but finding the correct theory is an empirical matter. This relies on an understanding of truth in two ways. In identifying the logical vocabulary (in accordance with the principle of coherence); and in finding the correct non-recursive axioms specifying the satisfaction conditions of predicates (and the reference of proper names), which determine when the

³³Cf. TP, p. 15.

³⁴TP, p. 28

³⁵Cf. TP, 55. Davidson compares this to what theories of measurement do by putting constraints on a central, undefined notion.

simplest sentences (of subject-predicate form) are true (in accordance with the principle of correspondence and the causal constraints).

While radically interpreting someone, we try to find a truth theory applying to his speech behaviour. We have already described how Davidson conceives of radical interpretation in chapter 1 and 2. We also discussed there how Davidson conceives of the theory of radical interpretation as part of a theory of rational behaviour in general. In *Truth and Predication* Davidson gives an outline of a more comprehensive pattern that informs normative explanation of behaviour, which is a combination of the theory of interpretation and decision theory.³⁶

In what sense does Davidson's story about radical interpretation shed light on the concept of truth? One may complain that it only shows how truth conditions are assigned to sentences, and that it gives a sophisticated account of the way the concept of truth is employed in a theory of behaviour, but that it does not say anything about what it means for a sentence to be true. What does truth consist of?

4.3.2 Davidson's behavioural approach to truth

One distinctive feature of Davidson's conception of truth is that truth, and the fact that a truth value attaches to certain entities, can only be understood in the context of human interactions. To see this, consider the kind of answer Davidson can offer (following from the above) to the question what it means for a sentence to be true (or false).

The (beginning of) the answer that Davidson offers, based on Tarski's biconditionals, is apparently trivial: the sentence is true *because* its truth conditions obtain. This answer means (according to Davidson) that a sentence is true because of just two things: what the sentence means and how the world is arranged.³⁷ However, Davidson gives this explanation more substance by offering an account of *why* a sentence has a certain truth condition. That account is given by adopting the perspective of a radical interpreter, who is trying to find a truth theory for the language the sentence is part of. By this move Davidson shifts the attention from the question what it means to say that a sentence is true to the question why a sentence has a certain truth condition.³⁸

Davidson's sketch of radical interpretation offers an account of why sentences have the truth conditions they have, by looking at them through the eyes of an interpreter. What the interpreter has to go on (initially) is just intentionally produced noises (utterances) held true. The interpreter abstracts over those noises by introducing types (sentences). She endows those (very elementary) sentences with truth conditions and tries to find a

³⁶Cf. TP, p. 74.

³⁷Cf. CTTK, p. 139.

³⁸For an analysis similar to the one given here, cf. Okrent (2011).

pattern among the sentences in order to construct a truth theory and to see the utterances as belonging to a language.

One may object that these considerations reveal nothing about why sentences are entities with truth conditions (or why they have certain truth conditions rather than others), because there is a fact of the matter about their truth conditions independent of what is revealed by looking at them as a radical interpreter, i.e. the sentences have truth conditions and truth values independent of their employment by speakers. In a sense this is true, because a sentence is only a sentence if we assume it is part of a language, whose semantics for Davidson can be stated by means of a truth theory; and a truth theory, it seems, exists independently of whether it gives the semantics of an actually spoken language (it is an abstract entity, that can be instantiated). However, consider what Davidson writes in *Truth and Predication*: ‘Nothing in the world, no object or event, would be true or false if there were no thinking creatures’.³⁹ In other words: without humans there are no entities carrying a truth value. There could be no truth theories either, since they rely on there being at least one actually spoken (meta)language to state them in.

For Davidson the fact that an entity has a truth value supervenes (directly or indirectly) on attitudes humans have towards that entity. Human beings endow entities (e.g. types of noises or ink formations) with truth conditions by holding them true under certain conditions, with certain linguistic intentions. Only because of the role of these entities in an explanation of how humans behave do they have a truth value. Truth bearers are truth bearers because they are held true and because they are interpreted as such. That they are in fact interpreted as having certain truth conditions by an interpreter is a requirement for them to have those truth conditions, since, as we saw in chapter 3, neither the speaker alone nor the speaker and the world can fix the truth conditions of a sentence: this requires another mind. Truth and falsity therefore emerge when interpretation takes place.

Nevertheless, sentences that are never used may also be true or false. An obvious example of such sentences are the sentences of a speaker’s language that are never uttered by the speaker. Their truth or falsity must be traced back to the explicit endowment of entities with truth conditions, by a speaker and an interpreter. They derive their truth conditions from the fact that they can be seen as part of a truth theory relating them to actually used sentences.⁴⁰

³⁹Cf. TP, p. 7.

⁴⁰As Okrent puts it: ‘Strictly speaking, sentences in a language, taken as potential utterance types in that language, that are never instantiated in any actual utterances, can on this view [Davidson’s view] also be true or false. On Davidson’s view, however, that there are such truth-bearing uninstantiated utterance types in a language is a function of the fact that the best truth theory for that language assigns truth conditions to those types of potential utterances. And, since truth conditions pertain to any utterance type only in light of the actual utterances of actual teleological rational agents, were there

Davidson therefore takes language to be the locus of truth and falsity, and the fact that entities have truth values is to be explained from the role those entities play in human interactions and the normative principles that govern such interactions. But what does this mean for *beliefs*? One motivation for the introduction of propositions is that speakers of different languages can have the same belief, and it seems plausible then to say that the contents of our beliefs (as what is shared by different speakers) are true or false in the most primary way, and the sentences only in a derivative way.

For Davidson, however, the order of explanation is the other way around: it is communication that makes it possible that we can have beliefs in the first place, and that two person can have the same belief. Therefore, the truth of a belief must be explained in terms of the truth of sentences, and in terms of interpretation. For Davidson propositionalist accounts rest on the mistaken idea that there are *objects of thought*: ‘The central idea I wish to attack is that these [the objects of the propositional attitudes] are entities that the mind can ‘entertain’, ‘grasp’, ‘have before it’, or be ‘acquainted’ with.’⁴¹

Davidson thinks of beliefs as certain states of a person that can be measured by sentences of various languages, in the same way as weight can be measured by using various units of measurement.⁴² In identifying and reporting our own beliefs we obviously use the sentences of our own language. That another person can have the same belief as I have depends on the fact that another person can interpret me as holding true a sentence that translates a sentence he also holds true. Therefore, what counts as having the same belief (and therefore the individuation of the contents of our beliefs) is to be explained in terms of interpretation: two person can only have the same belief if they are (potentially) able to interpret each other as having the same belief. Communication makes having beliefs and thought possible for Davidson, whence the truth of beliefs is secondary to the truth of sentences: ‘But though possession of a thought is necessarily individual, its content is not. The thoughts we form and entertain are located conceptually in the world we inhabit, and know we inhabit, with others.’⁴³

Radical interpretation relies on a pre-theoretical understanding of the concept of truth. This understanding is manifested in the interpreter’s ability to recognize hold true attitudes, in knowing the difference between what is held true and what is true, and in matching sentences she holds true with sentences the speaker holds true. However, we still seem not closer to a philosophical *analysis* of this pre-theoretical understanding of truth.

Here Davidson’s answer is radical. Davidson simply thinks we should

no such agents nothing could have such truth conditions and nothing could be a truth vehicle.’ Cf. Okrent (2011), footnote 18.

⁴¹KOOM, p. 35

⁴²Cf. IA, p. 74.

⁴³TVK, p. 218

not go further in trying to analyse truth than giving empirical content to Tarskian truth theories.⁴⁴ Davidson takes the concept of truth to be a primitive concept. He writes that truth is a ‘beautifully transparent’ concept and that ‘Truth is one of the clearest and most basic concepts we have, so it is fruitless to dream of eliminating it in favor of something simpler or more fundamental.’⁴⁵

4.3.3 Truth as primitive

Davidson wants to abstain from any attempt to define or further analyse truth (in as far as analysis is supposed to lead to a definition, or, if not a definition, an understanding in terms of other concepts). Moreover, he does not only reject a definition of truth, but also any attempt to capture something of its essence by means of general and vague formulae. Davidson argues that it does not make sense to try to define the most fundamental concepts we have, like truth. He considers any such attempt futile because of the circularity that is often involved in such definitions: when a philosopher tries to analyse a certain concept on its own, she has no choice but to take an understanding of the other fundamental concepts (equally potential matter of philosophical analysis) temporarily for granted in order to try to define the concept at stake in terms of the others. Davidson thinks there is a hopeless circularity in such an approach.⁴⁶

To judge from what Davidson *does* say on the concept of truth, his alternative to a definition is to illuminate the concept not by defining it, but by describing *why* it is important, what role it plays in our understanding of the world, and (hence) how it is related to the other fundamental concepts that philosophy is concerned with (like those of meaning and of belief).

A particular reason not to try to define truth for Davidson may be not to fall victim to circularity himself. The reason is that Davidson more or less analyses meaning, a fundamental concept that he deems much more obscure than truth, in terms of truth conditions (although he does not *define* meaning either). If a definition of truth then would (implicitly) rely on an understanding of the concept of meaning, we would end up with the kind of circularity Davidson thinks is problematic.

Davidson has another argument for the undefinability of truth: he takes it to follow from Tarski’s undefinability theorem.⁴⁷ Tarski’s theorem says that true-in- L cannot be defined in L . Davidson seems to infer from this that a definition of the concept of truth is impossible because such a definition would have to be stated in a language L , and would then implicitly also define true-in- L , which contradicts Tarski’s theorem. It is questionable

⁴⁴Cf. EPT, p. 190.

⁴⁵CTTK, p. 139 and TP, p. 55

⁴⁶Cf. FTDT, pp. 19-21.

⁴⁷Cf. FTDT, p. 23.

whether that as such is a good reason for abstaining from a *philosophical* analysis of truth. Natural language (the language of philosophical discourse) has inconsistent sentences anyway, which does not prevent us from philosophizing.⁴⁸ Even if we know beforehand that we will not succeed in giving a satisfactory *definition* of truth, our failures may bring to light important features of the concept.⁴⁹

However, maybe it is not a surprise that the concept of truth resists analysis if it is, as Davidson assumes, the core concept of rationality. For Davidson grasping the concept of truth and grasping propositional content are interdependent, as we saw. Having the concept of truth depends on linguistic communication, but understanding language again depends on understanding truth. It seems to naturally follow from this that the only thing we can do is describing how truth carries over from language to language, by relying on a prior understanding of truth, as Tarski showed. Davidson writes that it gradually dawned on him that whereas Tarski took for granted an understanding of meaning to explore the notion of truth, he himself had in mind the reverse: taking for granted the concept of truth, he tried to shed light on meaning.⁵⁰ Davidson thinks further analysis of truth cannot bring to light more than we already know:⁵¹

Truth, as applied to utterances of sentences, shows the quotational feature enshrined in Tarski's Convention T, and that is enough to fix its domain of application. Relative to a language or a speaker, of course, so there is more to truth than Convention T; there is whatever carries over from language to language or speaker to speaker. What Convention T [...] reveal[s] is that the truth of an utterance depends on just two things: what the words as spoken mean, and how the world is arranged. There is no further relativism to a conceptual scheme, a way of viewing things, or a perspective. Two interpreters, as unlike in culture, language, and point of view as you please, can disagree over whether an utterance is true, but only if they differ on how things are in the world they share, or what the utterance means.

If Davidson is right about this, it explains the mistakes of deflationists and correspondence theorists. Deflationists argue, based on the apparent

⁴⁸The undefinability theorem did not prevent Tarski himself from giving a vague definition of truth in terms of correspondence, which he thought was given content to by his theory.

⁴⁹Davidson ridicules Plato's attempts to define fundamental concepts: one of the few Platonic definitional achievements was to define 'mud' as 'water + earth'. Cf. FTDT, pp. 19-21. Nevertheless, Gettier's proof that Plato failed to define knowledge was considered highly significant and inspired sophisticated scrutiny of the concept.

⁵⁰Cf. Davidson (2001b), p. xvi introduction.

⁵¹CTTK, p. 139

triviality of *T*-sentences and their resistance to analysis, that truth is not only transparent, but in fact trivial. Correspondence theorists, on the contrary, have tried (in vain, if Davidson is right) to analyse the relation between meaning and ‘how the world is arranged’ in terms of entities in the world corresponding to sentences. A sentence, however, is not true because of some entity in the world, but simply because of the way the world is, and the only way to express the latter is by using the sentence itself (it is therefore not possible to analyse what the truth of a sentence consists of, since that brings us back to the sentence itself).⁵²

4.3.4 First truth, then ontology

Davidson thinks of truth theories as revealing the ontological commitments of a language. In *The Method of Truth in Metaphysics*, Davidson argues that this can only be done by constructing an adequate truth theory that reveals what entities are quantified over in a language.

For Davidson truth is therefore conceptually (and evidentially) prior to ontology. We observe speakers holding true sentences, and on the basis of a truth theory constructed in accordance with the observed behaviour we find out what entities their language is quantifying over.

Another, related way in which Davidson draws conclusions from the features of truth theories is by looking at what entities play semantic roles. This is his main concern in *Truth and Predication*. Davidson here discusses what he calls ‘the problem of predication’, which is the question what the unity of sentences consists of: what role do their constituent parts play in making the sentence a meaningful whole? For Davidson this amounts to explaining how the truth condition of a sentence can be explained in terms of its constituent parts. In Davidson’s conception of how this should be achieved we see very clearly how his programme in semantics mirrors his general philosophical views.

Predicates play a crucial role in explaining the unity of a sentence.⁵³ A name does not bear a truth value, but once combined with a predicate it does. Davidson attacks at length various approaches that have tried to explain the role of the predicate by assuming that it expresses an entity that, together with the entity referred to by the name, combines into something

⁵²Davidson argues there is some irony here. It seems that if the postulation of facts makes sense, it is only so for the simplest sentences. The truth of complex sentences could then be explained supposedly in terms of the facts corresponding to simple sentences (in terms of conjunctions or disjunctions of simple sentences). However, *if* this is possible (cf. the substitutional interpretation of quantification) the postulation of facts (or any ontology) is not needed: it suffices to specify the true sentences. Only if this is *not* possible, if quantification adds something essential (as is so according to Tarski’s objectual interpretation of quantification) it is needed to introduce ontology (but not facts), see next section. Cf. TF, p. 52.

⁵³Cf. TP, particularly chapters 4-7 for Davidson’s discussion of the following issues.

that is true or false. Davidson thinks this does not bring us any closer to an explanation: why would two entities taken together (whether or not there is a certain relation between them) be true or false? Davidson thinks that such a method (which has been pursued by philosophers ranging from Plato to Russell, and seems inherent to any methodology that tries to explain truth in terms of the (representational) properties of certain (abstract) entities)⁵⁴ reverses what should be the order of explanation: *that* a sentence is true or false is a primitive fact, something we can find out as radical interpreters.⁵⁵ What a semantic theory needs to do is not, in the first place, to explain the truth or falsity of a sentence, taken in isolation, in terms of something else (like a proposition), but rather to explain the truth conditional pattern among all the sentences of a language. Doing this involves the introduction of ontology (satisfaction functions), but—as Tarski shows—*not* of entities expressed by predicates: predicates are simply true *of* entities (there may be quantification over properties or relations, but there is no specific *semantic* role to play for an entity expressed by a predicate). For Davidson, Tarski also showed that *facts* play no semantic role (and this is, apart from the slingshot argument, another argument against fact-based correspondence theory that Davidson could appeal to): truth theories do not need facts to explain truth (even though facts may be a particular category in the ontology of discourse).

4.4 Davidson’s realism

4.4.1 Verification-transcendence

Davidson’s position on truth (and the accompanying metaphysical implications) has (have) been characterized variously, ranging from ‘naive realism’, to anti-realism, and pragmatism.⁵⁶

⁵⁴Plato is concerned with ‘the problem of predication’ in *Theaetetus* and *Sophist*. Cf. *Theaetetus* and *Sophist* in Plato, translated by Jowett (1892). Cf. Russell (1912), chapter XII and XIII. Cf. TP, chapter 5 for Davidson’s discussion of propositionalist accounts of truth.

⁵⁵Cf. TP, chapter 6 and 7.

⁵⁶For a discussion of the first (without a negative connotation), see Malpas (1992), chapter 7, and Okrent (2011) (Malpas and Okrent compare Davidson’s view with Heidegger’s *aletheia*). For an anti-realist interpretation, see Stoutland (1982) (I and II). Stoutland argues that Davidson is an anti-realist since not ontology, but people’s holding true sentences would determine which sentences are true and false. Stoutland thinks that for Davidson sentences are (held) true because of how well they serve extra-linguistic purposes (II, p. 41). Stoutland’s interpretation is, of course, unacceptable: for Davidson sentences are held true because of their analytical or inductive relations to other sentences held true or because we are caused to hold them true (in typical cases by what is constitutive of their truth conditions). Stoutland’s interpretation is also incompatible with Davidson’s idea that truth is verification-transcendent. In fact, his interpretation is pragmatic rather than anti-realist. For an overview see Glanzberg (2013).

Presumably this shows that Davidson's position cannot be captured in terms of traditional labels. What we think is distinctive of Davidson's position is that he combines the idea that belief is veridical with the idea that truth is verification-transcendent. There is at least an apparent tension between these positions: if the truth of my beliefs does not in any way depend on how I could verify or justify them, it seems that I cannot exclude that many of my beliefs are false. For Davidson this tension is dissolved by the idea that the conditions that typically cause me to believe something are the conditions under which the belief is true. This makes belief veridical.

Therefore, we could say that, even though Davidson rejected the omniscient interpreter argument to argue for the veridicality of belief, he is committed to the position that there is a fact of the matter as to what is true or false from an omniscient (or fully informed) perspective, which is presupposed by the idea that truth is verification-transcendent. The opposite, to restrict truth to what is verifiable, is unintelligible for Davidson, as we saw above. The constraints on interpretation, which are also the constraints that are constitutive of meaning and belief, guarantee that such an omniscient interpreter would judge that we are mostly right about the world.

The reason is that the content of our beliefs is informed by what causes us to entertain beliefs. This presupposes that the features or things that cause us to entertain beliefs exist (at least to an extent) independently of the mind: otherwise Davidson's position could hardly be seen as *externalist*. Indeed, for Davidson the content of our concepts (or at least of some of our concepts) is determined by what we typically apply them to. The priority therefore seems to lie in the world: we do not 'shape' the world by employing our concepts, but the world gives content to our concepts. However, Davidson would reject the idea that there is a real question as to whether the mind or the world has priority in the individuation of our concepts. The question (and especially the conceivability of the answer that the mind has priority) relies on a scheme-content dualism, an idea that Davidson famously attacked in *On The Very Idea of A Conceptual Scheme*.

Davidson writes that communication relies on an innate ability to classify certain things, based on an idea that some things are more similar than other things.⁵⁷ Such an ability is needed to communicate about the macroscopic features of reality, which forms the starting point of all communication. (It seems that the Principle of Correspondence is based on this ability.) For Davidson this does not mean that our mind somehow imposes a conceptualization on reality. The only conception we have of reality is in terms of these concepts: a world of water, mountains, and tables. Those are the things that cause us to entertain perceptual beliefs. There is no reality *beyond* that, which is transformed by the mind into the world as we

⁵⁷Cf. SP, p. 120 and TVK, p. 212.

perceive it. Rather, we are more sensitive to some features of the world than other features (which Davidson's suggests is innate and a product of evolution, and is therefore grounded in and constrained by the physical world).⁵⁸ For example, we are foremost sensitive to the macroscopic reality, and as intentional creatures we have a potential ability to classify macroscopic phenomena into types (shapes, tables, etc.). We further have a tendency to deem certain utterances as being of the same type. Such abilities allow us to communicate, which again is a prerequisite for having thought. As we saw in chapter 3, a solitary may also potentially have these abilities, but she cannot have thought for Davidson. The reason is not that she needs others to impose a conceptual structure on the world (the child does not need the teacher to *organize* the world), but rather that, even though she is situated in a mind-independent, objective world, whose features she is potentially able to classify, she cannot have an external criterion for the adequacy of her classifications.⁵⁹ Having concepts for Davidson is therefore not having an apparatus to organize the world. It is rather an ability to classify things already existing independently of us. The ability to use concepts can only be actualized by communication. We saw in chapter 3 that in order to classify things there must be an external criterion of correctness, provided by other speakers having roughly the same concepts. Such a community of speakers provides what Davidson calls a 'measure', the possibility to think about the world: 'A community of minds [...] provides the measure of all things. It makes no sense to question the adequacy of this measure.'⁶⁰ The idea that man is the measure of things is perfectly compatible with realism for Davidson, given that it makes no sense to doubt the adequacy of the measure in the absence of an intelligible alternative. The measure does not distort the measured for Davidson.

One may object that some imaginative tribe living in the same physical reality may form entirely different concepts, and that in that sense concepts do organize reality. Davidson thinks that such an idea is of no relevance to the question of realism, if it is not unintelligible in the first place. The only idea of the world we have is in terms of the concepts we share with other people, and it is on the basis of the features specified by those concepts that we interpret each other. Someone may have slightly different concepts, but if we cannot identify what typically causes someone to hold true his sentences (on the basis of a shared ontology), we can simply not conclude that such a person is a speaker talking about the same objective reality. Any interpretation must be based on features that we recognize, and we cannot imagine a speaker being caused to hold true sentence by features we do not

⁵⁸Cf. SP, p. 120 and TVK, p. 212.

⁵⁹Cf. EE, pp. 203: 'A noncommunicating creature may be seen by us as responding to an *objective* world; but we are not justified in attributing thoughts about our world (or any other) to it.' (Our emphasis.)

⁶⁰TVK, p. 218

(potentially) recognize. For Davidson this gives us enough reasons to say that the features we normally talk about are real and that language (or the mind) does not distort reality. We do not see the world *through* language (as if it were a filter), but *with* language. In this sense then Davidson is a realist. But maybe it is more suitable to say that Davidson is a ‘common sense’ realist, given his resistance to metaphysical labelling: there is no sensible question of what is real, what is real is just the world as we see it around us.⁶¹

⁶¹It is not clear, however, how Davidson’s realism relates to Putnam’s essentialist realism, which is of importance to an understanding of Davidson’s externalism. Cf. Putnam (1973). At some point Davidson writes that if an inhabitant of earth were to move to twin-earth (a place exactly like earth, except that on twin earth, unbeknownst to the speaker, there is no water but twater, a liquid with a different chemical structure but as far as concerns features discernible for the speaker exactly like water) he would be wrong if he said (pointing to a glass of twater) ‘that is water’. (Cf. KOOM, p. 22.) This is explicable from the fact that in the past the speaker used the word to refer to water, and he still intends the word to refer to water (although it seems that for Davidson after spending some time on twin earth the speaker’s word ‘water’ may start referring to twater). However, it seems that if we restrict the story to earth and we suppose that the liquid the speaker has always called ‘water’ on earth is actually not just water, but in a minority of cases it was twater, it is not easy to say what Davidson would say about the reference. Were the speakers partly wrong before they discovered this or did they mean both liquids? Putnam would probably adopt the first solution, but for Davidson this is not clear, since essences seem not important to his causal theory. Davidson writes somewhere that similarity of microstructure does not necessarily determine the reference of the word ‘water’. It seems that for Davidson there need not be an independent essence (like sameness of molecular structure) that plays a crucial role in individuating the causes that give content to our beliefs. (Cf. EE, p. 198 and further.) Only if the speaker intends that sameness of microstructure matters to the what his words apply to, does it matter. This, it seems, excludes that for Davidson we could be wrong about causes that could be individuated according to criteria we (nor any experts in our linguistic community) are not (yet) aware of, unless that is part of the speaker’s intention. (However, in Putnam’s scenario nobody in the linguistic communities on the two planets is aware of the microstructure of water, whence it seems that Davidson would have to say that after moving to twin-earth the speaker’s word ‘water’ also refers to twater, unless maybe it is implicitly part of the speaker’s assumption that the liquid on twin-earth is the same natural kind as the one he applied his words to on earth according to some conception of ‘natural kind’ that would make the difference: the speaker intends to be talking about stuff that is the same stuff as water on earth (i.e. the speaker intends that is water), even though he does not know what exactly the criteria for being water are.) The point of Davidson’s externalism is therefore not that mind-independent *essences* that we have been in causal contact with determine the meaning of our words, but rather that what we mean (how we intend to be interpreted) is a product of how a fully informed interpreter could reconstruct that intention on the basis of how we are causally connected to things in our environment. For Davidson these causes (even if not individuated according to essences) are objective. Further, externalism for Davidson pertains not primarily to specific words, but to our belief system as a whole. Cf. EE. p. 198. Cf. Child (2001) for a discussion of this issue.

4.4.2 Indeterminacy and realism

In this last section we shall briefly go into the indeterminacy of interpretation, which seems to be of importance to the understanding of Davidson's realism.⁶²

Davidson has an instrumentalist conception of the reference and satisfaction axioms of a truth theory. Various truth theories may equally satisfy the empirical demands even though they differ in their axioms. There is, however, not much choice for the recursive axioms for Davidson: any theory must have the recursive structure of quantification theory. However, all theories giving the right truth conditions are good theories, where the principles of interpretation (the empirical and formal constraints discussed in chapter 1 and 2) determine what the right truth conditions are.⁶³ The axioms are merely theoretical posits: any way of assigning referents to names and satisfaction conditions to predicates that (together with the recursive axioms) yields a theory satisfying the constraints is equally right. As a consequence, reference is inscrutable.

This is one kind of indeterminacy: there are always various empirically adequate theories that assign the same truth conditions to all the sentences but have different axioms.

In some cases, however, there is another kind of indeterminacy: there may be several empirically equivalent theories differing in their assignments of truth conditions to particular sentences.⁶⁴

The former indeterminacy is easy to accept if we accept that all the evidence relating to meaning is evidence pertaining to the sentential level. The

⁶²For Davidson's discussion of the issues in this section cf. IA, pp. 77-81.

⁶³As we discussed in chapter 2, this cannot mean that any truth theory that assigns truth conditions that are *extensionally* equivalent to those assigned by an adequate theory (e.g. by adding a conjunct with a tautology to the satisfaction conditions of a predicate) is itself also adequate. Instead, for any theory that assigns truth conditions that are formally *and* empirically adequate, there are numerous other theories that assign truth conditions that are formally and empirically adequate. However, Davidson is not very clear on this issue in IA, he might also be read there as saying that any theory that assigns truth conditions that are extensionally equivalent to those assigned by a given empirically adequate theory is also empirically adequate (it depends on what he means by a theory's role in determining *the* truth conditions of sentences, cf. IA, p. 79: are *the* truth conditions merely extensionally adequate, or do they have to satisfy additional constraints?). But that would leave unexplained why empirical constraints are needed next to formal constraints, which we discussed in chapter 1 and 2 (which seemed to be so because extensional adequacy does not suffice for the theory's being interpretive), as Davidson himself admitted. The examples Davidson gives in IA depend on the possibility of changing the satisfaction conditions of predicates and compensating this by changing the reference of names. (Cf. IA, pp. 78-80.) This is a possibility that you would not want to be ruled out by an empirical constraint (if that were possible at all), we think, so it is compatible with the idea that there are empirical constraints on truth theories, which (we think) do not bear on the kind of indeterminacy discussed by Davidson. Cf. our note to 2.3.2.

⁶⁴Cf. IA, pp. 80-81.

second kind of indeterminacy is more radical, however: how could there be a fact of the matter as to whether one of my beliefs is true or false, if this depends on which truth theory an interpreter uses to interpret the sentence expressing that belief? For Davidson the latter indeterminacy follows from the rejection of the analytic/synthetic distinction. In some cases there is no way to find out whether disagreement between speaker and interpreter is due to difference in meaning or difference in beliefs (both explanations can be accounted for by proper adjustments of the theory of behaviour as a whole). The example Davidson gives is of a quarrel over how to classify colours.⁶⁵ The interpreter can either decide that the speaker uses slightly different colour concepts or that he has slightly different beliefs on the colour of objects. Davidson thinks that in such a case there is no fact of the matter which explanation is right. Presumably such dispute is only possible against a background of agreement on concepts, and is therefore the exception rather than the rule. Nevertheless, this seems to imply that for at least some of a person's beliefs there is no definite answer as to whether they are true or false, since if they express the belief by an utterance, the utterance can be correctly interpreted both as a true utterance and as a false utterance. However, a more suitable way to express what follows from the indeterminacy at issue is that what belief exactly is expressed by an utterance can only be understood in the context of an holistic interpretation. For some utterances this means that, depending on the theory of what a speaker means and believes as a whole, the interpreter can either attribute a true belief and slightly deviant concepts (from her perspective) to the speaker or a false belief and the same concepts. This then shows that belief is truly holistic for Davidson: individual beliefs (conceived of as what is expressed by a sincere assertion of the speaker) only have a clear identity against the background of a broader interpretation of what the speaker means and believes. For some utterances of beliefs, this matters to their truth value. This would be so even if an omniscient interpreter interpreted the speaker. But on any interpretation, most of the speaker's beliefs are true. The veridicality thesis must also be understood in this light.

⁶⁵Cf. IA, pp. 80-81.

Chapter 5

Davidson's attitude towards the sceptic

5.1 Introduction to chapter 5

In the fifth and final chapter we shall discuss Davidson's attitude towards the sceptic. Davidson's veridicality thesis seems to imply a refutation of some important forms of scepticism. Davidson's position on scepticism is widely debated in the literature and Davidson himself has struggled with formulating his view on it.¹ The two most important questions are whether Davidson is able to provide a sound and non-question begging answer to the sceptic and—if so—what exactly it means for scepticism: does it imply that the sceptic is simply mistaken, or rather that scepticism cannot even be formulated consistently? We shall try to answer these questions.

First, we shall discuss the kinds of scepticism that we think Davidson's ideas have implications for: scepticism regarding the veridicality of belief and scepticism regarding the possibility of knowledge. (5.2) We shall then argue that Davidson's argument for the veridicality thesis, construed as a direct argument against the sceptic on her own terms, begs the question (at least at first sight). (5.3) Instead, Davidson is concerned with formulating an epistemological theory which has as a consequence that we can be said to know that the sceptic is wrong. For Davidson sceptical scenarios are incompatible with the nature of belief, and that they are is a condition for knowledge to be possible. (5.4) Therefore, (traditional) scepticism cannot get off the ground from a Davidsonian epistemological perspective. (5.5)

¹The issue is for example discussed by A.C. Genova (Genova (1999)), Thomas Nagel (Nagel (1999)), Barry Stroud (Stroud (1999)), Anita Avramides (Avramides 2013), and Duncan Pritchard (Pritchard (2013)). Cf. RTS, p. 163 for Davidson's reflection on his position.

5.2 Veridicality and scepticism

It seems that the sceptic that could be refuted by the veridicality thesis is the one who maintains the opposite of the veridicality thesis, namely the idea that it may be the case that all our beliefs are false together, and that this is a possibility to be taken seriously, since we do not know (we do not have a good argument) that this is not actually the case. The sceptic may for example argue that we have no good reason to believe that we are not brains in the vat, and that *if* we are a brain in the vat, many of our beliefs (for example the belief that we have a body) are false, whence we have no good reason to believe that most of our beliefs are true.

In this sketch of the position of the sceptic we assume that, even though the sceptic believes that she may be in a situation in which all her beliefs are false, she still has those beliefs. However, these considerations may be a reason for the sceptic to suspend judgement, and to abstain from knowledge claims. We shall see that this is the kind of sceptic that for Davidson is the most interesting: the sceptic who challenges the possibility of knowledge. In the sceptical tradition, this challenge is mostly formulated in a somewhat different way: as the question how knowledge claims can be appropriately justified. The sceptic argues that there are only three possibilities for the justification of knowledge claims: justification rests on some ultimate unquestionable evidence (or belief), justification is circular, or the process of justification is infinite.² All of these are (prima facie) problematic, since each seems to imply that a knowledge claim cannot be founded on something that is certain (unless something is dogmatically postulated to be certain). The sceptic therefore argues that our knowledge claims cannot be justified, and that we therefore should not claim to have knowledge.

Davidson does not directly address this sceptical challenge in the form we stated it here. In most essays, Davidson is rather concerned with *belief* and with arguing that belief is veridical than with *knowledge*. Even if belief were veridical (and even if we knew this) this does not directly imply that we have any knowledge (except *that* belief is veridical). However, it seems that Davidson, instead of answering the sceptic's question, wants to undermine a presupposition of the challenge: the idea that knowledge claims must be based on something certain by means of justification. Davidson wants to argue that the veridicality of belief explains why we can have knowledge.

Thomas Nagel has suitably explained Davidson's position as a reformulation of the Cartesian *Cogito ergo sum*.³ Whereas for Descartes the existence of thought implies the existence of the thinker, for Davidson the existence

²Cf. Klein (2014) for a good overview. There are two important traditions: Academic scepticism, maintaining that we know that we cannot have knowledge of empirical propositions, and the more radical Pyrrhonian scepticism, maintaining that we do not even know whether knowledge of empirical propositions is possible.

³See Nagel (1999).

of thought implies that we have knowledge, not only of our own mind, but also of the world, and of (the content of) other minds.⁴

5.3 The veridicality thesis as argument against scepticism

Davidson's argument that belief is veridical is often considered to be an argument against scepticism. Davidson himself is not directly concerned with the sceptic. For example, in *A Coherence Theory of Truth and Knowledge*, where Davidson most explicitly advocates that belief is veridical, he does not present his argument as an attempt to refute the sceptic, and he explicitly denies that this is his aim.⁵ Nevertheless, the thesis seems to imply a rejection of (forms of) scepticism, and Davidson is, of course, aware of this. And as we shall see, Davidson thinks that a proper epistemological theory *should* exclude scepticism to a considerable extent.

It is an interesting question whom Davidson's argumentation in *A Coherence Theory of Truth and Knowledge* is supposed to convince. Davidson writes (suggesting that he is interested in scepticism after all) that his argument can rescue us from a standard form of scepticism by showing why it is impossible for all our beliefs to be false together.⁶ But whom does Davidson want to convince by arguing that belief is veridical? Davidson at some point suggests his veridicality argument is supposed to refute the sceptical possibility that we are wrong that there is a world inhabited by macroscopic objects, as we normally conceive of it.⁷ However, suppose that before reading *A Coherence Theory of Truth and Knowledge* we consider it to be possible that there is no external world, and that all our beliefs regarding it may therefore be false. Would we be convinced by Davidson's arguments that we are mistaken?

One reason to think we should not be convinced is that Davidson's argument presupposes rather than proves the existence of a world independent of our mind. Davidson writes at the beginning of the essay what he thinks is part of the nature of belief: 'beliefs for me are states of people with intentions, desires, sense organs; they are states that are caused by, and cause, events inside and outside the bodies of their entertainers.' He then sets out to argue that it follows from the way our beliefs have content that most of our beliefs are true. However, if Davidson assumes a priori that we have

⁴In the following, when we are concerned with scepticism, it is one of the kinds of scepticism identified explicitly. We are for example not concerned with any implications for scepticism with regard to induction when writing that the sceptical position is inconsistent for Davidson.

⁵Cf. AF, pp. 157.

⁶Cf. CTTK, p. 153.

⁷Cf. RTG, pp. 193-194.

bodies, and that our beliefs are causally connected to events inside and outside those bodies, he seems to exclude a priori the sceptical scenario (there is no external world) which we want to be shown to be impossible by an argument.

Genova therefore rightly points out that an external world sceptic (or a sceptic that maintains that all our beliefs could be false) would not accept the premisses of Davidson's argument. Therefore, Davidson's argument at most proves wrong a mild sceptic, one who assumes (like Davidson) that there is an external world, and who accepts Davidson's externalism, but nevertheless argues that all our beliefs *could* be false (in as far as they are not already presupposed to be true by accepting the premisses). Davidson then argues that this is impossible.

A radical sceptic may, rather than giving up the sceptical possibility, reject the idea that every belief system is radically interpretable by an omniscient interpreter in the way specified by Davidson, or (equivalently), that causal connections inform our beliefs so as to make them globally veridical.

Nevertheless, strictly speaking this would mean that Davidson and the radical sceptic disagree on the nature of belief, and that therefore they cannot really argue on the same terms. We shall see that this is indeed what Davidson's position implies, but to fully appreciate Davidson's argument, we have to consider first Davidson's project in epistemology.

5.4 Davidson's new *Cogito* against the sceptic: context and motivation

Instead of refuting the sceptic, Davidson is rather concerned with the question how we can have beliefs about the world. Davidson follows Quine's example in adopting a naturalistic approach to epistemology. This amounts to looking at epistemology from (what Davidson calls) a third person perspective. By this Davidson means that the task of epistemology is to explain—in common sense or scientific terms—how agents acquire beliefs and knowledge about the world. Naturalized epistemology is concerned with an empirical explanation of how actual agents come to entertain beliefs and knowledge.⁸

This is in direct opposition to a first person approach, the most prominent example of which is the approach of doubt, that Descartes adopted: a first person approach starts from 'within the mind' and assumes nothing but that we have thought, and then tries to find out what follows from the fact that we think.⁹ A naturalistic approach presupposes what a first person approach doubts: an external world, inhabited by minds that acquire knowledge.

⁸Cf. EE, pp. 193-194.

⁹Cf. Descartes (1644).

For Davidson, however, the first person approach is not merely a choice that he rejects, he thinks of it as a route that is philosophically flawed, based on a mistaken idea of epistemic priorities: Davidson holds that we cannot without contradiction think that we have thoughts *without* by implication assuming that there is an external world with other minds we communicated with: this is a precondition for having thought. A consequence of this is for Davidson that Cartesian scepticism, in as far as it says that one's having thought is the most primary certainty, and that all other things are less certain (or may all be false) is inconsistent.

While applauding Quine's third person approach, Davidson rejects Quine's foundationalism, which he thinks is a relict of the first person approach.¹⁰ For Quine our beliefs are ultimately all based on something other than belief, namely on what goes through our sense organs. Davidson rejects the idea that our beliefs are based on (or justifiable in terms of) something other than beliefs. The main motivation for Davidson to reject Quine's foundationalism is that for him the idea of an unconceptualized given, that our beliefs are based on, makes no sense: for something to justify a belief, it must have propositional content, which Davidson thinks can only be another belief. Davidson thinks Quine's foundationalism is also inherently sceptical (apart from that it is unintelligible): if our beliefs are based on something, this something is an epistemic intermediary, in between the things and our beliefs, and may be unreliable. If Quine is right that our beliefs are based on sense data, whether our beliefs are reliable depends on whether our senses are reliable.¹¹

Davidson not only rejects Quine's foundationalism (with its inherent scheme-content dualism) but any epistemological theory that bases beliefs on some kind of evidence, be it something outside the realm of beliefs or on some special kind of beliefs that constitute the basis of all other beliefs. Davidson does think that we acquire knowledge of the world through perception (and the beliefs that are engendered by it), but he thinks both that all perceptual beliefs are revisable (and are hence not an unquestionable foundation, but open to support by other beliefs) and that there is no clear distinction between perceptual and theoretical beliefs on epistemological grounds.

In A Coherence Theory of Truth and Knowledge Davidson proposes an alternative to foundationalism, motivated by the idea that there must be some other way than the one proposed by foundationalism in which our beliefs are connected to external reality. According to this alternative, the connection of beliefs to reality does not consist of our basing our beliefs on something other than belief, but rather of the fact that beliefs—especially perceptual beliefs—are veridical by nature. This is motivated by the idea

¹⁰Cf. EE, p. 194.

¹¹Cf. 1.4.

that beliefs are normally engendered under the conditions that make them true. The connection between beliefs and the world for Davidson does not consist of *epistemic*, but merely of *causal* intermediaries.

By means of his triangulation theory, Davidson gives a further elaboration of his causal theory of belief: it explains how there can be a fact of the matter as to what in the world causes our beliefs. The triangulation theory (which we discussed in chapter 3) embodies a view on epistemology, which—as opposed to Quine’s—is not merely third person but essentially social.

In Davidson’s triangular epistemology, there is no kind of knowledge that has priority. In an essay by the same name, which is a good overview of his epistemological views, Davidson distinguishes three varieties of knowledge:¹² knowledge of one’s own mind, knowledge of the world, and knowledge of other minds. All of these three kinds of knowledge concern a vertex of the triangle that is essential to language and thought. For Davidson these three kinds of knowledge are all equally basic, mutually dependent, and (hence) irreducible to each other. Davidson begins his essay by briefly discussing various approaches to epistemology that have tried to take one of these kinds of knowledge as basic, and to see the other ones as derived, or dependent on it.¹³ One example of those approaches is the Cartesian approach, which takes knowledge of one’s own mind as basic. Platonism and rationalism are other examples of approaches that distrust empirical beliefs and construe knowledge as a priori. Empiricism, on the other hand, conceives of all knowledge as based on experience, whence knowledge of the external world is conceptually prior to other kinds of knowledge (specifically to knowledge of another mind). Behaviourists (also) try to base knowledge of (other) minds on knowledge of the world, and hence to construe it as a derived kind of knowledge. According to Davidson’s triangular externalism, on the other hand, all three kinds of knowledge are equally essential for thought to be possible. The existence of thought presupposes a mind of the thinker, an external world, and other minds, all of which are essential for thought to be objective and to have content. Knowledge (by the thinker) of all of these three is essential to have thought. None of these three kinds of knowledge is conceptually prior to the others. Knowledge of the world would not be possible if we had no knowledge of (the contents of) another mind, without which we could not communicate in order to individuate the content of our thought. Knowledge of other minds, however, depends on knowledge of the world, because we cannot directly observe the contents of another mind, but only on the basis of intentional behaviour manifested in the world. Knowledge of one’s own mind depends on knowledge of the world and knowledge of (the contents of) another mind, since thought depends on

¹²Cf. TVK.

¹³Cf. TVK, pp. 205-208.

communication.¹⁴ Therefore, for Davidson scepticism with regard to other minds (or the contents of other minds) is inconsistent, unless it is to imply scepticism with regard to (the contents of) one's own mind. If Davidson's triangular account is right, we cannot not have beliefs if there were no other minds, whence it would be inconsistent to believe that there may be no other minds.

All these kinds of knowledge must come at once to have propositional thought, and we cannot coherently deny one of them, while retaining the others (for Davidson that would amount to the same kind of circularity manifested by defining basic concepts like truth in terms of other basic concepts). The difference between the kinds of knowledge is just in fallibility, not in priority. Only knowledge of one's own mind is—to a large extent—infallible: regarding knowledge of one's beliefs, it just amounts to knowing what one's sentences expressing those beliefs mean, and this again is determined by the conditions under which we typically hold true those sentences. Knowledge of one's own mind is therefore trivial for Davidson: it is manifested by the speaker's uttering sentences that express his beliefs in such a way that those beliefs can be 'measured' by interpreters in terms of their own sentences. For Davidson it is only because knowledge of our own mind is immediate and infallible, that there is something for others to interpret: how we intend to be interpreted.¹⁵

We have also discussed in chapter three that not only need we presuppose all these three kinds of knowledge to explain thought for Davidson, it also follows that one's beliefs concerning all three sectors of knowledge must be largely true.

What does this mean for Davidson's answer to the sceptic?

Thomas Nagel thinks that Davidson, instead of refuting the sceptic, has shown that scepticism cannot be expressed.¹⁶ The sceptic holds that all our beliefs could be false. Davidson's position implies that our beliefs would not have content if they were all false: if the world external to our thought is not pretty much as we think it is, it is inexplicable why our beliefs have the content they have.

Nagel thinks of Davidson's position on scepticism as replacing the Cartesian *je pense, donc je suis* by *je pense, donc je sais*: the existence of thought does not only require that the thinker exists, but also that the thinker is right about her own mind, the world, and other minds.¹⁷

¹⁴Cf. TVK, p. 213.

¹⁵Cf. TVK, p. 213. It is often thought that externalism is incompatible with infallible access to one's own mind. Davidson argues against this idea at length in KOOM, arguing that knowledge of one's own mind is trivial: it simply amounts to knowing under which conditions one holds true one's sentences.

¹⁶Nagel (1999). Nagel's essay is mostly based on A Coherence Theory of Truth and Knowledge. His analysis could be further supported by Davidson's ideas on triangulation.

¹⁷Cf. Nagel (1999), p. 196. Davidson himself writes in his reply that he had previously qualified his position as 'I think, therefore you are' Cf. RTN, p. 207.

Therefore, Davidson challenges the sceptic's claim that she could still have beliefs if she were completely mistaken about the world. For Davidson the sceptical position is inconceivable: if we were completely mistaken about the world, we could not have the concepts required to believe that the world is one way rather than another. In fact, we could not have any concepts. Of course, it is possible to learn many concepts without ever having correctly applied them to anything in the world, simply by learning their definitions, in terms of concepts already known. However, for this to be possible, there must be concepts to use in those definitions, and at least some of those must have been learned directly, in a triangular learning process. And the way we learned to classify objects and events by means of those concepts that we acquired directly is what constitutes their meaning. If a sceptic argues that she does not know how to apply correctly any of her concepts—for example because she thinks she does not know whether there is any world with objects and events in the extension of those concepts—she cannot, for Davidson, consistently maintain that she entertains any beliefs.

Let us have a look at the brain in the vat scenario to illustrate Davidson's position.¹⁸ Putnam argued that we could not be a brain in the vat *and* have beliefs about a world outside the vat.¹⁹ Davidson does not himself go into this particular scenario, but his position implies an answer at least as anti-sceptical as Putnam's. In the brain in the vat case all our beliefs about the world are assumed to be induced by the electrical impulses fed to our brain by an evil scientist. In this scenario there is an external world (the world of the evil scientist), but our beliefs about it are manipulated (we are not in 'direct touch' with the world) and—hence—to an extent, at least, false (for example, the belief that we have a body is false); or at least that is supposed to be possible.

According to Putnam's solution, the problem with the scenario is that it presupposes that our empirical beliefs could be about a world that we have never been directly in touch with. For Putnam (and for Davidson), however, what causes our beliefs is constitutive of their content. This means that *if* in the scenario our beliefs are caused by electrical impulses fed to our brain (and not by the macroscopic features of the world outside of the vat), they are not about the world outside of the vat. Therefore, our word 'tree' in that case refers not to trees in the external world, but to something else. The same goes for our words 'brain' and 'vat' (they do not refer to brains and vats in the external world), whence our thought 'I am a brain in the vat' would be false. Therefore, we are not a brain in the vat.

This answer seems not entirely satisfactory. The problem is that it gives the feeling that an *epistemological* problem is refuted on the basis of *semantic*

¹⁸Cf. RTN, p. 205.

¹⁹Cf. Putnam (1981), chapter 1. Putnam uses the experiment to argue against metaphysical realism, the idea that the content of our beliefs is fully independent of the world.

considerations. One may grant that even though one's thought 'I am a brain in the vat' is false, this is merely so because we cannot express that we live in an illusion, since even the meaning of our words is manipulated. Putnam's solution therefore feels like a trick: we cannot consistently believe that we are a brain in vat, but the evil scientist may point to the vat and say 'That's a brain in the vat. He is deluded, and he cannot even consistently think that this is so, because his words derive their content from the illusion' (even though the vatted brain cannot express this).

Whether Putnam would agree that this is possible, does not concern us here. We think, however, that Davidson can add an additional element to his answer, to give it a more epistemological character: the need for another mind.²⁰ For Davidson we cannot speak of *the* cause of our beliefs (and hence not of our having any beliefs) unless we communicated with others on the basis of a common cause. Therefore, if we suppose for the moment that we could be a brain in the vat, there is no answer to the question what our beliefs are about unless this is determined intersubjectively: we must have communicated with others in order to determine what is the common cause of our beliefs. But this makes no sense for a brain the vat: even if there were something that could sensibly be *the* cause of our beliefs if we were a brain the vat, there is no other mind for the brain to communicate with. (The brain cannot communicate with the scientist: the scientist may interpret what the brain is 'thinking' about, but the brain cannot communicate with the scientist: the envatted brain may not even 'know' that there is a scientist, it is supposed to live in illusion created by the scientist.) If that is so, we can reject the scenario as inconsistent even without considering what we might say if we were a brain in the vat: we could not even have beliefs (or think meaningful words) in the first place.

One may object to this that the illusion of communication may also be induced by the evil scientist. The point is, however, that that presupposes that the brain in the vat could have beliefs, and hence could believe that he communicated with others. But if Davidson is right, he could not believe that in the first place if he did not *really* communicate with others. We could object here that this answer is as much of a trick as Putnam's: we could for example argue that it is a property of a brain that it normally belongs to someone who is able to learn to speak and acquire beliefs. If the scientist feeds the appropriate stimuli to the brain, it should be possible to create the *illusion* of beliefs inside the brain (otherwise it is inexplicable how embodied brains work, who may get exactly the same stimuli and acquire the capability of thought). This seems true, but then still Davidson's answer is radical (and more radical than Putnam's): according to Davidson we would

²⁰Davidson himself, however, does not do this. In fact, he writes—which is (arguably) inconsistent with our analysis—that he endorses the idea that the brain's thoughts are just about what typically causes them, cf. EE 201.

have to say that even though the brain may entertain the illusion of having beliefs (and of having communicated), those beliefs are not about anything, since the brain did not *really* communicate. Therefore, we do not—like Putnam does—have to consider the words we were to use if we were a brain in the vat. We can a priori exclude that those words mean anything. On the other hand, the Davidsonian position may be seen as worse than Putnam’s position, since whereas for Putnam we cannot be a brain in the vat and think that we are one (whence we can say that we are not a brain in the vat), Davidson’s position may be seen as implying that we could be a brain in the vat after all, but then we could not maintain that we have beliefs: those beliefs are devoid of content. Nevertheless, this would still be an answer to the sceptic: the sceptic cannot consistently maintain that he is a brain in the vat, since *if* he is, his thought (the *Cogito*) is an illusion. Indeed, we shall see below that Davidson thinks that the *Cogito* is an *empirical* assumption, which by implication eliminates sceptical scenarios. For Davidson therefore assuming that one thinks is assuming that one does not live in an illusion, since thought is not possible in an illusion. Adopting Davidson’s position, we can therefore point to a new paradox in the sceptical position, slightly deviating from the way Putnam approaches the sceptic.²¹

Nagel says he can only think of two alternatives that remain for the sceptic to avoid Davidson’s conclusions. One is adopting Platonism with regard to our concepts: the idea that we ‘grasp’ concepts independently of learning how to use them, without communication. In that case, it seems, we could have the notions required to entertain beliefs without ever having

²¹A way out for the sceptic may be to define a scenario in which there are in a sense *really* other minds in the illusion (like in *The Matrix* (L. and A. P. Wachowski, Warner Bros: 1999)) with the possessors of which we communicate. However, in as far as the persons in the illusion are *controlled* mentally by an evil scientist or something else in the world outside of the illusion, we cannot say that the causes of their beliefs are to be located in the illusion, whence it still seems to make no sense to say that their beliefs have content (which is required for communication within the illusion). (In *The Matrix* that seems to be so: the minds within the illusion are causally connected to (and influenced by) machines in the external world.) If the persons within the illusion are *not* causally controlled by something outside it, this seems to deflate the scenario: it would be like saying that the universe *as a whole* is an illusion controlled by someone (maybe by God), but nevertheless the beliefs we entertain are caused by things within the illusion and we truly communicate with other minds. This is far less extreme than being a brain in the vat, if it makes sense at all. (What does it mean to say that the universe we live in as a whole is an illusion, if this does not bear on content and the veridicality of our beliefs?)

Another way out for the sceptic may be to argue that our brain may be recently plugged into the scientist’s computer (after we were kidnapped by him and the brain was taken out of our skull, without our noticing it). We would then have been in the real world, communicating with other people, to acquire beliefs. However, the question then arises what our perceptual beliefs would be about. They are now induced by electrical impulses, i.e. by other things than they were caused by in the external world, and we have not objectively located those new causes by means of communication. It seems that we can therefore not have beliefs in such a case.

correctly classified things in the world. We could then consistently assume to be thinkers without assuming anything about the world. The other refuge for the sceptic, according to Nagel, would be to change the subject matter and to be sceptical that our beliefs have content at all (i.e. to doubt even the *Cogito*), precisely because assuming that they have content amounts to assuming one has many true beliefs about the world.²²

For Nagel the conclusion of Davidson's veridicality argument is intuitively too strong not to elicit a new form of scepticism:²³

[I]t is absolutely amazing that there should be an a priori argument which proves that *this* set of propositions, which I believe, covering vast tracks of history, natural science, and ordinary lore about the world, is largely true.

However, we think Nagel's characterization of Davidson's argument is a little too ambitious. Davidson himself concedes that the veridicality thesis (or the premisses it is based on) follow(s) partly from empirical assumptions (in as far as the empirical — a priori distinction is a sensible one for Davidson), indeed, from a view on 'the way we are'.²⁴ This suggests that, rather than a thesis with huge empirical implications, the veridicality thesis is a conceptual truth that we are committed to believe given our best ideas on the nature of belief. If this is so, scepticism cannot get off the ground.

5.5 Does Davidson leave room for the sceptic?

5.5.1 Two readings of the veridicality thesis

We think that what exactly Davidson's attitude towards the sceptic is, comes to light best in Barry Stroud's essay on Davidson's anti-scepticism and Davidson's reply to it.²⁵

Stroud applauds Davidson's rejection of scepticism. However, regarding the veridicality thesis he thinks that Davidson vacillates between a weaker and a stronger reading.

According to the weaker reading, belief ascription is globally veridical from the perspective of the ascriber and (hence) *if* we interpret someone, we cannot but conclude that most of her beliefs are true (and if someone interpreted us, she would conclude the same about our beliefs). It is impossible to reach the verdict (in particular not by an omniscient interpreter)

²²Cf. Nagel (1999), pp. 2014-205. We shall not go into the question whether a Platonist can maintain that he may be a brain in the vat, but at least a Platonist could maintain to be a solipsist.

²³Cf. Nagel (1999), p. 202.

²⁴Cf. RTG, p. 194.

²⁵Stroud (1999) and RTS.

that a person's beliefs are false in the main. According to the stronger reading, anyone's beliefs are *in fact* mostly true. In other words: not only do interpreter's necessarily conclude that the interpretee is largely right, but beliefs are *really* largely true.²⁶ Stroud accepts the veridicality thesis on the weaker reading, but rejects the stronger reading.

For Stroud the stronger reading cannot be established, and the weaker reading for him suffices to eliminate any serious threat from scepticism. It seems therefore that Stroud accepts Davidson's agreement thesis, and accepts Davidson's argument that the conditions of interpretation urge us to believe that others' (and our own beliefs) are largely true, but does not (fully) accept Davidson's argument 'from agreement to veridicality'. Stroud thinks that both of Davidson's attempts in *A Coherence Theory of Truth and Knowledge* to extend agreement to veridicality fail (and in general he thinks that, as a philosophical enterprise, it is mistaken to try to establish the stronger reading).

For Stroud the omniscient interpreter argument fails because even though it is right that such an omniscient interpreter would find our beliefs largely true *if she interpreted us*, whence they would in fact be largely true, we do not know if our beliefs are interpretable by an omniscient interpreter.²⁷

The other of Davidson's arguments simply relied on 'the complicated causal truth' that makes us the believers we are: roughly, our beliefs are typically caused by what they are about, and hence are typically true.²⁸ Stroud thinks that we have no compelling reason to think that our beliefs are caused by what makes them true, even if Davidson is right that interpretation relies on such causes. Specifically, he argues that it is not necessary for successful interpretation that the contents of our beliefs are determined by what *actually* typically causes them. Stroud agrees that we do interpret each other by identifying conditions that cause speakers to hold true their sentences. However, he thinks this does not require that the conditions as we identify them are *in fact* the causes of the beliefs. He gives one example of this. To correctly interpret a speaker's sentence as being about a rabbit, it is not required that the speaker's holding true the sentence is caused by the actual presence of a rabbit that is observed by speaker and interpreter: it is instead required that the speaker *and* the interpreter both *take* something to be a rabbit (possibly mistakenly) and take that to be the cause of each other's belief that there is a rabbit and of the speaker's holding true the related sentence. That would suffice for successful communication. Therefore, Stroud argues, communication does not necessarily arise where causes converge, but where causes are *taken* to converge and where they are

²⁶Cf. Stroud (1999), pp. 144-145.

²⁷Cf. Stroud (1999), section V. As we discussed in chapter 3, Genova takes this to be an implicit assumption, which Stroud thinks begs the question.

²⁸Cf. Stroud (1999), section VI. Cf. CTTK, p. 153.

identified as a common cause.²⁹

Nevertheless, Stroud thinks that the weaker reading suffices to establish that the possibilities proposed by the sceptic are inconceivable and hence can be rejected. Stroud writes: ‘an enquirer’s relation to the apparently innocent possibility from which a sceptical threat is thought eventually to arise is therefore parallel to a speaker’s relation to the possibility expressed in the paradoxical sentence ‘I believe that it is raining, and it is not raining.’³⁰ That sentence may be true, but cannot consistently be believed to be true. The same goes, according to Stroud, for the scenario’s sketched by sceptics. This means that what the sceptic asserts (e.g. that all our beliefs may be false) may be true (all our beliefs could in fact be false), but cannot consistently be *believed* to be true. Stroud writes: ‘There is a difference between something’s being simply inconsistent or impossible (which someone’s believing that it’s raining, and its not raining, is not) and something’s being impossible for anyone to believe or discover.’³¹ Davidson’s arguments for Stroud make us appreciate that the ‘sceptical possibility’ is inconceivable, but not that it is impossible. What Stroud has in mind seems to be the following: it is an inconsistent idea to think that a set of beliefs is completely mistaken and to think that those are someone’s (or our own) beliefs, since it is inconsistent to attribute a largely false set of beliefs. Nevertheless, that this is inconsistent to do for us (given the way belief ascription works) does not mean that there is no possibility that we entertain beliefs that are largely false: rather, knowing that a set of beliefs is largely false, we could not attribute that set: not only not to ourselves, also not to others, since it goes against the principles of belief attribution to attribute largely false beliefs. It seems that Stroud’s ‘unbelievable’ possibility could be actual in only two scenarios: either our beliefs are largely false and we have never communicated with someone, or our beliefs are largely false and we have communicated with others, but (by the agreement thesis) on the basis of shared largely false beliefs (but Stroud would then add that it is inconceivable to think that either of these is actual, since were we to think that, we could not anymore believe that we had those false beliefs.) For Davidson, however, neither of these is possible.

However, the point of scepticism is not that the sceptic believes that his beliefs are false (as Stroud’s rain sentence suggests), but rather that she thinks she does not have sufficient reason to think they are (mostly) true. The sceptic therefore has the meta-belief that all her beliefs *may* be false, or (if she takes that to be inconsistent) she suspends judgement and has no beliefs (except meta-beliefs).³² It seems therefore that Stroud’s reading of

²⁹We shall come back to this point in the next section.

³⁰Stroud (1999), p. 156. Cf. below for our comment on this comparison.

³¹Stroud (1999), p. 156.

³²In as far as philosophers are interested in scepticism and think there is no good argument against it, it seems that they still entertain their normal beliefs but also believe

the veridicality thesis does not suffice to prove the sceptic wrong: it merely implies that it is inconsistent to think that all one's beliefs are false, but the sceptic does not believe this (and one may argue that it is obviously inconsistent to think this, it does not need a sophisticated argument like Davidson's). Maybe what Stroud means is that the sceptic could not even consistently tell us that our beliefs may be largely false, because that would imply that the sceptic has identified our beliefs *not* as largely true, which would imply that he interpreted us wrongly. That would get close to what we think is what Davidson means by the thesis (see below), but it would still not explain why Descartes's position (or any traditional sceptical position) is inconsistent, since Descartes suspends judgement about all things, except the *Cogito*.³³

To support his reading of the thesis, Stroud appeals to Davidson's idea that a community of minds is the measure of all things. For Stroud this is the core insight of Davidson's epistemology: we cannot step outside the realm of beliefs and judge about them. Stroud construes scepticism as a negative verdict on a body of putative knowledge made from an impossible position outside it.³⁴ Stroud thinks Davidson should refrain from making a positive verdict from the same position, against the sceptic. Stroud therefore thinks we do not *know* that all or most of the things we believe are true.³⁵

The problem with Stroud's interpretation is that Davidson himself is clearly after a stronger answer to the sceptic than Stroud, as is clear from Davidson's reply to Stroud's essay. In the reply he argues that Stroud demands too much of knowledge and that we do know that the sceptic is false, something that Stroud explicitly denies. Stroud on the other hand overestimates what Davidson thinks is the power of the argument for veridicality. Davidson's reply to Stroud (and his disagreement) sheds light on how 'strong' the veridicality thesis for Davidson is and also what Davidson is after in arguing for the veridicality of belief.

5.5.2 Veridicality and the possibility of knowledge

What Davidson is after, according to his reply, is showing that we can *justifiably* believe that belief is veridical, and that we therefore should believe this. For Davidson it follows from our conception of belief that belief is veridical. To see why this is so, however, is not trivial, and requires a sophisticated argument. For Davidson the argument shows that we have

that in some way (that does not affect their life) they may be completely wrong and they do not *know* whether this is not so. Maybe an occurrent sceptical meta-belief would require a temporary suspension of one's first-order beliefs. However, we shall see that Davidson rejects such a position as inconsistent.

³³In a sense it is trivially so that if a sceptic says things to us she belies her own position, since communication relies on beliefs.

³⁴Cf. Stroud (1999), p. 157.

³⁵Cf. Stroud (1999), p. 157.

reason to believe that the world is pretty much as it is and in fact, contrary to Stroud's position, that we *know* this.³⁶

Sometimes Davidson suggests that the anti-sceptical argument is a fortunate consequence of his position.³⁷ However, in his reply to Stroud Davidson rather suggests that a proper epistemological theory *should* imply that the sceptic is mistaken. Davidson writes (concerning Stroud's view): 'if we do not know this [that the world is pretty much as we think it is], it seems pointless to say that we know *anything at all* about the world around us. It seems to me that any position that assumes, or leads to, this conclusion must be demanding more of knowledge than is justified.'³⁸ (Emphasis is ours.) For Davidson Stroud's position (that we do not *know* that our beliefs are largely true) would imply that we have no reason to think that we know anything at all about the world, and would therefore put too strong conditions on knowledge. For Davidson it is a precondition for an explanation of the possibility of knowledge that it offers an argument against scepticism. This is a controversial idea and there are (apparent) reason not to accept it.

One reason is that if we need to know that the sceptic is mistaken in order to have any knowledge at all, it would seem that people who never thought about scepticism, or people who do not accept any anti-sceptical argument, would be deprived of any knowledge whatsoever. This bizarre idea is presumably not what Davidson is after.³⁹

Therefore, it seems we should rather take Davidson to hold that a (or the) right epistemological theory should deliver an argument against the sceptic. If we have the right theory on the nature of belief and knowledge, we *know* that the sceptic is wrong (we have a justified belief and whatever more it takes to have knowledge). People who are not familiar with the theory could then be said to implicitly know this, or rather, we can explain why they know the world is pretty much as they believe it is, since we understand why the nature of knowledge is such that this is so. But why would a theory of 'the possibility of knowledge', as Davidson calls it, have to explain that the sceptic is mistaken, or confused? Davidson seems to think this is obvious, but it is in fact controversial. Stroud thinks philosophy cannot and should not try to explain this. Robert Nozick developed a theory of knowledge that is able to incorporate the sceptical possibility, which allows us to know all kinds of particular things without excluding that the sceptic is mistaken.⁴⁰ And one may argue that there is no contradiction in the idea of having many justified beliefs about the world (which consti-

³⁶Cf. RTS, pp. 162-164.

³⁷Cf. RTG, p. 192. Pritchard also suggests this (Pritchard (2013), pp. 528-530).

³⁸RTS, p. 162.

³⁹Davidson recognizes this difficulty: cf. RTN, p. 208.

⁴⁰Cf. Nozick (1981), part 3: Epistemology. Nozick proposes an analysis of knowledge in terms of counterfactuals: our beliefs that constitute knowledge must be preserved under appropriate counterfactual conditions.

tute knowledge), but nevertheless not having any justification (and hence no knowledge) that most of one's beliefs are true (even if one believes they are), since that seems to require more than the sum of the justifications of all of one's particular knowledge claims (which are ultimately all based on beliefs).⁴¹ This would mean that one can have a sceptical meta-belief: the belief that one's first-order beliefs, even though justified, may all be false, because having a justification does not guarantee that they are true. For Davidson, however, the question how first-order beliefs can be justified (and hence how knowledge about the world is possible) cannot be separated from the question how our body of beliefs as a whole is connected to the world and to other minds. The reason is that for Davidson the veridicality of belief explains why knowledge is possible. It is only because belief is globally veridical that justification of first-order beliefs by other beliefs makes sense and that acquiring knowledge is possible. For Davidson it is because beliefs have 'a presumption in favour of their truth', and that they are supported by many other beliefs that have a veridical presumption, that they can amount to knowledge.⁴² In fact, for perceptual beliefs we do not have to justify on independent grounds that we are not deceived in order to have perceptual knowledge: '[W]e can have perceptual knowledge, say that we are seeing a cow, without having to ascertain independently that we are not deceived in one way rather than another', '[W]e are *justified* in taking our perceptual beliefs to be true, even when they are not, and so when they are true, they constitute knowledge.' (Davidson's emphasis.)⁴³ Independent justification (i.e. a justification that we are not deceived), even if possible to some extent, cannot give a foundation on something other than (perceptual) beliefs. For Davidson, the veridicality of perceptual beliefs is essential to an explanation of the possibility of knowledge, and for someone to have perceptual knowledge he must just 'recognize situations in which he is justified' to believe what he sees.⁴⁴

⁴¹ Although there is a contradiction if one thinks knowledge is closed under implication (which Nozick indeed rejects), since any piece of knowledge, since it is veridical, implies that things are the case which are incompatible with what the sceptics says might be the case and hence implies that the sceptic is mistaken (for example, if I know that I am sitting on a chair and if knowledge is closed under implication, this means that I know that there is an external world with material objects (since a chair is a material object in the external world)). According to Nozick's theory, I can know that I am sitting on a chair (and that I have a body) without knowing that I am *not* a brain in the vat (where Nozick assumes that Putnam is wrong that our words in this scenario would have such a meaning as to be true). But my sitting on a chair implies that I am not a brain in the vat: indeed, Nozick rejects the idea that knowledge is closed under implication. However, a wish to avoid rejection of the closure of knowledge under implication is not (directly) Davidson's reason for requiring that a theory of knowledge be anti-sceptical, see below.

⁴² Cf. CTTK, p. 153.

⁴³ EE, p. 199 and Reply to Nagel p. 208 respectively.

⁴⁴ EE, p. 201. In fact, it is vague what Davidson means here by 'recognizing situations in which one is justified to believe that one sees something'. The sceptic may ask here

Anyone who accepts perceptual externalism [the idea that perceptual beliefs are about what (typically) causes them, GWE] knows he cannot be systematically deceived about whether there are such things as cows, people, water, stars, and chewing gum. Knowing that this is the case, he must recognize situations in which he is justified in believing he is seeing water or a cow. In those cases where he is right, he knows he is seeing water or a cow.

Davidson therefore argues that in as far as we take knowledge to depend on perception, we do not have to look for an ultimate and certain foundation, as the sceptic demands. It suffices for Davidson that we can justify our knowledge claims on the basis of perceptual beliefs, and accept that perceptual beliefs are by nature veridical.

Of course, this still leaves many potential targets for the sceptic: the reliability of indirectly acquired knowledge (based on sources), the problem of induction, etc. (These are problems that Davidson is not concerned with when arguing that belief is veridical.) Davidson can therefore certainly not eliminate the sceptic's demand for a proper justification on the basis of these considerations. However, in as far as the sceptic demands a justification for our perceptual knowledge claims that gives certainty, Davidson rejects the idea that that is needed to explain knowledge (nor is it possible to give such a justification).

Davidson therefore thinks that if belief were not veridical by nature, any form of justification would be senseless. For a foundationalist the question of justification and the question of global veridicality can be separated: justification could be construed for a foundationalist in terms of a relationship of beliefs to what constitutes the foundation, whereas globally a foundationalist can be a sceptic, since justification fails to guarantee veridicality. For Davidson it is exactly because belief is *globally* veridical that justification of particular beliefs is grounded and that knowledge claims are rational.⁴⁵

For Davidson therefore the fact that we know many things about the world implies that we know that the world is as we believe it is. For Stroud

how she could recognize such a situation. Doesn't that require a justification after all? Nevertheless, Davidson's position implies that we cannot be systematically or globally mistaken, so even if we cannot precisely specify situations in which we are justified to entertain a perceptual belief, we cannot be wrong globally.

⁴⁵That Davidson's position implies that we have not merely true beliefs but knowledge about the world is obvious in another way (which is just the linguistic variant of the argument we have already discussed). It is plausible that we must *know* what our sentences mean in order to use them, particularly what our perceptual sentences mean. But if they are about what typically causes us to hold them true, then we have knowledge that the truth conditions of those sentences at least sometimes obtain, which means that we have knowledge that we live in a world roughly like the one we talk about (if we deny this, we cannot consistently maintain that our sentences have meaning).

this is not so: we could know many things, but we do not *know* that most of our beliefs are true.

5.5.3 Assuming that one thinks

For Davidson the veridicality thesis is primarily the cornerstone of an epistemological theory, based on the best intuitions we have about what beliefs and knowledge are, rather than a fortunate argument particularly devised to beat the sceptic on her own terms. For Davidson it does not make sense to prove the sceptic wrong *in addition*.⁴⁶

The argument for the veridicality of belief depends on premisses that are empirical in the sense that they depend on a view on ‘the way we are’ for Davidson.⁴⁷ One of the premisses is Davidson’s externalism:⁴⁸

What I propose is a modest form of externalism. If our past—the causal processes that gave our words and thoughts the content they have—had been different, those contents would have been different, even if our present state happens to be what it would have been had that past been different. [...] my case against scepticism depends on it.

Another important premiss is the idea that thought depends on communication. Together these two premisses combine into Davidson’s ‘triangular externalism’. For Davidson, when considering what belief is, we cannot make sense of our having beliefs without assuming that we communicate with others in a shared world.

Davidson does not claim that we know this with the certainty demanded by the sceptic: ‘It is only after the question [explaining, philosophically, how knowledge is possible] has been answered that it is wrong to keep asking, but how do I know *that*?’⁴⁹ This is an interesting position on scepticism. Davidson does of course not mean that we are not allowed to question the validity of his triangular externalism (his premiss), but rather that he would not accept that if we do that on purely sceptical considerations. This is the difference between Davidson and the sceptic: Davidson’s concern is to bring to light our best beliefs about the nature of belief and knowledge, and (not surprisingly, since it had to be so for Davidson) he finds an argument why

⁴⁶Cf. RTS, p. 163.

⁴⁷Cf. RTG, p. 194.

⁴⁸RTS, pp. 165

⁴⁹Cf. Avramides (2013), pp. 558-559 and RTS, p. 163. Davidson also writes that we do not know that we know the sceptic is false. This seems strange: if I know that belief is veridical, doesn’t that imply that I know that I know that belief is veridical? Or does it merely imply that I believe that I know that believe is veridical? We shall not go into this, since it is not clear what exactly Davidson means here, but his general argument, as presented in the main text, is clear.

belief is veridical by nature. He does not investigate the nature of belief as a sceptic: he does not suspend judgement.

The strength of Davidson's veridicality thesis is somewhere in between Stroud's weaker reading and his stronger reading (even though Davidson himself sometimes pretends that the stronger reading is right).

The weaker reading, or Stroud's reading, is too weak because it does not fully capture the spirit of Davidson's argument. Stroud for example argues that we may be mistaken about what causes our beliefs and that the only thing that matters is what we take to be the causes. However, this cannot be so *systematically*.⁵⁰ To have beliefs about the world at all, we must have located, by means of communication, the features in the world those beliefs are about. This requires both that these features really exist in typical cases and that the communicators intend each other to be speaking about those really existing features, otherwise there could be no content to their beliefs. (If we assume (by contradiction) that the communicators systematically communicate on the basis of a shared mistaken common cause, it seems that we would either have to say that their communication is just harmonious empty babbling (without meaning) that is not about anything, or (more plausibly) that they are not mistaken at all: if they both systematically call a rabbit 'hare', then 'hare' simply means rabbit for them.)

Stroud is nevertheless right in a sense: Davidson rather proves that scepticism is incoherent, given a certain conception of belief, than that it is false. Davidson himself agreed with Rorty that he tells the sceptic to get lost, instead of refuting her.⁵¹ If we thought of Davidson's argument for the veridicality of belief as *refuting* the sceptic, we would have to assume that Davidson and the sceptic share their notion of belief and operate on common terms. This is not so. Davidson's position implies that the sceptic's idea is incoherent.

There is something paradoxical about this position. How could we possibly have knowledge that our beliefs are mostly true, if the only thing we could such knowledge on is beliefs?⁵² Belief in the veridicality thesis, and the premisses it is based on, is after all itself a belief (or knowledge) based on empirical premisses. It seems that Davidson must acknowledge that there is a possibility that the veridicality thesis is wrong.

Davidson probably admits this by conceding that we do not have a guarantee (we do not know that we know) that the thesis is true.⁵³ But the same goes for *any* kind of knowledge for Davidson. However, this does not mean that there is room for the sceptic after all *within* Davidson's epistemological framework: if we are Davidsonians about epistemology, we are committed

⁵⁰Cf. RTS, pp. 164-165.

⁵¹Cf. AF, p. 157, RTS, p. 163, and Rorty (1986).

⁵²Especially as a foundational principle for a philosophical theory of knowledge (as discussed above) it may seem to presuppose what it is supposed to prove.

⁵³Cf. RTS, p. 163.

to say that scepticism is incoherent. In order to consistently formulate a sceptical scenario that goes against the veridicality thesis, we have to set aside Davidson's epistemology. This is presumably what Stroud aims at when writing that we do not know that the sceptic is wrong: we may be mistaken about what belief is. Davidson thinks we do not have to exclude that possibility, since we are justified in having our conception of belief: it is based on our best intuitions and on philosophical reasoning. Nevertheless, the sceptic may insist that Davidson's epistemological views may be wrong. A Davidsonian may grant this, but nevertheless insist on her turn that this is inconceivable for her as long as the sceptic does not explain *why* she is (or may be) wrong. As long as the sceptic has no convincing argument (or can merely appeal to epistemological frameworks more friendly to scepticism but unacceptable for a Davidsonian), scepticism (the idea that all our beliefs may be false) cannot be understood by the Davidsonian. The sceptic should first convince the Davidsonian to change her notion of belief.

Nagel argues that we may accept Davidson's argument and be sceptics regarding content instead. Davidson's veridicality thesis, as we saw, arises from considerations on what is needed for beliefs to have objective truth conditions, and hence content.⁵⁴ The sceptic may concede that it is inexplicable how we could think without our having an adequate picture of the world, but she could then proceed and argue that this is sufficient reason to doubt what is the most fundamental premiss of Davidson's argument: the *Cogito*, the idea that our thought is meaningful and objective, which is presupposed by Davidson's account of what is needed for this to be so. Indeed, Davidson acknowledges that the assumption that one has thought is an empirical premiss, and 'not a small one', as he puts it.⁵⁵ The very point of Davidson's position is that having thought implies having empirical knowledge, which in a sense is a new application of the rejection of the analytic/synthetic distinction: the assumption of thought (and of having concepts) has immediate empirical implications. In fact, content scepticism seems not to be so different from saying that we may be a brain in the vat (which seems to mean that we live in an illusion, and in that case (we might say) our thinking would also be an illusion). However, the idea that our thought lacks any determinate content seems incoherent, and even if we entertain this idea, we cannot without contradiction express it. The sceptic that Davidson is interested in is the one that assumes the *Cogito*, i.e. assumes that her thought has content, but thinks her picture of the world may be fundamentally mistaken.

Nevertheless, even if we take Davidson's argument to be a valid philosophical argument, it may fail to be compelling psychologically. Someone may be convinced that everything, including his thinking, is an illusion. As

⁵⁴Cf. Nagel (1999), p. 205.

⁵⁵Cf. RTN, p. 209.

long as this person is not planning to develop a proper philosophical argument on the basis of that conviction (as suggested by Nagel), we would advise him to visit the psychologist.

Conclusions and suggestions for further research

For Davidson belief is veridical because the conditions that typically cause us to entertain a belief are the conditions under which the belief is true. For Davidson the veridicality of belief explains why we can have knowledge.

Davidson elaborated this causal theory of content by the idea of ‘triangulation’: only linguistic communication can settle what the causes of our beliefs are. We have seen that Davidson tries to understand truth and belief from a third person perspective, by looking at the role these notions play in a theory of rational behaviour. For Davidson this perspective is justified by the idea that the primary application of these notions is in an explanation of linguistic communication. Davidson writes in *Thought and Talk* that belief as a private attitude is unintelligible ‘except as an adjustment to the public norm provided by language’.¹ For Davidson this gives rise to the idea that three kinds of knowledge (knowledge of one’s own mind, knowledge of another mind, and knowledge of the world) are interdependent.

These considerations enable Davidson to explain his position on truth: truth is verification-transcendent and belief is constrained by what is true. These considerations also seem to imply that certain kinds of scepticism are unintelligible, but we also saw that Davidson’s position on scepticism is subtle and not easy to reconstruct.

Davidson has offered an interesting account of meaning, truth, belief, and the possibility of knowledge. However, there are various issues that need to be subject to further scrutiny in order to achieve a comprehensive assessment of Davidson’s thought on these central concepts in philosophy. We shall mention four of them.

As a first point, Davidson has not defined a very precise theory of perceptual knowledge (in *Epistemology Externalized* Davidson merely suggests that he agrees with Tyler Burge on this issue), and further work could be done here, since it is quite crucial for an upgrade of the veridicality thesis to a theory of knowledge, and for an adequate answer to the sceptic. What is particularly needed is an account of why a belief that is *not* caused by what

¹TT, p. 170

typically causes beliefs of a similar kind is still a belief of the same kind as those that are caused by the typical cause.²

A second issue that needs further exploration is the relationship between Davidson's holistic externalism and on the one hand the causal externalism proposed by Hilary Putnam and on the other hand the social externalism proposed by Tyler Burge.

A third point concerns Davidson's truth-conditional understanding of meaning. It would be an interesting question to what extent Davidson's account of truth depends on his truth-conditional analysis of meaning. It seems that for Davidson truth is so important in a theory of rationality because of its role in the understanding of meaning and hence its role in communication. As we saw, Davidson thinks of the concept of truth (together with the concept of belief) as the most basic concept(s), without which we could not have any other concept. However, if Davidson's programme in truth-conditional semantics fails, does this mean that Davidson was wrong about the primacy of the concept of truth? Or can the question of the general philosophical importance of truth be separated from the question of the importance of truth in a (formal) analysis of meaning for Davidson (and the success of such an analysis)? Davidson wants to appeal to truth in an analysis of meaning because he thinks truth is a clearer concept than meaning, which seems to suggest that truth is somehow more basic. However, we also saw that for Davidson having the concept of truth depends on communicating propositional content (and hence on understanding meaning) and communicating propositional content depends on having the concept of truth, which means that truth is not more basic than meaning but that the notions are interdependent and can only 'arise' simultaneously. Truth therefore seems merely an 'easier' concept than meaning, which for Davidson is suitable as a tool for analysis of natural language, since what Davidson thinks is distinctive of the unity of sentences is that they have a truth value and truth conditions (and what makes natural language obscure, is that the logical structure is often hidden and not easy to reconstruct, which seems to be the reason that meaning is obscure for Davidson). It seems to us that what is crucial to Davidson's position is his idea that our pre-theoretical understanding of truth is manifested in the interpretation of others. Even if Davidson's semantic programme fails—for example because not all phenomena in natural language can be captured by a truth theory—a Davidsonian could still maintain that truth plays a crucial role in interpretation (and more in general in making sense of rational behaviour), because of the role of belief and an understanding of the difference between what is true and what is held true in the interpretive process (and in the theory of rational behaviour). However, further clarification of this issue is needed (and is probably quite complicated).

²Cf. EE, p. 199-120.

A fourth example of a subject that could be explored in more detail is Davidson's argument against the idea of a solitary thinker. Especially in *The Second Person*, Davidson suggests that his argument against a private language is based on considerations similar to the ones discussed by Wittgenstein and by Kripke in his book on Wittgenstein. However, there are at least two differences between Wittgenstein's private language argument and Davidson's 'triangulation argument'. The first is that Davidson focusses primarily on belief rather than on language: for Davidson not merely a private language, but also a solitary believer is unintelligible (but this may be merely a difference in approach to the same issue, although we are not sure about that). The second difference is in Davidson's conception of meaning, which, as opposed to (the later) Wittgenstein's, is systematic, homogeneous, referential, and individualistic. Davidson advocates a strict separation of meaning and use and he construes normativity in linguistic communication not in terms of shared practices or conventions (which he thinks play no essential role in linguistic communication) but in terms of the intention to be interpretable. Despite these significant differences (Davidson's views on language are in a sense a denial of what for Wittgenstein explains the normativity of language), Davidson proposes an argument against private language (which for Davidson means: an uninterpreted language) based on considerations similar to Wittgenstein's (or at least, that seems so). It could be further explored how Davidson's ideas here relate to those of Wittgenstein.

Works by Donald Davidson, with abbreviations

Below the reader can find a list of the publications by Davidson that have been used for this thesis. References to works other than by Davidson can be found in the general bibliography. The abbreviations as specified at the left side are used to refer to them in the footnotes. When referring to page numbers of the publications, we refer to the page numbers of the publications as collected in the following volumes (*not* to the page numbers of the original publications). The essays in the collections sometimes contain slight modifications, like the addition of footnotes:

Davidson, D. (2001a). *Essays on Actions and Events*. Oxford: Oxford University Press.

Davidson, D. (2001b). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.

Davidson, D. (2001c). *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.

Davidson, D. (2004). *Problems of Rationality*. Oxford: Oxford University Press.

Davidson, D. (2005). *Truth, Language, and History*. Oxford: Oxford University Press.

In the list below we mention in between brackets the page numbers of the publications as they appear in the four volumes of collected essays.

- TM Truth and Meaning. *Synthese* 17 (1967): pp. 304-323. (2001b: pp. 17-36)
- CR Causal Relations. *Journal of Philosophy* 64 (1967): pp. 691-703. (2001a: pp. 150-163)
- ST On Saying That. *Synthese* 19 (1968-1969): pp. 130-146. (2001b: pp. 93-108)

- TF True to the Facts. *Journal of Philosophy* 66 (1969): pp. 748-764. (2001b: 43-54)
- RI Radical Interpretation. *Dialectica* 27 (1969): pp. 748-764. (2001b: pp. 125-140)
- DCT In Defence of Convention T. In: H. Leblanc (ed.), 1973, *Truth, Syntax and Modality*. Amsterdam: North-Holland, pp. 76-86. (2001b: pp. 65-76)
- BBM Belief and the Basis of Meaning. *Synthese* 27 (1974): pp. 309-323. (2001b: pp. 141-154)
- RLQ Replies to David Lewis and W.V. Quine. *Synthese* 27 (1974): pp. 345-349. (2001b: pp. 171-179)
- VICS On The Very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association* 47 (1974): pp. 5-20. (2001b: pp. 183-198)
- TT Thought and Talk. In: J. Guttenplan (ed.), 1975, *Mind and Language*. Oxford: Oxford University Press, pp. 7-23. (2001b: pp. 155-170)
- RF Reply to Foster. In: G. Evans and J. McDowell (eds.), 1976, *Truth and Meaning*. Oxford: Oxford University Press, pp. 33-41. (2001b: pp. 171-179)
- MTM The Method of Truth in Metaphysics. In: P.A. French, T.E. Uehling, and H. Wettstein (eds.), 1977, *Midwest Studies in Philosophy, II: Studies in Metaphysics*. Minneapolis, Minnesota: University of Minnesota Press, pp. 244-254. (2001b: pp. 199-214)
- IR The Inscrutability of Reference. *The Southwestern Journal of Philosophy* 10 (1979): pp. 7-19. (2001b: pp. 227-241)
- MP Moods and Performances. In: A. Margalit (ed.), 1979, *Meaning and Use*. Dordrecht, The Netherlands: D. Reidel Publishing Company, pp. 441-458. (2001b: pp. 109-121)
- RA Rational Animals. *Dialectica* 36 (1982): pp. 317-327. (2001c: pp. 95-106)
- CTTK A Coherence Theory of Truth and Knowledge. In: D. Henrich (ed.), 1983, *Kant oder Hegel?*. Stuttgart: Klett-Cotta, pp. 428-438. (2001c: pp. 137-153)

- NDE A Nice Derangement of Epitaphs. In: E. Lepore (ed.), 1986, *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 433-446. (2005: pp. 89-108)
- KOOM Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association* 1987: pp. 441-458. (2001c: pp. 15-38)
- AF Afterthoughts, 1987. In: A. Malachowski (ed.), 1990, *Reading Rorty*. Cambridge: Blackwell, pp. 134-138. (2001c: pp. 154-158)
- EPT Epistemology and Truth. In: *Proceedings of the Fourth Panamerican Philosophy Conference*. Cordoba: National University of Cordoba, 1988. (2001c: pp. 177-192)
- MTE Meaning, Truth and Evidence. In: R. Barret and R. Gibson (eds.), 1990, *Perspectives on Quine*. Oxford: Blackwell. (2005: pp. 47-62)
- SP The Second Person. *Midwest Studies in Philosophy* 17 (1992): pp. 255-267. (2001c: pp. 107-122)¹
- EE Epistemology Externalized. *Dialectica* 45 2-3 (1991): pp. 191-202. (2001c: pp. 193-204)
- RII Radical Interpretation Interpreted. *Philosophical Perspectives* 8 (1994): pp. 121-128.
- PO The Problem of Objectivity. *Tijdschrift voor Filosofie* 1995: pp. 203-220. (2004: pp. 3-18)
- PCT Pursuit of the Concept of Truth. In: P. Leonardi and M. Santambrogio (ed.), 1995, *On Quine*. Cambridge: Cambridge University Press. (2005: pp. 63-80)
- TVK Three Varieties of Knowledge. In: A.P. Griffiths (ed.), 1997, *A.J. Ayer: Memorial Essays, Royal Institute of Philosophy Supplement*. New York: Cambridge University Press, pp. 153-166. (2001c: pp. 205-220)
- IA Indeterminism and Antirealism. In: C.B. Kulp (ed.), 1997, *Realism/Antirealism and Epistemology*. Lanham, Maryland: Rowman & Littlefield, pp. 441-458. (2001c: pp. 69-84)

¹The last four pages in The Second Person as reprinted in Subjective, Intersubjective, Objective were taken from a paper called 'The Conditions of Thought', published in *Le Cahier du Collège International de Philosophie* 7 (1989): pp. 165-171.

- FTDT The Folly of Trying to Define Truth. *Journal of Philosophy* 94 (1997): pp. 263-278. (2005: pp. 19-38)
- STL Seeing Through Language. In: J.M. Preston (ed.), 1997, *Thought and Language*. Cambridge: Cambridge University Press, pp. 15-27. (2005: pp. 127-142)
- RTG Reply to A.C. Genova. In: L.E. Hahn (ed.), 1999, *The Philosophy of Donald Davidson*. Chicago: Open Court, pp. 192-194.
- RTN Reply to Thomas Nagel. In: L.E. Hahn (ed.), 1999, *The Philosophy of Donald Davidson*. Chicago: Open Court, pp. 207-210.
- RTS Reply to Barry Stroud. In: L.E. Hahn (ed.), 1999, *The Philosophy of Donald Davidson*. Chicago: Open Court, pp. 162-166.
- ET The Emergence of Thought. *Erkenntniss* 51 (1999): pp. 7-17. (2001c: pp. 123-134)
- TR Truth Rehabilitated, in: R.B. Brandom (ed.), 2000, *Rorty and his Critics*. Oxford: Blackwell, pp. 65-74. (2005: pp. 3-18)
- TP *Truth and Predication*. Harvard: Harvard University Press, 2005.

Bibliography

- Avramides, A. (2013). Knowledge of Other Minds in Davidson's Philosophy. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 550-564.
- Bennett, J. (1985). Critical notice of Davidson's *Inquiries into Truth and Interpretation*. *Mind* 94: pp. 601-626.
- Bernecker, S. (2013). Triangular Externalism. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 443-455.
- Carpenter, A.N. (2003). Davidson's transcendental argumentation: externalism, interpretation, and the veridicality of belief. In: J. Malpas (ed.), *From Kant to Davidson: Philosophy and the Idea of the Transcendental*. London: Routledge, pp. 219-237.
- Child, W. (2001). Davidson, realism and natural kinds. *Dialectica* 55 (1): pp. 29-50.
- Descartes, R. (1644, with additional material from the French translation of 1647). *Principia philosophiae (Principles of Philosophy)*. Translation with explanatory notes by V. Rodger and R. P. Miller. Dordrecht: Reidel Publishing Company, 1983.
- Dummett, M. (1959). Truth. *Proceedings of the Aristotelian Society* Supplementary Volume LIX: pp. 141-162. Reprinted in: *Proceedings of the Aristotelian Society* Virtual Issue 1 (2013): pp. 82-98.²
- Fodor, J. and Lepore, E. (1994), Is Radical Interpretation Possible?, *Philosophical Perspectives* 8: pp. 101-119.
- Genova, A.C. (1992). The Very Idea of Massive Truth. In: L.E. Hahn (ed.), *The Philosophy of Donald Davidson*. Chicago: Open Court, pp. 167-191.

²We refer to the page numbers as in the issue from 2013.

- Glanzberg, M. (2013). The Concept of Truth. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 156-172.
- Glock, H.J. (1996). On safari with Wittgenstein, Quine and Davidson. In: R.L. Arrington and H.J. Glock (eds.), *Wittgenstein and Quine*. Oxford: Wiley Blackwell, pp. 144-172.
- Glüer, K. (2007). Critical Notice: Donald Davidson's *Collected Essays*. *Dialectica* 61: pp. 275-284.
- Goldberg, N. (2008). Tension within Triangulation. *The Southern Journal of Philosophy* 46. No. 3: pp. 363-383.
- Goldberg, M. (2012). Swampman, response-dependence, and learning. In: G. Preyer (ed.), *Davidson's Philosophy: Truth, Meaning, and the Mental*. Oxford: Oxford University Press, pp. 148-163.
- Hacker, P.M.S. (1998), Davidson on Intentionality and Externalism. *Philosophy* 73 (286): pp. 539-552.
- Hahn, (2003), When Swampmen Get Arthritis. "Externalism" in Burge and Davidson. In: M. Hahn and B. Ramberg (eds.) (2003), *Reflections and Replies. Essays on the Philosophy of Tyler Burge*. Oxford: Clarendon Press, pp. 29-58.
- Joseph, M.A. (2004). *Donald Davidson*. Montreal, Quebec and Kingston, Ontario: McGill-Queen's Press.
- Klein, P. (2014). Skepticism. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2014 Edition)* (<http://plato.stanford.edu/archives/sum2014/entries/skepticism/>).
- Künne, W. (2003). *Conceptions of Truth*. Oxford: Clarendon Press.
- Lepore, E. and Ludwig, K. (2005), *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Oxford University Press.
- (2013). Truth in the Theory of Meaning. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 175-190.
- Lycan, W.G. (2013). Davidson and "The Method of Truth" in Metaphysics. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 141-155.
- Malpas, J.E. (1992). *Donald Davidson and the Mirror of Meaning*. Cambridge: Cambridge University Press.

- McGrath, M. (2014). Propositions. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2014 Edition)* (<http://plato.stanford.edu/archives/spr2014/entries/propositions/>).
- Nagel, T. (1999). Davidson's New *Cogito*. In: L.E. Hahn (ed.), *The Philosophy of Donald Davidson*. Chicago: Open Court, pp. 195-206.
- Neale, A. (2001). *Facing Facts*. Oxford: Clarendon Press
- Nozick, R. (1981). *Philosophical Explanations*. Cambridge, Massachusetts.: Harvard University Press.
- Okrent, M. (2011). Davidson, Heidegger, and Truth. In: J. Malpas (ed.). *Dialogues with Davidson: Acting, Interpreting, Understanding*. Massachusetts: MIT Press, pp. 87-111.
- Pagin, P. (2013). Radical Interpretation and the Principle of Charity. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 225-246.
- Plato. *Dialogues Vol. 4: Parmenides. Theaetetus. Sophist. Statesman. Philebus*. Translated by B. Jowett. Oxford: Oxford Clarendon Press, 1892.
- Putnam, H. (1973). Meaning and Reference. *Journal of Philosophy* 70: pp. 699-711.
- (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.
- Pritchard, M. (2013). Davidson and Radical Scepticism. In: E. Lepore, and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 521-532.
- Ramsey, F.P. (1950). Facts and Propositions. *Proceedings of the Aristotelian Society (Supplementary)* 7: pp. 153-170.
- Rorty, R. (1986). Pragmatism, Davidson, and Truth. In: E. Lepore (ed.). *Truth and Interpretation*. New-York: Blackwell, pp. 333-355.
- Russell, B. (1912). *The Problems of Philosophy*. London: Williams and Norgate; New York: Henry Holt and Company.
- Schantz, R. (1993). Davidson on Truth. In: R. Stoecker (ed.). *Reflecting Davidson. Donald Davidson Responding to an International Forum of Philosophers*. Hawthorne: de Gruyter, pp. 25-35.
- Speaks, J. (2013). Predication. In: E. Lepore and K. Ludwig (eds.). *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 328-338.

- Stoljar, D. and Damnjanovic, N. (2014). The Deflationary Theory of Truth. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2014 Edition)* (<http://plato.stanford.edu/archives/fall2014/entries/truth-deflationary/>).
- Stoutland, F. (1982). Realism and Anti-Realism in Davidson's Philosophy of Language I. *Critica* 14 (41): pp. 13-53.
- Stoutland, F. (1982). Realism and Anti-Realism in Davidson's Philosophy of Language II. *Critica* 14 (42): pp. 19-48.
- Stroud, B. (1999). Radical Interpretation and Philosophical Scepticism. In: L.E. Hahn (ed.), *The Philosophy of Donald Davidson*. Chicago: Open Court, pp. 139-161.
- Tarski, A. (1944). The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research* 4: pp. 341-376.
- Verheggen, C. (2013). Triangulation. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 456-471.
- Williams, J.R.G. (2013). Reference. In: E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, pp. 264-286.
- Quine, W.V.O. (1960). *Word and Object*. Cambridge, Massachusetts: MIT Press, pp. 26-79.