Kripke's Wittgenstein's Sceptical Solution and Donald Davidson's Philosophy of Language

Ali Hossein Khani

a thesis submitted for the degree of

Doctor of Philosophy

at the University of Otago, Dunedin,

New Zealand.

October 2016

To Saeedeh

Abstract

This thesis is an attempt to investigate the relation between the views of Wittgenstein as presented by Kripke (Kripke's Wittgenstein) and Donald Davidson on meaning and linguistic understanding. Kripke's Wittgenstein, via his sceptical argument, argues that there is no fact about which rule a speaker is following in using a linguistic expression. Now, if one urges that meaning something by a word is essentially a matter of following one rule rather than another, the sceptical argument leads to the radical sceptical conclusion that there is no such thing as meaning anything by any word. According to the solution Kripke's Wittgenstein proposes, we must instead concentrate on the ordinary practice of meaning-attribution, that is, on the conditions under which we can justifiably ascribe meaning to each other and the utility such a practice has in our life. Davidson has also argued that following rules is neither necessary nor sufficient for explaining success in the practice of meaning something by an utterance. According to his alternative view of meaning, a speaker's success in this practice is fundamentally a matter of his utterance being successfully interpreted by an interpreter in the way the speaker intended. On the basis of these remarks, Davidson raises objections to Kripke's Wittgenstein's sceptical argument and solution.

In this thesis, I will argue that Davidson has failed to fully grasp the essentially sceptical nature of the argument and solution proposed by Kripke's Wittgenstein. I will argue that as a result of this Davidson's objections and his alternative solution to Kripke's Wittgenstein's sceptical argument are mistaken. These criticisms are pursued via an investigation of Davidson's problematic reading of Quine's sceptical arguments for the thesis of the indeterminacy of translation. Having criticized Davidson's actual response to Kripke's Wittgenstein, I will claim that Davidson's best option for resisting the sceptical problem is to adopt a form of non-reductionism about meaning. Claudine Verheggen's recent claim that Davidson's use of the notion of triangulation will help to establish non-reductionism will be argued to be a failure. I will urge that the main obstacle in defending a non-reductionist view is the problem of accounting for the nature of self-knowledge of meaning and understanding. After discussing Davidson's account of self-knowledge and Crispin Wright's objection to this account, I will argue

that, although Wright's objection is ultimately unsuccessful, Davidson's account fails for other reasons. Finally, I tentatively suggest that the resources for an alternative response to the sceptical problem can possibly be extracted from Davidson's account of intending, which has some features suggestive of a judgement-dependent account of meaning and intention.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my primary supervisor, Prof. Alex Miller, for the continuous support of my Ph.D study, for his encouragement, patience, and motivation throughout my research project. Without his precious guidance and counsel, this thesis would not have been possible. I would also like to thank my secondary supervisor Assoc. Prof. Greg Dawes for his advice and helpful comments on the drafts of the thesis.

I am grateful to Prof. Penelope Mackie for her insightful feedback on the thesis. For their helpful discussions, I would also like to thank Dr. Kirk Michaelian, Prof. Rod Girle, and Prof. Max Cresswell. I wish to express my appreciation to my friends who have contributed and supported me by their stimulating discussions, especially Daniel Wee, Takahiro Yamada, and Ali Kalantari.

I owe a special thank you to my family, to my mother, for her constant patience, love, and care, and to my brothers and sister for supporting me in whatever way they could during my study. I especially thank my brother, Hassan, for always being there to support me in every difficult decision I have made.

Most importantly, I would like to express my appreciation to my wife and best friend, Saeedeh Shahmir, who has always been a constant source of strength and inspiration for me. Without her continuous encouragement and support, I would not have been able to complete this thesis at all.

I am also grateful to the University of Otago for financially supporting this thesis by granting me a University of Otago Doctoral Scholarship.

Contents

Introduction 1

1. Chapter 1: Kripke's Wittgenstein 9

Introduction 9

1.1. The Negative Part: Kripke's Wittgenstein's Sceptic's Sceptical Argument ... 10

1.1.1. The Sceptical Challenge 12

1.1.2. The Sceptical Challenge and Traditional Realism 14

1.2. The Candidate Facts 16

1.2.1. Previous Behaviour 16

1.2.2. Internalized Instructions 17

1.2.3. Mathematical Laws 18

1.2.4. Mental Images 18

1.2.5. Dispositions 20

1.2.6. Machines Embodying Intentions 23

1.2.7. Simplicity 24

1.2.8. Meaning as a Unique, Irreducible Qualitative Experience 27

1.2.9. Meaning Something as a Primitive Fact 28

1.2.10. Fregean Senses 30

1.3. The Sceptical Conclusions 31

1.4. KW's Sceptic's Sceptical Argument 32

1.4.1. The First Strand of the Sceptical Argument 33

1.4.2. The Second Strand of the Sceptical Argument 34

1.5. The Positive Part: Kripke's Wittgenstein's Sceptical Solution 36

1.5.1. Straight Solution vs. Sceptical Solution 36

1.5.2. The Sceptical Solution 39

1.5.3. Assertability Conditions 40

1.5.4. The Role of Agreement and the Private Language 42

1.5.5. The Role and Utility of Our Linguistic Practices 46

1.6. Conclusion 48

2. Chapter 2: The Later Davidson 50

Introduction 50

2.1. The Negative Part 51

2.1.1. The Common View 52

2.1.2. Davidson's Argument against the Common View 53

(a) Criticizing the First Version of the Common View 54

(b) Criticizing the Second Version of the Common View 56

(c) Criticizing the Third Version of the Common View 58

2.1.3. The Argument from Malapropism 59

(a) First Meaning vs. Conventional Meaning 60

(b) The Common View Re-Characterized 61

(c) The Phenomenon of Malapropism 63

2.1.4. Prior and Passing Theories 66

2.1.5. Passing Theories as Systematic Theories 71

2.1.6. "There Is No Language" Claim 75

2.1.7. Linguistic Error 78

2.1.8. Davidson's General Argument against the Common View 79

2.2. The Positive Part: Davidson's Alternative View 80

2.2.1. Davidson's Alternative View of Meaning 80

(a) Intention and Mutual Interpretation 82

(b) First, Conventional, and Speaker Meaning, and Ulterior Purposes 82

(c) Davidson's Interpersonal View 84

2.2.2. Davidson's Responses to Objections 87

(a) Being Obliged to Semantic Norms 87

(b) Practical Possibility of Communication without Convention 89

(c) The Wittgensteinian Seems Right/Is Right Distinction 89

2.2.3. Davidson's Criterion for Success in Communication 91

(a) An Initial Problem 93

2.2.4. Rational Animals: The Interdependence of Language and Thought 97

(a) The Argument from Surprise 100

2.2.5. Triangulation, Objectivity, and Cause-Determination 103

(a) Primitive Triangulation 106

(b) Linguistic Triangulation 111

2.2.6. Conclusion 115

3. Chapter 3: Davidson on Quine's Indeterminacy of Translation Thesis 117

Introduction 117

3.1. Quine's Arguments for the Indeterminacy of Translation 118

3.1.1. Quine's Physicalism 118

3.1.2. The Argument from Below 119

3.1.3. The Argument from Above 122

3.2. Davidson on Quine's Indeterminacy Arguments 127

3.2.1. The Measurement Scales Analogy 129

3.2.2. The Puzzle 131

3.3. Conclusion: A Dilemma for Davidson 134

4. Chapter 4: Davidson on Kripke's Wittgenstein 138

Introduction 138

4.1. Davidson on KW's Sceptical Argument and Sceptical Solution 138

4.1.1. The First Place: The Rule-Following Picture of Meaning 140

(a) Problems 142

4.1.2. The Second Place: "The"-Language-Determination 144

(a) Problems 146

4.1.3. The Third Place: The "Cause"-Determination 149

(a) Problems 153

4.2. Mark Joseph on Davidson's Alternative Solution 157

4.3. Conclusion 160

5. Chapter 5: Davidson's Non-Reductionism 162

Introduction 162

5.1. Verheggen on Davidson's Non-Reductionism 163

5.1.1. On Verheggen's View of Davidson's Non-Reductionism 171

5.2. Davidson on First-Person Authority 177

5.3. Wright's Objection to Davidson's Account of Self-Knowledge 182

5.4. Evaluating Wright's Objection 186

5.4.1. Davidson's Explanation of the Asymmetry 187

5.4.2. Wright's Objection Revisited 192

5.4.3. Is Wright's Objection Plausible? 196

5.5. Objections to Davidson's Account of Self-Knowledge 197

5.6. Wright's Response to KW's Sceptic 202

5.7. Wright's Judgement-Dependent Account of Intention/Meaning 209

5.8. Davidson's Judgement-Dependent Account of Intention/Meaning 215

5.8.1. Verheggen's Reading of Davidson's Non-Reductionism Revisited 216

5.9. Davidson's Judgement-Dependent Account of Intending 217

5.9.1. Intentional Actions 218

5.9.2. Prima-Facie Judgements vs. All-Out Judgements 222

5.9.3. Intending as Judgement-Dependent 224

5.10. Conclusion 226

6. Chapter 6: Summary and Conclusions 229

Appendix: The Early Davidson 235

References 270

Introduction

The philosophy of language can be regarded as an attempt to provide systematic explanations of our most basic linguistic practices, such as the practice of meaning something by an utterance. Saul Kripke and Donald Davidson are among the most influential philosophers of language who have been concerned with the main obstacles for providing such explanations. Kripke's book *Wittgenstein on Rules and Private Language* (1982) demonstrates one such attempt. In this book, he interprets Wittgenstein's *Philosophical Investigations* as presenting a sceptical argument against a certain sort of understanding of the practice of meaning something by an utterance, together with a sceptical solution to this sceptical problem. The view, against which Kripke's Wittgenstein (henceforth "KW") proposes his argument, implies that if a speaker means something by her words, then there is a rule, or a state of affairs, which determines the correct application of the words for the speaker. KW's sceptic – the figure who presents the sceptical argument in the first two chapters of Kripke's book – argues that there is no such thing as meaning anything by any word, since no fact about the speaker can be found capable of constituting the fact that she means one thing rather than another by her words or determining which rule she is following. KW's sceptical solution replaces such a problematic view with an alternative account of meaning, according to which all we need to do is to specify the conditions under which the sentences of the form "John means such-and-such by 'plus'" can be justifiably asserted and to illustrate the utility that asserting such sentences under such conditions has in our lives. According to this view, we are justified in asserting that a speaker means something by an expression if we can observe, in enough cases, that the speaker uses that expression in agreement with the way we would use it. Such a practice of attributing meanings to others has endless benefits in our lives: if such assertions can be legitimately made about a speaker, she can be accepted as a reliable speaker of our speech community and receive the benefit of interacting with other members of this community in a variety of ways.

Davidson (1984b, 1986, 1991b, 1992, 1994) has also argued against a certain sort of view of our linguistic practices. According to the view which Davidson rejects, speaking a language is to follow certain rules or conventions, which are fixed in advance of any particular conversation that two speakers may have. Davidson argues that the requirement of following such rules is neither necessary nor sufficient for a

plausible explanation of the practice of meaning something by an utterance. His alternative account suggests that success in this practice can be properly explained in terms of success in mutual interpretation: a speaker can be said to mean something by her words, no matter how she uses them, if she intends her utterance of the words to be understood in a particular way and the interpreter successfully interprets the speaker as such.

Davidson's and Kripke's views have been the subject of vigorous discussions for more than thirty years. Nonetheless, there is a considerable gap in the exegetical literature on both: there has not yet been a systematic investigation of the potential relationship between the two. Although their views have been individually explored, an investigation of how they may face each other has been absent. This thesis is intended as a modest contribution to filling this gap.[1] Although Davidson's later work on meaning has been inspired by Wittgenstein's *Investigations*, especially by his remarks on ostensive learning and private language, Davidson by no means sympathizes with the sceptical argument and solution that Kripke takes Wittgenstein to be offering. Rather, he attempts to provide an account of meaning, which can play the role of an alternative to KW's view. However, while Davidson's alternative view is anti-individualist, it falls short of the strong form of communitarianism proposed by KW.

To provide an overview of the main topics discussed in this dissertation, I start with a general outline of the thesis's chapters and then indicate the methodology I employed in my work.

Chapter 1: Kripke's Wittgenstein

In this chapter, I introduce Kripke's interpretation of the later Wittgenstein's view of meaning and linguistic understanding. This chapter will be divided into two general sections, the "Negative Part" and the "Positive Part". In the first part, KW's sceptic's sceptical argument is introduced. I will explain the steps through which the sceptical

[1] I can mention two books relatively on this topic, none of which has yet been published. In the last stages of my work, I noticed a forthcoming book by C. Verheggen and R. Myers (2016) on Davidson's remarks on triangulation in which Verheggen discusses the relation between Davidson's and Kripke's Wittgenstein's views. I was able to access a manuscript copy of this book and therefore able to discuss it in Chapter 5 of this thesis. Just very recently, I noticed that a volume on Davidson and Wittgenstein (*Wittgenstein and Davidson on Thought and Action*, edited by Verheggen. Cambridge University Press) is also in press. Unfortunately, this book will arrive too late for me to be able to discuss it in this thesis.

argument is established and consider ten possible responses to this argument, together with the sceptic's arguments against them. At the end, two different strands of the sceptical argument will be characterized, each of which takes for granted a different characterization of the view that the sceptical argument aims to reject. It will be shown that both strands, through different paths, lead to the same sort of radical sceptical conclusion about meaning, according to which there is no such thing as meaning anything by any word. In the Positive Part, KW's sceptical solution is presented and characterized with respect to the different strands of the sceptical argument.

Chapter 2: Donald Davidson

This chapter will explicate the main characteristics of Davidson's later account of meaning. Davidson deals with two main issues: firstly, the problems with a widely accepted view about meaning, which we will call the "Common View", and secondly, providing an alternative to the Common View. Accordingly, this chapter will be divided into two general parts: the "Negative Part" which introduces Davidson's argument against the Common View and the "Positive Part" in which Davidson's alternative account of meaning is presented. According to the Common View, if a speaker means something by a word, then there is necessarily a rule which determines the correct application of the word for the speaker and in terms of a shared grasp of which the success in the speaker's communication with others can be explained. Davidson introduces and rejects three versions of this view, each of which will be explored in this part of the chapter.

As an alternative to the Common View, Davidson proposes his new account of meaning, according to which it is success in mutual interpretation that can plausibly explain success in the practice of meaning something by an utterance. In the Positive Part of the chapter, we will also consider Davidson's response to three main objections to his view, which lead to his discussion of the notion of triangulation. Davidson's remarks on the notion of triangulation offer a causal explanation of the process of meaning-determination. I will try to provide a clear characterization of these remarks in this chapter.[2]

[2] The discussion of triangulation will appear again in Chapter Four, where we critically discuss the plausibility of Davidson's claim that employing the notion of triangulation can support his alternative solution to KW's sceptic's sceptical problem.

Chapter 3: Davidson on Quine's Indeterminacy of Translation Thesis

Chapter Three introduces and evaluates Davidson's grasp of Quine's arguments for the indeterminacy of translation. In this chapter, it will be argued that Davidson has actually missed the "sceptical" part of Quine's sceptical arguments. The critical points made in this chapter will then support our claim in Chapter Four that Davidson has neglected the "sceptical" part of KW's sceptical argument too. In this regard, this chapter will begin by introducing Quine's famous arguments for the indeterminacy of translation, i.e. the "Argument from Below" and the "Argument from Above". By way of providing a clear characterization of these arguments, it will be shown that they are intended by Quine to raise scepticism about fine-grained meaning facts: among all the available facts about meaning, which, according to Quine's physicalism, are the facts about the speaker's behaviour and the goings-on in her environment, we can find no fact about fine-grained meanings, i.e. the unique meaning that we intuitively expect each sentence to possess.

Davidson, however, disagrees with such a sceptical conclusion, though he accepts the main premises of Quine's arguments, i.e. Quine's physicalistic view and the indeterminacy of translation thesis. I will, however, argue that such a reading of Quine's arguments is highly puzzling and indeed leaves Davidson facing a dilemma: he must either remain Quinean, which results in embracing the sceptical conclusions of Quine's arguments, or give up on Quine's project, which leads to a radical divergence from Quine, more extreme than any divergence that Davidson is prepared to admit. It will also be argued that the main rationale Davidson introduces to justify his reading of Quine, i.e. the "measurement scales analogy", fails to resolve this puzzle: I will argue that it blurs the essential distinction between the epistemological problem of the underdetermination of theory and the metaphysical/constitutive problem of the indeterminacy of translation. Davidson thus appears to make the very same mistake that Quine accused Chomsky of making in his discussion of the indeterminacy of translation.

Chapter 4: Davidson on Kripke's Wittgenstein

In this chapter, I examine Davidson's grasp of KW's sceptical argument and solution. Two general questions are asked in this chapter: (I) Are Davidson's criticisms of KW's

sceptical argument and solution plausible? Answering this question involves an investigation of whether Davidson's grasp of KW's sceptical argument and solution is appropriate. Considering Davidson's presentation of KW's view, we detect three main places in Davidson's later writings in which he deals with KW's sceptical argument and solution. In each case, Davidson's criticisms are introduced and critically discussed. In this regard, (i) I will argue that Davidson's interpretation of KW's view is highly problematic and suffers from the same sort of problem as Davidson's treatment of Quine's arguments: he neglects the *sceptical* outcomes of the arguments. (ii) I will then argue that Davidson's objections to KW are also misplaced. To put it in general terms, Davidson criticizes the rule-following picture of meaning, which, according to Davidson, KW's sceptic takes for granted in his sceptical argument.[3] By appealing to the two strands of KW's sceptical argument characterized in Chapter One, I shall argue that the success of KW's sceptical argument does not necessarily depend on presupposing the rule-following picture of meaning. Rather, the second strand of the sceptical argument, which presupposes the existence of meaning facts rather than rules, can still establish the sceptic's desired sceptical conclusion. Therefore, Davidson's objection that the sceptical argument can be blocked by rejecting the rule-following picture of meaning is arguably mistaken.

(II) The second general question regarding Davidson's response to KW is whether Davidson's alternative solution to KW's sceptic's sceptical problem is successful. Davidson's alternative view of meaning is supposed to provide a solution to KW's sceptical problem. This solution appeals to the process of interpretation and induction: a speaker means something by her expressions if she intends her utterance of them to be interpreted in a particular way and the interpreter successfully interprets the speaker's utterance in that way. This account would not imply that speakers follow shared rules or conventions, which, in Davidson's view, KW's sceptical solution implies. In this regard, (i) I will argue that Davidson's alternative solution fails to resist KW's sceptic's sceptical problem. The sceptic is concerned with the metaphysical question of what it is that makes it the case that the speaker intends to be interpreted in one way rather than another, not with the epistemological question of how we can know what the speaker intends. Unless Davidson can properly answer the metaphysical question, there would be no justification whatsoever for any claim about success in interpretation. (ii)

[3] This criticism stems from Davidson's argument against the Common View, or Conventionalism, which is introduced in detail in Chapter Two.

Davidson's solution is supposed to be supported by the causal explanation of the process of interpretation, which Davidson offers in his discussion of the notion of triangulation. According to Davidson's externalism about meaning, for linguistic responses to be meaningful, the actual cause of the responses must be fixed, and such cause-determination essentially depends on the speaker's having linguistic interactions with at least one other person. I will argue that Davidson's use of the notion of triangulation cannot help him to resist KW's sceptic's sceptical problem, since the sceptic, even after presupposing the sort of causal facts that Davidson introduces in his discussion of the notion of triangulation, would still be able to run his sceptical argument. Having argued that Davidson's actual response to KW fails, this chapter ends with the question whether there are potential resources in Davidson's writings which could enable him to provide a plausible response to KW's sceptic's sceptical problem.

Chapter 5: Davidson's Non-Reductionism

This chapter seeks an answer to the above question and begins by claiming that the best way for Davidson to withstand KW's sceptic's sceptical argument is to hold non-reductionism about meaning, according to which the fact that a speaker means something by her words is itself a primitive fact. First of all, we will consider the most recent claim on this matter made by Claudine Verheggen who claims that Davidson's discussion of the notion of triangulation can establish or justify a non-reductionist response to KW's sceptic. I will argue that Verheggen's claim fails, since Davidson's use of the notion of triangulation presupposes non-reductionism and does not establish or justify it. It will be suggested that instead of looking at Davidson's remarks on triangulation, a persuasive defense of non-reductionism requires an investigation of his account of first-person authority. The reason is that KW's sceptic tried to rule out non-reductionism about meaning by an "Argument from Queerness", according to which non-reductionism leaves the primitive state of meaning something by an utterance, and the way we know such a state, completely mysterious. This argument, hence, leads directly to the problem of self-knowledge.

Davidson's account of self-knowledge will be introduced in the second part of this chapter. According to this account, the necessary conditions on the possibility of interpretation imply that speakers have noninferential knowledge of what they mean and believe. Introducing and examining Crispin Wright's objection to Davidson's account is

the next target of this chapter. According to Wright, Davidson's account fails, since it leads to a dilemma: speakers' authoritative knowledge of their beliefs must either be granted *before* interpretation takes place, in which case self-knowledge is presupposed rather than explained, or be credited *after* interpretation takes place, in which case there will be no real difference between the way the speaker knows himself and the way the interpreter knows the speaker. I will argue that Wright's objection fails because it neglects Davidson's actual explanation of self-knowledge. Nonetheless, I will also argue that Davidson's account fails for other reasons. This means that Davidson's account of self-knowledge would not be successful in dealing with KW's sceptic's argument from queerness.

As the final part of this chapter, through considering Wright's judgement-dependent account of meaning and intention as an alternative response to KW's sceptic, I will suggest that Davidson's account of intending manifests the essential features of such a judgement-dependent account. According to Davidson's view, when an agent intends to φ, she makes an unconditional, all-out judgement that doing φ is desirable for her. I will suggest, tentatively, that this type of account is capable of being extended to the case of meaning. The chapter hence concludes that, although Davidson's actual response to KW's sceptic fails, he may have the resources to provide us with a non-reductionist, judgement-dependent account of meaning, which has the potential to resist KW's sceptic's sceptical argument.

Chapter 6: Summary and Conclusions

This chapter brings together the main claims and criticisms made throughout the thesis.

Appendix: The Early Davidson

Although Davidson's discussion of KW's view emerges in his later work, which can be investigated more or less independently of his earlier writings, the roots of some of his later doctrines can be traced back to his earlier ones. Since the thesis presupposes a knowledge of Davidson's early philosophy of language, I will introduce the main themes of his earlier works in this appendix including Davidson's remarks on Tarski-style theories of truth, the procedure of radical interpretation, the interdependence of meaning and belief, holism, the principle of charity, and the indeterminacy of

interpretation. Nonetheless, those who are already familiar with Davidson's earlier view can skip this chapter.

Methodological approach and plan

This thesis aims to put together the views of two influential philosophers in order to evaluate their actual, as well as potential, responses to each other. To draw such an important comparison, it is essential to have a clear characterization of the views and arguments involved in their debates. Providing such characterizations, in turn, invites a careful investigation of the sort of views they are criticizing and the sort of views they are defending. I will pursue this goal in the first two chapters of the thesis. Regarding the characterization of the views in question, I will concentrate on the main works of these philosophers, though I will use the secondary literature on both where more clarification is required. An important methodological consideration concerning this project is to illustrate the *potential* relationship between Davidson's and KW's views and the way they could have responded to each other's criticisms. I will pursue this task in the last three chapters of the thesis.

1. Chapter 1: Kripke's Wittgenstein

Introduction

Wittgenstein, as presented by Kripke in his well-known book, *Wittgenstein on Rules and Private Language* (1982), is taken to be supporting a sceptical argument against a certain sort of understanding of the practice of meaning something by words, as well as a sceptical solution to the sceptical problem.[4] According to the view which is challenged by Kripke's Wittgenstein's sceptic,[5] if a speaker means something specific by a word, then there is a rule, or a possible meaning fact, which determines the correct application of the word for the speaker.[6] The sceptical argument aims to arrive at the conclusion that such a conception of meaning leads to a radical sceptical conclusion about meaning, according to which there is no such thing as meaning anything by any word at all. KW, however, does not accept such an insane sceptical conclusion; instead, he proposes an alternative to the problematic view which, according to KW, brings about such an intolerable conclusion.[7]

Accordingly, this chapter will be divided into two general parts, "the Negative Part" and "the Positive Part". In the Negative Part, I am going to provide an exposition of KW's sceptic's sceptical argument, as presented in the first two chapters of Kripke's book. I will characterize two strands of this argument, one of which rests on the rule-following picture of meaning and the other takes for granted the existence of meaning facts. Both will be shown to lead to the radical sceptical conclusion. The sceptical

[4] Kripke emphasizes that this interpretation of Wittgenstein's work is not what Wittgenstein himself necessarily endorses: "my method is to present the argument as it struck me, as it presented a problem for me, rather than to concentrate on the exegesis of specific passages" (1982, p. viii).

[5] KW's sceptic is the figure running the main argument, i.e. the sceptical argument, in the second chapter of Kripke's book. KW himself is the character proposing the sceptical solution to the sceptical problem in the third chapter of the book.

[6] The view against which KW's sceptic argues is sometimes described as Traditional, or Classical, Realism about meaning, as Kripke himself says "In order for Wittgenstein's sceptical solution of his paradox to be intelligible, the 'realistic' or 'representational' picture of language must be undermined by another picture (in the first part)" (Kripke, 1982, p. 83). See also (Wilson, 1994, p. 373; 1998, p. 105-106).

[7] It is important to note that there are different readings of KW's sceptical argument and solution, among which my characterization of KW's view is closer to Wilson (2006, 1998, 1994), than others. For different interpretations of KW's sceptical argument and solution, see e.g. McGinn (1984), Boghossian (1989), Hattiangadi (2007), Baker and Hacker (1984), McDowell (1984, 1993), Zalabardo (1989), Byrne (1996), Kusch (2006), Davis (1998), Miller (2010), and Wright (1989, 1984).

solution proposed by KW will be introduced in the Positive Part. I shall characterize the sceptical solution with respect to the two mentioned strands of the sceptical argument and illustrate the sort of alternative view of meaning which it offers.

1.1. The Negative Part: Kripke's Wittgenstein's Sceptic's Sceptical Argument

KW's sceptic starts his challenge by introducing the sort of common sense conception of the practice of meaning something by words that we intuitively take for granted in our ordinary talk of this practice. To illustrate this conception, the sceptic brings in the practice of following rules. Based on such a conception, there are specific rules that we learn or grasp in one way or another, e.g. by being taught by our parents or teachers, which are supposed to determine the correct application of our words for us. It is indeed because of our grasp of such rules that we know how we should use the words in the future. KW's sceptic gives an arithmetical example to clarify this point. For example, by learning how to use "+" we thereby grasp a rule ("x + y = z" is true if and only if the number denoted by "x" added to the number denoted by "y" yields the number denoted by "z"). This rule "determines my answer for indefinitely many new sums that I have never previously considered" (Kripke, 1982, p. 7). Hence, successful communication will be dependent on the communicators' grasp of such rules: we all use our words in a certain way because we have grasped certain rules, which determine the conditions under which those words can be used correctly. For instance, as Kripke says, "I, like almost all English speakers, use the word "plus" and the symbol "+" to denote a well-known mathematical function, addition" (1982, p. 7). In this sense, we all master the rule regarding the correct use of the word "plus", or the plus sign, "+". According to this rule, the plus sign "+" refers to a certain mathematical function, that is, the addition function.

The important characteristic of these rules is that they are general in their character: once we grasp such rules, we will be able to apply them to a potentially indefinite number of future cases. In the case of the addition function, for example, "The function is defined for all pairs of positive integers" (Kripke, 1982, p. 7), so that, when we learn this rule, we learn a rule that is applicable to an infinite number of cases. KW's sceptic invites us to assume that the number 57 is larger than any number we have previously faced. According to the rule-following picture of meaning, although I have not yet

computed the addition problem “57 + 68=?”, I know that there is just one unique and correct answer to this arithmetical query *if* “+” is supposed to denote the addition function and “57” and “68” are supposed to have their standard denotations – “57” denotes the number 57 and “68” denotes the number 68. Similarly, as a result of grasping the rule for addition, I know that the addition problem “2 + 2=?” has a unique answer, “4”, and if “+” denotes the same function that it did in the case of “2 + 2=?”, I know that there would be one correct answer to “57 + 68=?” as well.

KW’s sceptic holds that the same story is true in the case of the practice of meaning something by words in general. If “green” has so far been used by me to mean *green*, then it will be correct to apply this word to certain things and incorrect to apply it to certain other things. It is what I meant by the word in the past that determines how I ought to use the word in the future. Putting it in terms of rules, if I have successfully grasped the rule determining the correct application of “green”, then I ought to use the word in a certain way in the future if I want to remain faithful to that rule. In this sense, the meaning of the word, or the relevant rule I have grasped regarding the application of the word, determines the correct use of the word in the future: if I meant *green* by “green” in the past, then I ought to apply “green” to certain (green) things in the future. Similarly, “57 plus 68” will have a unique correct answer if I am supposed to mean the same thing that I did by the word in the past, that is, *plus*. However, what is the answer to the question “57 + 68 = ?”?

Obviously, after performing the computation and checking the result, I will come up with the answer “125”. According to Kripke,

> It is correct both in the arithmetical sense that 125 is the sum of 68 and 57, and in the metalinguistic sense that ‘plus’, as I intended to use that word in the past, denoted a function which, when applied to the numbers I called ‘68’ and ‘57’, yields the value 125. (1982, p. 8)

As the result of following the general rule for addition, “Ordinarily, … I do not simply make an unjustified leap in the dark”; rather, “I follow directions I previously gave myself that uniquely determine that in this new instance I should say “125”” (Kripke, 1982, p. 10). Hence, “*if* I meant plus, then unless I wish to change my usage, I am justified in answering (indeed compelled to answer) “125”, not “5”” (Kripke, 1982, p. 11). What I meant by my words in the past determines how I ought to respond in the future.

1.1.1. The Sceptical Challenge

The bizarre sceptic, however, entirely rules out my answer, i.e. "125", to be the correct one. He thinks that I am now giving a wrong answer because I actually mean something different from addition by "+". The sceptic believes that the correct answer ought to be "5", not "125". According to the sceptic, if we are to remain faithful to what we meant by "plus" in the past, the answer we ought to give is nothing but "5", since, as he says, "in the past I used 'plus' and '+' to denote a function which I will call 'quus' and symbolize by '⊕'. It is defined by:

$x \oplus y = x + y$, if $x, y < 57$

$= 5$ otherwise" (1982, pp. 8-9).

If I answer "125" to the mentioned arithmetic query, it is because "I am now misinterpreting my own previous usage. By 'plus' … I *always meant* quus; now, under the influence of some insane frenzy, or a bout of LSD, I have come to misinterpret my own previous usage" (Kripke, 1982, p. 9).

The important question that the sceptic asks is: "Who is to say that this is not the function I previously meant by '+'?" (1982, p. 9). The sceptic continues "wild it indubitably is, no doubt it is false; but if it is false, there must be some fact about my past usage that can be cited to refute it. For although the hypothesis is wild, it does not seem to be *a priori* impossible" (1982, p. 9). The sceptic's challenge for us, hence, is to introduce a fact about our past usage of the word "plus" that can rule out his sceptical hypotheses, e.g. that we meant *quus* by "plus" in the past. What facts do we have to offer in order to prove that he is wrong, that what we meant in the past by "plus" has been *plus*, not *quus* or anything else? If we meant *plus* by "plus", then "125" is the correct answer, while if we meant *quus* by "plus", then the sceptic is right to insist that "5" is correct.[8] What do we have in terms of which we can argue that the correct use of "plus" results in answering "125", and not "5"? The second step of KW's sceptic's sceptical argument starts with discussing and criticizing the possible candidate facts that

[8] We should note that the sceptic's problem is not a problem about the arithmetical procedures we follow to compute the answer so that we can disprove it by offering some mathematical proof. The problem, again, is meta-linguistic. Indeed, even if we can say something constructive about the mathematical procedure leading us to the arithmetic value 125 as the answer to the arithmetic query, we still fail to argue that our linguistic answer "125" to the query is correct *if* we fail to determine whether we actually meant *plus* or *quus* by "plus" in the past.

we may introduce in order to rule out the sceptical hypothesis. Obviously, if no fact of the matter as to what I meant by my word in the past can be found, there would be no fact of the matter about what I mean by the word now or in the future. As KW's sceptic says, "if there can be no fact about which particular function I meant in the *past*, there can be none in the *present* either" (1982, p. 12). We can, hence, see that the sceptical problem is a *metaphysical* problem, rather than an epistemological one: the sceptic is not concerned with the question how we can know what the speaker means by his linguistic expressions; rather, he asks whether there is any meaning to be known at all, that is, whether there is any fact of the matter as to what the speaker means by his expressions. As Kripke clarifies, "it is clear that the sceptical challenge is not really an epistemological one. It purports to show that nothing in mental history of past behavior – not even what an omniscient God would know – could establish whether I meant plus or quus" (1982, p. 21; see also p. 39).

There is, however, an important point to note regarding what the sceptic indeed demands. The sceptic requires us to introduce some fact about what we meant by "plus" in the past, which can rule out the sceptical hypotheses about the meaning of the word, such as "plus" meaning *quus*. Such a fact, however, must meet two important conditions. According to the first condition, this fact must be able to determine what we meant by the word in the past. As Kripke puts it, the sceptic, first of all, "questions whether there is any fact that I meant plus, not quus, that will answer his sceptical challenge" (1982, p. 11). Any suggested fact must dismiss all the sceptical hypotheses about what I meant by "plus" in the past. Secondly, this fact must be able to tell us what we ought to do in the future, that is, it must determine the correct use of the word in future cases. In other words, the sceptic "questions whether I have any reason to be so confident that now I should answer '125' rather than '5'" (Kripke, 1982, p.11). Any fact we may offer as to what we meant by "plus" in the past must tell us that, in the present and future cases, the word ought to be used in a certain way. We can say that, in this sense, the suggested facts must satisfy the "normative" feature of meaning: it is not enough that we cite a fact about our past usage of words, without taking into account that it has to determine the correct use of the words regarding future cases. If a suggested fact successfully determines what we meant by our words in the past, but fails to determine how we ought to use them in the future, then it indeed fails to properly deal with the sceptical challenge. The second condition naturally stems from the common sense conception of meaning that the sceptic took for granted in the beginning: when we

learn how to use "plus", we grasp a general rule which determines the unique, or correct, answer to a potentially infinite number of arithmetical queries in which "+" features. Similarly, if there is any fact as to what we meant by our words in the past, it must tell us how we ought to use them in the future. As Kripke summarizes, the answer to the sceptic "must [first] give an account of what fact it is … that constitutes my meaning plus, not quus. [Secondly] it must, in some sense, show how I am justified in giving the answer '125' to '68+57'" (1982, p.11).

The last point we need to consider is about the sort of facts which the sceptic allows us to employ in order to respond to his challenge. It is emphasized by Kripke that the sceptic does not confine us to behavioural facts, that is, the observable behaviour of the speaker; he imposes no constraint on the sort of fact we may appeal to:

> Another important rule of the game is that there are no limitations, in particular, no behaviorist limitations, on the facts that may be cited to answer the sceptic. The evidence is not to be confined to that available to an external observer, who can observe my overt behavior but not my internal mental state. (1982, p. 14)

Before discussing in detail how the sceptic rules out the plausibility of the various possible responses to his sceptical problem, it is important to look briefly at the view the sceptic aims to reject, which, I think, can be properly called "Traditional Realism" about meaning.

1.1.2. The Sceptical Challenge and Traditional Realism

If the sceptic can successfully show that there is no fact of the matter as to what we mean by our words, then he can seriously challenge a certain sort of realist view about meaning, which we can call "Traditional Realism". This view, however, can be characterized in two ways: in terms of the rule-following picture of meaning and in terms of the presumption that there are meaning facts. We saw that, according to the rule-following picture of meaning, there are rules determining the correct application of words. According to Traditional Realism characterized in terms of the rule-following picture,

> if the speaker, S, means *green* by his utterance of "green", then there is a rule, R, (e.g. that "green" applies to green things and does not apply to non-green things) that determines the correct application of "green" for S.

There are fixed rules, for Traditional Realists, which speakers of a language follow and shared grasp of which explains the speakers' success in linguistic communication. Since, for instance, we have all grasped the rule for addition, we all come up with the correct answer "125" to the query "57 + 68 = ?". The sceptic's question, however, was: Is there anything about the speaker's behavioural or mental life that determines which rule the speaker followed? The defenders of Traditional Realism, by introducing some fact about the speaker, must try to rule out the sceptic's sceptical hypotheses, such as the hypothesis that, in the past, the speaker actually followed the rule R*, rather than R, regarding the use of the word "green", where, according to R, "green" applies to green things only, while R* indicates that "green" applies to green things up to time *t* and to blue things after *t*. In the former case, "green" means *green*, while, in the latter, it means something else, e.g. *grue*.[9] If the speaker can be said to be following R, and hence to mean *green* by "green", then the application of "green" to a blue thing at t^+ is incorrect. However, if the speaker is following R*, and hence meaning *grue* by "green", such a use of the word will be correct. The realists, thus, face the challenge of introducing a fact about the speaker making it the case that the speaker follows R, rather than R*.

Traditional Realism can also be characterized without appealing to the rule-following picture of meaning, rather in terms of the claim that there exist meaning facts determining the correct application of words. In other words, there are facts of the matter as to what a speaker means by his words, which determine the correct use of the words for the speaker:

> If S means *green* by his utterance of "green", then there is some possible meaning fact, or a state of affairs, F, (e.g. that S means *green*, not anything else, by "green") that determines the correct application of "green" for S.

Based on this characterization, if I mean *green* by "green", then there is a state of affairs, or a possible fact, e.g. that by "green" I mean *green*, and not *grue* or anything else, which determines the correctness conditions for the application of "green" for me. KW's sceptic can now be regarded asking: Who is to say that it is the state of affairs F, i.e. that "green" means *green*, and not the state of affairs F*, i.e. that "green" means *grue*, that determines the correct application of "green" for the speaker? What KW's sceptic demands is the introduction of some fact about the speaker, behavioural or otherwise, which constitutes the fact that he means e.g. *green*, and not *grue*, by "green".

[9] The "grue" example was originally suggested by Nelson Goodman (1983).

Hence, Traditional Realism implies that there are facts of the matter about what the speaker means by his words, or about which rule he follows, which determine the correct application of the words for the speaker. The sceptic claims that no suitable meaning-constituting fact can be found among the facts about the speaker's behaviour and mental life. Any fact that the advocates of Traditional Realism offer, as mentioned above (Section 1.1.1, pp 13-14), must satisfy two conditions: it must (I) determine that the speaker meant *green*, not *grue* or anything else, by the term "green" and (II) be capable of justifying the application of "green" to a green object or telling the speaker how "green" ought to be applied. KW's sceptic, however, maintains that there cannot be found any fact about the speaker which can satisfy the above two conditions. We can now consider the candidate facts which Traditional Realists introduce and the sceptic rejects.

1.2. The Candidate Facts

KW's sceptic aims to argue that no fact about the speaker can be found which determines which rule he followed or constitutes the fact that he meant one thing rather than another by his words in the past. In this regard, the second step of the sceptical argument begins by introducing and rejecting ten candidate facts which may be suggested in response to the sceptical challenge.

1.2.1. Previous Behaviour

The first candidate is the behaviour of the speaker in the past. As KW's sceptic said, "I have performed … only finitely many computations in the past" (1982, p. 8). My past behaviour has been the same whether I followed the addition function or quaddition function, since, for the numbers smaller than 57, their sum and their quum have been the same and I have not dealt with numbers larger than or equal to 57 until now. Hence, nothing in my past behaviour can show whether I meant *plus* or *quus* by "plus": my past behaviour has been compatible with both meaning *plus* and meaning *quus* by "plus".

1.2.2. Internalized Instructions

Another candidate can be introduced by adverting to some other instructions or rules that I have learnt for adding numbers in the past, such as "the counting rule". As KW's sceptic explains this response:

> I learned – and internalized instructions for – a *rule* which determines how addition is to be continued. What was the rule? Well, say, to take it in its most primitive form: suppose we wish to add *x* and *y*. Take a huge bunch of marbles. First count out *x* marbles in one heap. Then count out *y* marbles in another. Put the two heaps together and count out the number of marbles in the union thus formed. The result is $x + y$. (1982, p. 15)

Hence, instead of talking about the addition rule, we can appeal to a different sort of rule, e.g. the counting procedure, which can help us to establish that we have been following the addition rule.

The sceptic's response to this suggestion, however, is as before. As with "plus" itself, I have applied "count" to some finite number of past cases. The sceptic can put forward the sceptical hypothesis that, perhaps, by "count" in the past I meant *quount*, not *count*, which can be specified as follows: "To 'quount' a heap is to count it in the ordinary sense, unless the heap was formed as the union of two heaps, one of which has 57 or more items, in which case one must automatically give the answer '5'" (Kripke, 1982, p. 16). The answer, hence, is "5" again. The chief problem with this sort of response is that it appeals to another rule to explain the rule in question. And it is obvious that the new rule will just inherit the same sort of sceptical problem: "the point is perfectly general: if 'plus' is explained in terms of 'counting', a non-standard interpretation of the latter will yield a non-standard interpretation of the former" (Kripke, 1982, p. 16). Appealing to other rules or instructions will trap us in a vicious regress of rules interpreting other rules. At the end, the process of "a rule for interpreting a rule" must stop at some point by introducing a basic rule, and the sceptic can again mount his sceptical challenge: "How can I justify my present application of such a [basic] rule, when a sceptic could easily interpret it so as to yield any of an indefinite number of other results?" (Kripke, 1982, p. 71). The sceptical problem would not be answered by appealing to some further instructions, rules, norms, procedures, algorithms, and the like. Therefore, it seems that my application of my words is just "an unjustified stab in the dark" (Kripke, 1982, p. 17).

1.2.3. Mathematical Laws

A suggestion for resisting the sceptic's non-standard interpretation of "+" is given by those mathematicians who think that taking the plus sign to be denoting quaddition cannot satisfy some important mathematical laws, which seem to work only if "+" denotes the addition function. According to this response, the addition function is the only function that satisfies, for example, the law "$x + y' = (x + y)'$", where x' is the successor function. The reason is that, if "+" stands for the quaddition function, we will gain an unacceptable result. For instance, if "+" denotes the addition function, we have: $57 + 68' = 57 + 69 = 126$, which is the same result we get from the right-hand side of the equation too: $(57 + 68)' = 125' = 126$. But, if "+" denotes the quaddition function, we have: $57 \oplus 68' = 57 \oplus 69 = 5$, which is not equivalent to the result we come up with from the right-hand side of the equation: $(67 \oplus 12)' = 5' = 6$. Also, there are other mathematical rules that will be wrong if "+" denotes the quaddition function, such as the law "$(x + 0 = x)$"[10].

However, the sceptical problem cannot be met by appealing to such mathematical laws, since the variables, the equality sign, the plus sign, and the like, are all open to non-standard interpretations. For instance, the first law is actually a law with quantifiers: $(\forall x)(\forall y)(x + y' = (x + y)')$. The other signs in this law, such as the universal quantifier or the equality sign, have been applied in a limited number of cases, and the sceptic can propose some non-standard interpretations of them compatible with his non-standard interpretation of "+". For instance, as Kripke says, the universal quantifier, $(\forall x)$, "might mean for every $x < h$, where h is some upper bound to the instances where universal instantiation has hitherto been applied, and similarly for equality" (1982, p. 16-17, fn. 12).

1.2.4. Mental Images

The fourth candidate fact about the speaker is proposed by appealing to some internal mental images that a speaker has when he applies a word, such as "green", to certain objects. "It has been supposed that all I need to do to determine my use of the word

[10] While, in the case of "+" denoting the addition function, we have $58 + 0 = 58$, we have, in the case of "+" denoting the quaddition function, $58 \oplus 0 = 5$.

‘green’ is to have an image, a sample, of green that I bring to mind whenever I apply the word in the future” (Kripke, 1982, p. 20). According to this suggestion, my use of “green” has been associated with a mental image of greenness, so that I have a sample of this colour in my mind. Now, when I want to apply this word to a new object, I just need to bring to mind that mental image and check if this object has the same colour that the mental image I have in my mind does.

The problem with this response is that the mental image of a green object, or of greenness, works just like a sign, which needs to be interpreted in a certain way and, hence, *is* open to interpretation. The sceptic’s question is: What is this sample a sample for? Suppose that up until time t, my use of “green” has been associated with a certain mental image of greenness. Now, at time t_1, I encounter an emerald. The sceptic’s question is whether “green” applies to this object, i.e. the emerald. We think that, according to the mental image that we have of greenness, “green” applies to the emerald correctly. However, the sceptic claims that this mental image can be interpreted as standing for grueness, or grue objects, so that it is incorrect to apply “green” to the emerald at t_1. What is it that makes it the case that the mental image we have in mind stands for green rather than grue objects? The mental image in our mind does not by itself say anything constructive about how we ought to use the word in the future: “Perhaps by ‘green’, in the past I meant grue, and the color image, which indeed was grue, was meant to direct me to apply the word ‘green’ to grue objects always. If the blue object before me now is grue, then it falls in the extension of ‘green’, as I meant it in the past” (Kripke, 1982, p. 20).

However, one may claim that we should not treat the sample as a sign but just as a physical sample or a physical object. In this case, we can say that “green” applies to an object if and only iff that object has a certain relation to our physical sample, e.g. if that object has “the same colour” as the sample does. Hence, it may be claimed that “in the past I stipulated that ‘green’ was to apply to all and only those things ‘of the same color as’ the sample” (Kripke, 1982, p. 20). The reason why this response does not work is obvious: “The sceptic can reinterpret ‘same color’ as same *schmolor*, where things have the same schmolor if …” (Kripke, 1982, p. 20). The sceptic can always introduce new non-standard interpretations of any relation that one may claim the object and the sample have. For instance, the sceptic can say that “green” applies to x if and only if x has the same schmolor as the sample, where having the same schmolor is defined as follows: if $t < t_1$, x has the same colour as the sample; otherwise, x has the same colour

as the sample of blue.

1.2.5. Dispositions

According to Dispositionalism,

> to mean addition by '+' is to be disposed, when asked for any sum '*x* + *y*' to give the sum of *x* and *y* as the answer (in particular, to say '125' when queried about '68 + 57'); to mean quus is to be disposed when queried about any arguments, to respond with their *quum* (in particular to answer '5' when queried about '68 + 57'). (Kripke, 1982, pp. 22-23)

The speaker has always been disposed to apply "green" only to green things, because of the way he has been taught by his teachers, parents, etc. . In the present case, he is disposed to apply "green" to this green object and, in the future, he will be disposed to apply it only to green things. The speaker, it is claimed by the dispositionalist, has not been disposed to apply "green" to blue things, and hence what he has meant by "green" has always been *green*, not *grue* or anything else. Hence, the fact that the speaker has been disposed to respond in a certain way determines what he meant by his word.

Kripke's sceptic rejects this suggestion by claiming that it cannot satisfy the two conditions that a proper candidate fact must meet. First of all, dispositions cannot constitute the fact that I mean *plus* by "plus", and not anything else, because they are finite in their nature: "Not only my actual performance, but also the totality of my dispositions, is finite" (Kripke, 1982, p. 26). We can call this problem the Finiteness Problem. KW's sceptic claims that "it is not true, for example, that if queried about the sum of any two numbers, no matter how large, I will reply with their actual sum, for some pairs of numbers are simply too large for my mind – or my brain – to grasp" (1982, pp. 26-27). There are always numbers that are too large for our mind to be able to compute their sum, given that our life-span and processing capacity is finite. If that is true, then the sceptic can claim that the plus sign, "+", stands for a function called *skaddition*, which can be characterized as follows:

x * y = x + y, if x, y are small enough for us to compute in our life time,

= 5, otherwise.

In this situation, even if the speaker *actually* responds by the sum of numbers in his life, he cannot be said to mean *plus* by "plus". Rather, the totality of the speaker's

dispositions is compatible with "plus" meaning *skaddition*. As a conclusion, the facts about a speaker's dispositions to respond in certain ways cannot determine what the speaker meant by his words. This view cannot "enable us to 'read off' which function I mean by a given function symbol from my disposition" (Kripke, 1982, p. 26).

There is, however, a more sophisticated version of Dispositionalism, which is introduced by the proponents of this view to deal with the finiteness problem. Dispositionalists claim that the finiteness problem arises because the notion of disposition that the sceptic employs in his criticism is too simple and crude. What we need is a more sophisticated notion of disposition, one that concerns how speakers are disposed to act under ideal or *ceteris paribus* conditions: "*Ceteris paribus*, I surely will respond with the sum of any two numbers when queried" (Kripke, 1982, p. 27). However, the sceptic's problem is: How must we interpret the "*ceteris paribus*" phrase? What are these ideal conditions? There are two readings of these conditions that the sceptic discusses and criticizes.

According to one interpretation, the idealized conditions are supposed to be conditions specified as follows:

> [I]f my brain had been stuffed with sufficient extra matter to grasp large enough numbers, and if it were given enough capacity to perform such a large addition, and if my life (in a healthy state) were prolonged enough, then given an addition problem involving two large numbers, *m* and *n*, I would respond with their sum, and not with the result according to some quus-like rule. (Kripke, 1982, p.27)

Under the above conditions, I would respond to the sum, and not the quum, of any two numbers, no matter how large they are. The problem with this reading of the idealized conditions is that "we have no idea what the results of such experiments would be. They might lead me to go insane, even to behave according to a quus-like rule. The outcome really is obviously indeterminate" (Kripke, 1982, p. 27).

There is, however, a more sophisticated reading of what the ideal conditions are, according to which,

> if I somehow were to be given the means to carry out my intentions with respect to numbers that presently are too long for me to add (or to grasp), and if I were to carry out these intentions, then if queried about '$m + n$' for some big *m* and *n*, I would respond with their sum (and not with their quum). (Kripke, 1982, p. 28)

However, KW's sceptic thinks that this move is hopeless because it is circular. Adding such idealized conditions begs the question against the sceptic's main challenge: these

conditions presuppose that I already have an *intention* to take "+" to denote some specific function, i.e. the addition function. In our discussion of the sceptical challenge (Section 1.1.1, pp. 12-14), we saw that the sceptic cast doubt on the relation between past meanings, or intentions, and the current use of words. If I intended to use "plus" in a certain way in the past, then I ought to use it in that way in the future. The problem was to determine what I intended to mean by "plus" in the past. Regarding the above idealized conditions, the sceptic's question is: What are my intentions regarding the kind of mathematical function that I would apply to the numbers *m* and *n*? The correctness of my response would depend on what I meant by "plus", or on the function I intended the plus sign to denote: if I meant *plus* by "plus", then I would be disposed to respond with the sum of the numbers, while if I meant *quus* by "plus", then I would be disposed to respond with their quum. What is our reason to favor one of the above options? The sceptic's claim is that we would not be able to choose one of them without circularity: we need to presuppose that the speaker meant *plus* by "plus" in order to argue that the speaker would respond by the sum of the numbers. Hence, if we accept the idealized conditions, "then the circularity of the procedure is evident. The idealized dispositions are determinate only because it is already settled which function I meant" (Kripke, 1982, p. 28). But, "of course the sceptic is challenging the existence of just such a fact" (Kripke, 1982, p .28).

There is, however, a second problem with the dispositional account of meaning:

> [Dispositionalism] misconceives the sceptic's problem – to find a past fact that *justifies* my present response. As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition on such a candidate, …, that it should *tell* me what I ought to do in each new instance. (Kripke, 1982, p. 24)

This problem with the dispositional account concerns the second condition that the sceptic introduced on candidate meaning facts. Dispositional facts fail to satisfy the normative feature of meaning, since they cannot justify our answering "125", rather than "5", to the query. As the sceptic asks, "is not the dispositional view simply an equation of performance and correctness?" (1982, p. 24). According to the dispositional account, what we are disposed to do (our performance) determines what we mean and, in this sense, determines whether or not we are using the word correctly (the correctness conditions). In this case, however, it is hard to accommodate systematic linguistic errors or mistakes: we want to say that if the speaker responds by "5", then he makes a mistake. But Dispositionalism has the consequence that it is the way the speaker is disposed to respond that determines what he means by his word. This will lead us to a

problem when we note that "most of us have dispositions to make mistakes" (1982, pp. 28-29). For instance, suppose that the speaker, for any reason, is disposed to respond slightly different to certain arithmetic queries: he responds to the query "6 + 5 = ?" by "10", to "6 + 6 = ?" by "11", and so on. The speaker, in other words, forgets to carry in adding the numbers. In these situations, we normally say that these people have made a mistake. But dispositionalists cannot make such a claim because, as we saw, "the function someone means is to be *read off* from his dispositions" (Kripke, 1982, p. 29). Instead of saying that the speaker means *plus* by "plus", but makes systematic mistakes, we have to say that the speaker means something different by "plus". We cannot say that the speaker is responding wrongly, since, by doing so, we are presupposing in advance that the speaker meant *plus* by "plus". Rather, we must say that what the speaker is disposed to do determines what he means by his word, in which case "a disposition to make a mistake is simply a disposition *to give an answer other than the one that accords with the function I meant*" (Kripke, 1982, p. 30). Thus, the dispositionalists cannot account for the notion of a systematic mistake. Moreover, dispositional facts are facts about how speakers *will* use expressions, whereas meaning facts are supposed to determine how speakers *ought* to use them. Hence, the dispositional account fails to accommodate the normative feature of meaning. As KW's sceptic puts it,

> the dispositionalist gives a *descriptive* account of this relation [i.e. the relation between meaning and future action]: if '+' meant addition, then I will answer '125'. But this is not the proper account of the relation, which is *normative*, not descriptive. The point is *not* that, if I meant addition by '+', I *will* answer '125', but that, if I intend to accord with my past meaning of '+', I *should* answer '125'. (1982, p. 37)

1.2.6. Machines Embodying Intentions

Another response, along with Dispositionalism, is given by appealing to the claim that a machine can be designed in a way to produce the sum of any two numbers as its outputs. As KW's sceptic introduces this response, "if I build such a machine, it will simply grind out the right answer, in any particular case, to any particular addition problem. The answer that the machine would give is, then, the answer that I intended" (1982, p.33). This is indeed a version of the dispositional account: it takes us as machines, which are disposed to give the sum of any two numbers, no matter how large they are. In this sense, we can say that this machine embodies our intention to mean *plus* by "plus", that is, that it produces the responses that we would give if we meant *plus* by

“plus”. The function which the machine embodies is nothing but the addition function, since it produces the sum, and not the quum, of numbers. This response, hence, aims to solve the finiteness problem: no matter how large the numbers are, the machine would automatically produce their sum, without being disposed to make mistakes.

The sceptic begins his criticism of this view by asking for clarification for the sort of “machines” we have in mind here. We can regard machines in two different ways, either as programs or as concrete objects. (I) If a machine is taken to be a program, then “exactly the same problems arise for the program as for the original symbol ‘+’: the sceptic can feign to believe that the program, too, ought to be interpreted in a quus-like manner” (1982, p. 33). The program, if it is some abstract program in the mind of the programmer, can be interpreted in a non-standard way, just as the rules and internal instructions in the mind of the speaker were open to re-interpretation. If this program is something that is written somewhere, then it is nothing but a set of signs and symbols, and the sceptic can again claim that, according to his interpretation, the plus sign in this program denotes the quaddition function, the equality sign denotes schomuality, and the like. (II) However, the machine may be claimed to be a concrete machine, something like a computer, a calculator, and the like. This way of regarding a machine will not help against the sceptic, though. First of all, when we want to talk about, or interpret, the behaviour of the machine and its functioning, we are bound to bring in instructions, that is, to use a language, some codes, and the like. In this situation, the sceptic will raise his standard objection: the program, the codes, or the instructions, can be interpreted in a quus-like, or in a skaddition-like, way. Secondly, as the sceptic says, “the machine is a finite object, accepting only finitely many numbers as input and yielding only finitely many as output – others are simply too big” (1982, p. 34). Thus, the sceptic can take the machine to be working in accordance with the skaddition function. And, there is a third problem with this response: “Actual machines can *malfunction*: through melting wires or slipping gears they may give the wrong answer” (Kripke, 1982, p. 34). In this case, the sceptic will argue that we cannot give a non-circular account of the conditions under which the machine is malfunctioning.

1.2.7. Simplicity

The next response to the sceptic is based on the alleged fact that the hypothesis that the speaker means *plus* by “plus” is “to be preferred as the *simplest* hypothesis” (Kripke,

1982, p. 38). According to this reply, we have a reason to prefer the hypothesis that the speaker means *plus* by "plus" over indefinitely many other hypotheses about what he might mean by the word. The sceptic, however, thinks that "such an appeal must be based either on a misunderstanding of the sceptical problem, or of the role of simplicity considerations, or both" (1982, p. 38). First of all, we should note that the sceptical problem was metaphysical, not epistemological: it was not a challenge about how we can *know* what the speaker means by his words. Rather, his claim was that "there is no fact as to what I meant" (Kripke, 1982, p. 38). A response to the sceptical problem, hence, cannot be given by reducing it to an epistemological problem and, thereby, presupposing that there are facts about meaning, which our hypotheses are about, so that we can prefer the simplest one to the rest. The sceptic's question is: What are the hypotheses about? Are they about some fact regarding what the speaker means by his words? The sceptic's answer is "no". There is no fact of the matter as to what a speaker means by his words, whether *plus* or *quus*. If we want to claim that one of the hypotheses about the things the speaker means by his word is the simplest, then we should first argue that there *is* such a fact about what the speaker means by the word, about which we have the competing hypotheses. We need to deal with the metaphysical problem first, that is, whether such a fact exists, before engaging in the epistemological problem.[11] If that is true, then the talk of the *simplest* hypothesis is irrelevant in this case, since "rival hypotheses" are hypotheses about some fact of the matter. But, in the case of meaning, the sceptic's claim is that there is no fact of the matter. The suggested hypotheses, hence, are not *genuine* ones.[12] As a result, the simplicity criterion cannot

[11] Hence, according to the sceptic, simplicity may help us to decide between some two hypotheses, but it does not help to determine what the hypotheses *are*, or are about: "Simplicity considerations can help us decide between competing hypotheses, but they obviously can never tell us what the competing hypotheses are. If we do not understand what two hypotheses state, what does it mean to say that one is 'more probable' because it is 'simpler'?" (1982, p. 38).

[12] The sceptic gives an example of the case, in which we have genuine hypotheses, so that we can employ the simplicity criterion: "Suppose there are two conflicting hypotheses about electrons, both confirmed by the experimental data" (1982, p. 38). While God, who is able to directly see the facts about electrons, does not need to appeal to any observation, evidence, or simplicity criterion, to decide between the rival hypotheses, we as human beings, without the God's powers, need to rely on our evidence and experiments to make such a decision. Now, "if two competing hypotheses are indistinguishable as far as their effects on gross objects are concerned, then *we* must fall back on simplicity considerations to decide between them" (1982, p. 38). In this case, the simplicity criterion helps us to decide between the two hypotheses because there is no dispute about the existence of the facts of the matter about electrons. Thus, simplicity considerations, according to the sceptic, work only if the problem was epistemological, not metaphysical: "Against *this* [metaphysical] claim simplicity considerations are irrelevant. Simplicity considerations would have been relevant against a sceptic who argued that the indirectness of our access to the facts of meaning and intention *prevents us ever from knowing* whether we mean plus or quus. But

provide us with a constructive suggestion about the existence of the facts as to what a speaker means by his words and, thereby, fails to meet the sceptic's first condition. [13]

The second problem with this response concerns the sceptic's second condition about the facts about meaning: they must justify our future uses of the words, that is, they must tell us how we ought to use the words in the future. "Recall that a fact as to what I mean now is supposed to *justify* my future actions, to make them *inevitable* if I wish to use words with the same meaning with which I used them before" (Kripke, 1982, p. 40). The sceptic, however, thinks that "no 'hypothetical' state could satisfy such a requirement" (1982, p. 40). According to the sceptic, when we take the fact that the speaker means *plus* by "plus" as a hypothesis, which is claimed to be the simplest one, we neglect the essential normative feature of meaning, since in this case the best we can say is that "in the future I can only proceed hesitatingly and hypothetically, *conjecturing* that I probably ought to answer '68+57' with '125' rather than '5'. Obviously, this is not an accurate account of the matter" (1982, p. 40). [14] Hence, appealing to the simplicity criterion, according to the sceptic, fails to provide a candidate fact that meets the two conditions imposed on candidate meaning facts.

---

such merely epistemological scepticism is *not* in question. … He claims that an omniscient being, with access to *all* available facts, still would not find any fact that differentiates between the plus and the quus hypotheses" (1982, p. 39).

[13] However, this objection from the sceptic seems to beg the question against the response from simplicity. One might ask: What does assure us that our hypotheses about electrons are *genuine*? Can't we run the sceptical argument in the case of "electrons"? How do we know that there are indeed the facts of the matter about them? Here, the sceptic presupposes, rather than arguing for, the sceptical claim that there is no fact of the matter about meaning. On the basis of this presupposition, he rejects the response from simplicity, according to which the fact that the hypothesis that the speaker means *plus* by "plus" is the simplest one can constitute the fact about what the speaker means by his word. It seems that if the sceptic wants to argue that this response is wrong, he has to argue that there is no fact of the matter about meaning and, thereby, there can be no genuine hypotheses about it. The sceptic, however, presupposes this claim. Wright (1984) makes this point clear: "It is only *after* the skeptical argument has come to its conclusion that the skeptic is entitled to the supposition that there is indeed no such fact of the matter. In the course of the argument, *he* cannot assume as much without begging the question" (1984, p. 773, fn. 5).

[14] Thus, according to the sceptic, a proper account of the normativity of meaning must take into account the fact that "I immediately and unhesitatingly calculate '68 + 57' as I do, and the meaning I assign to '+' is supposed to *justify* this procedure. I do not form tentative hypotheses, wondering what I should do if one hypothesis or another were true" (1982, p. 40).

1.2.8. Meaning as a Unique, Irreducible Qualitative Experience

Another response to the sceptical challenge would be to "argue that "meaning addition by 'plus'" denotes an irreducible experience, with its own special *quale*, known directly to each of us by introspection" (Kripke, 1982, p. 41). According to this suggestion, meaning something by a word is to be taken to be a sort of unique experience, which each speaker has when he uses a word. Conscious experiences are said to have their own special qualities, which are known as "qualia". The experience of a headache is a subjective experience that a subject has, and it has its own unique quality, which can be characterized as follows: the subject knows what it is like to be in that state, i.e. to have a headache. Similarly, the subject knows how it feels to be happy, how it feels to be anxious, and the like. These experiences are known to the person experiencing them by introspection. According to the suggestion under consideration, meaning something by a word is just like having a unique experience: "Each of us knows immediately and with fair certainty that he means addition by 'plus', presumably the view in question assumes we know this in the same way we know that we have headaches – by attending to the 'qualitative' character of our own experiences" (Kripke, 1982, p. 41). The experience of meaning addition by "plus" is supposed to have its own irreducible quality, which distinguishes it from other states. This fact about the speaker, hence, is offered as an irreducible fact about the speaker, not reducible to the facts about the speaker's dispositions, behaviour, mental images, intentions, and the like. As Kripke elaborates, "maybe I appear to be unable to reply just because the experience of meaning addition by 'plus' is as unique and irreducible as that of seeing yellow or feeling a headache, while the sceptic's challenge invites me to look for another fact or experience to which this can be reduced" (1982, p. 41).

According to the sceptic, however, this candidate fact fails to determine what the speaker meant by his words in the past and fails to capture the normative character of meaning. First of all, what the speaker has experienced so far can be taken to be compatible with his meaning *plus* by "plus" and his meaning *quus* by "plus", just as the speaker's behaviour could be taken to be compatible with both of these sceptical hypotheses. Regarding our example of "57 + 68", the speaker has never thought about or experienced attempting to answer to this query before, since, supposedly, 56 has been the largest number he has ever confronted. As the sceptic asks, "I had never thought of this particular addition before: is not an interpretation of the '+' sign as quus compatible with everything I thought?" (1982, p. 41). Before having this particular experience, all

of my past experiences have been compatible with the hypothesis that "plus" means *quus*. Also, suppose that when I think of "+", I have a special sort of feeling, analogous to what I have when I have a headache. The sceptic again asks: "How on earth would this headache help me figure out whether I ought to answer '125' or '5' when asked about '68 + 57'?" (1982, p. 42). The fact that I have a special sort of experience cannot determine how I ought to use the word in new cases. For instance, if it is claimed that, because of the feeling that I now have I should answer "125", the sceptic would reply that your experience or feeling is indeed telling you that you ought to answer "5". It is, "once again, a rule for interpreting a rule. No internal impression, with a *quale*, could possibly tell me in itself how it is to be applied in future cases" (Kripke, 1982, p. 43).

1.2.9. Meaning Something as a Primitive Fact

A different, yet irreducible, fact about the speaker is suggested by those who believe that the fact that the speaker means something by his words is itself a primitive fact, which must not be compared to the facts about the speaker's qualitative experiences. As the sceptic introduces this view,

> perhaps … meaning addition by 'plus' is a state even more *sui generis* than we have argued before. Perhaps it is simply a primitive state, not to be assimilated to sensations or headaches or any 'qualitative' states, nor to be assimilated to dispositions, but a state of a unique kind of its own. (1982, p. 51)

Hence, this response can be put in the following simple way: meaning something by words cannot be explained in terms of any other facts about the speaker; rather, it itself is a primitive fact about what the speaker means by his words in each new case of using them.

According to the sceptic, this suggestion "seems desperate" (1982, p. 51), since it does not illustrate what such a state is, that is, it leaves the nature of this state mysterious:

> It leaves the nature of this postulated primitive state – the primitive state of 'meaning addition by "plus"' – completely mysterious. It is not supposed to be an introspectible state, yet we supposedly are aware of it with some fair degree of certainty whenever it occurs. For how else can each of us be confident that he *does*, at present, mean addition by 'plus'? (1982, p. 51)

This argument is sometimes called "the Argument from Queerness".[15] According to this

[15] See, e.g. (Boghossian, 1989, p. 178).

argument, taking the fact that the speaker means *plus* by "plus" to be an irreducible primitive fact about the speaker will lead to the problem of explaining what this fact is and how we know about it. The suggested view above does not provide us with any constructive answer to this problem. What sort of state is this state, which is supposedly present in a potentially infinite number of new cases of use? If this state is not supposed to be introspectible, then how do we know it? According to the sceptic,

> such a state would have to be a finite object, contained in our finite minds. It does not consist in my explicitly thinking of each case of the addition table, nor even of my encoding each separate case in the brain: we lack the capacity for that. Yet [as Wittgenstein states] (§ 195) "in a queer way" each such case already is "in some sense present". (1982, p. 52)

As Boghossian (1989) puts it, "the first charge is that we would have no idea how to explain our ability to know our thoughts, if we endorsed a non-reductionist conception of their content" (1989, p. 179). If the primitive state of meaning something by words is not introspectible with its own special quality, then it is difficult to explain how we know such a state.[16] The second objection from the sceptic was that "it is utterly mysterious how there could be a finite state, realized in a finite mind, that nevertheless contains information about the correct applicability of a sign in literally no end of distinct situations" (Boghossian, 1989, pp. 179-180).[17] As Wright (1984) puts the problem from queerness, "How can there be a state which each of us knows about, in his own case at least, noninferentially and yet which is infinitely fecund, possessing specific directive content for no end of distinct situations?" (1984, p. 775). [18]

[16] However, Boghossian thinks that what we are asked to do is indeed to solve the difficult problem of self-knowledge, i.e. the fact that we have noninferential knowledge of our thoughts. But if that is true, then what the sceptic must do is to provide us with "a *proof* that no satisfactory epistemology was ultimately to be had" (Boghossian, 1989, p. 179). That is to say, the sceptic cannot reject this non-reductionist response just by claiming that it leaves us with the problem of self-knowledge. Rather, he needs to argue that the problem of self-knowledge *cannot* be solved. KW's sceptic, however, does not provide such an argument. As we will discuss in Chapter Five, Davidson and Wright both try to deal with this problem, though in different ways.

[17] As Boghossian emphasizes again, the sceptic has not provided us with any argument in favor of the claim that it is not possible that there exist states of mind with such general contents: "What Kripke needs, if he is to pull off an argument from queerness, is some substantive argument, distinct from his anti-reductionist considerations, why we should not countenance such states" (1989, p. 180).

[18] Wright takes this problem seriously and tries to provide an answer to it (see Wright, 1984; 2001, especially essays 4-7). We will also see, in Chapter Five, how Wright deals with this "Generality of Content" problem.

1.2.10. Fregean Senses

Another suggestion comes from the Platonists or mathematical realists who believe that mathematical entities are not mental entities and, hence, cannot be said to be present in the mind of the speaker; rather, they are objective entities in the world. According to this view, "the addition function is not in any particular mind, nor is it the common property of all minds. It has an independent, 'objective', existence" (Kripke, 1982, p. 53). If that is true, then what we mean by "+" is a matter of the word referring to an objective entity, i.e. the addition function. Its meaning has nothing to do with certain internal goings-on in the speaker's mind, which supposedly make it the case that "plus" means *plus*. Hence, we can say that "+" has a sense, which is objective and independent of the speaker's subjective thoughts. The sceptic's demand, hence, is blocked, that is, the demand of introducing a fact about the speaker as to what he means by "plus", or as to which function "+" refers to. Moreover, we can explain the fact that the addition function has as its consequence the sum of indefinitely many numbers, no matter how large they are, since it can be said that such a mathematical object is an infinite abstract object. It is simply in its nature to be so.

The sceptic's problem with this suggestion is put in the following question: How can we have access to such an objective mathematical object? It is the speaker who has to grasp such an objective sense, and it seems that our grasp of the sense of the plus sign relies on, or is mediated by, some "idea" in the mind of the speaker. The sceptic uses Frege's analysis of our use of the plus sign to clarify this point:

> Frege's analysis of the usage of the plus sign by an individual posits the following four elements: (a) the addition function, an 'objective' mathematical entity; (b) the addition sign '+', a linguistic entity; (c) the 'sense' of this sign, an 'objective' abstract entity like the function; (d) an idea in the individual's mind associated with the sign. (1982, p. 54)

While the sense of the plus sign is, in its nature, objective and the same for all individuals, each individual speaker has an idea in mind regarding this sign, which is subjective and private: "Each such individual grasps this sense by virtue of having an appropriate idea in his mind. The 'sense' in turn determines the addition function as the referent of the '+' sign" (Kripke, 1982, p. 54). Primarily, there is no essential problem in saying that the sense of "+" determines its referent, that is, the objective addition function as a mathematical entity. The problem, however, appears when we ask "how the existence in my mind of any mental entity or idea can *constitute* 'grasping' any

particular sense rather than another" (Kripke, 1982, p. 54). If it is through having a certain subjective idea in mind that the sense of the plus sign can be grasped, then the sceptic can claim that this idea is indeed fixing the quaddition function as the referent of "+", rather than the addition function. "The idea in my mind is a finite object: can it not be interpreted as determining a quus function, rather than a plus function?" (Kripke, 1982, p. 54). Indeed, the sceptic's sceptical challenge just re-emerges in this case in the form of the question what makes it the case that I grasp this sense of "+" rather than another. Although the senses and references are objective and independent of the individual, the idea in the mind of the individual is subjective, and no one can grasp the sense of the plus sign without having such an idea in mind. The problem with ideas, however, is similar to the problem with internal rules, instructions, mental images, and the like: they can be interpreted in non-standard ways. They cannot constitute the fact that the speaker means *plus* by "plus", or that he takes the plus sign to denote the addition function.

1.3. The Sceptical Conclusions

The sceptic has criticized and rejected the suggested candidate facts, which were introduced to deal with his sceptical problem. It is important to note that the sceptical challenge can be expanded throughout our language and applied to all available terms and expressions, from apparently truth-conditional sentences (such as "57 + 68 = 125") to predicates (such as "is a table", "is red", "is good", etc.), all nouns and names (such as "Pacific Ocean", "Guinness", "John", etc.), indexicals (such as "I", "He", "That", and so on), and even non-truth-conditional sentences (such as imperatives, interrogatives, and exclamatory expressions). As an example, the sceptic can claim that, when we use the imperative expression "Trust me!", what we really meant by "trust" in the past, and hence in the present, is *schomrust*, according to which, until *t*, apply "trust" to those whom you trust; after *t*, apply it to those whom you mistrust. As another example, the indexical "I" might be interpreted by the sceptic to be meaning *schwai*, in which case, up to the time *t*, "I" refers to me; afterwards, to any second person. Hence, the sceptical problem is indeed completely general.

According to the sceptic, then, we cannot find any fact about the speaker determining what he meant by "plus". And, if there is no fact as to what he meant by his words in the past, there is no fact as to what he means by his words now. Hence, if, following the

sceptic, we accept Traditional Realism's conception of meaning, according to which there are certain rules determining the correct use of words for us or there are certain meaning facts determining the correctness conditions for our words, then the sceptic has argued that there is no correct or incorrect use at all, since no fact about the speaker can determine which rule he is following or constitute the fact that he means one thing rather than another. Thus, on Traditional Realism's conception of meaning, without such rules or meaning facts, there would be no meaning at all. The sceptic puts the conclusion of his argument as follows:

> The sceptical argument, then, remains unanswered. There can be no such thing as meaning anything by any word. Each new application we make is a leap in the dark; any present intention could be interpreted so as to accord with anything we may choose to do. So there can be neither accord, nor conflict. (1982, p. 55)

Following Wilson (1998, pp. 107-109), we can distinguish between two sorts of sceptical conclusion, which the sceptical argument establishes: the "Basic Sceptical Conclusion" (BSC) and the "Radical Sceptical Conclusion" (RSC). The sceptic, after criticizing and rejecting the candidate facts, concluded that no fact about the speaker can be found which determines what he means by his words. This is the Basic Sceptical Conclusion of the sceptic's argument:

(BSC) There is no fact of the matter as to what someone means by his words.

However, the Radical Sceptical Conclusion led to a more radical claim:

(RSC) There is no such thing as meaning anything by any word at any time.

According to Wilson (1994), KW's sceptic accepts both the Basic Sceptical Conclusion and the Radical Sceptical Conclusion, whereas KW himself accepts the Basic Sceptical Conclusion but rejects the Radical Sceptical Conclusion as "insane and intolerable" (see Wilson, 1994, pp. 370-373). This distinction is important especially for our discussion of KW's sceptical solution.

1.4. KW's Sceptic's Sceptical Argument

Earlier in this chapter (Section 1.1.2), we saw that Traditional Realism can be characterized in two ways: (I) as taking for granted the rule-following picture of meaning and (II) as presupposing the existence of meaning facts. In this regard, we can introduce two strands of the sceptical argument, which, by proceeding through different

Basic Sceptical Conclusions, establish the same Radical Sceptical Conclusion that there is no such thing as meaning anything by any word.

1.4.1. The First Strand of the Sceptical Argument

The first version of the sceptical argument takes the first characterization of Traditional Realism as its first premise. Indeed, this version of Traditional Realism can be taken to be one characterization of the normativity of meaning thesis, which the sceptic took for granted at the beginning of his argument. As Boghossian characterizes the normativity thesis, "if I mean something by an expression, then the potential infinity of truths that are generated as a result are *normative* truths: they are truths about how I *ought* to apply the expression, if I am to apply it in accord with its meaning, not truths about how I *will* apply it" (Boghossian, 1998, p. 143). Regarding our example of "green", Boghossian clarifies this thesis as follows:

> Suppose the expression 'green' means *green.* It follows immediately that the expression 'green' applies *correctly* only to *these* things (the green ones) and not to *those* (the non-greens). The fact that the expression means something implies, that is, a whole set of *normative* truths about my behaviour with that expression: namely, that my use of it is correct in application to certain objects and not in application to others. (1998, p. 148)

Meaning something by a word determines the correctness conditions for the application of that word. The correctness conditions for the application of "green", if it is supposed to mean *green*, determine that it is correctly applicable to an object if and only if that object is green. However, the correctness conditions of "green", if it means *grue*, state that it is correctly applicable to an object if and only if that object is green, up until time *t*, and blue, after *t*. Hence, the normativity thesis can be summarized as the thesis stating that "meaningful expressions possess conditions of *correct use*" (Boghossian, 1998, p. 148). In this regard, we can say that our first characterization of Traditional Realism, and, hence, the first premise of the first strand of the sceptical argument, is a characterization of the normativity of meaning thesis in terms of the *rule-following picture of meaning*, which we will call "$\mathbf{N_{TR1}}$": if the speaker, S, means *green* by "green", then there is a rule determining the correct use of the word for him.

The sceptic challenged this claim. His "grounding constraint" (Wilson, 1998, p. 107), or his "Ground Demand" ($\mathbf{GD_{KW1}}$), was that the proponents of Traditional Realism must

introduce a fact about the speaker, which constitutes the fact that the speaker indeed follows R, rather than R*, with R* stating that "green" applies to green things before the time *t*, and applies to blue things after *t*. The sceptic argued that there can be no such fact about the speaker. This strand of the sceptic's sceptical argument, hence, can be characterized as follows:

(I) **N_{TR1}**: If S means *green* by his utterance of "green", then there is a rule, R, (e.g. that "green" applies to green things and does not apply to blue things), which determines the correct application of "green" for S.

(II) **GD_{KW1}**: If there is such a rule, R, then there is some fact about S's psychological, behavioral, or social states that makes R, rather than R* (e.g. that "green" applies to green things before the time *t*, and applies to blue things after *t* (in which case "green" means *grue*)), as what determines the correct use of "green" for S.

(III) **BSC_{KW1}**: There is no such a fact about S as to whether he follows R or R* [This is established by the arguments considered in Sections 1.2.1 – 1.2.10].

It would follow from these premises that the speaker does not mean anything by his utterance of "green". Since the argument generalizes to all words and expressions, it follows that:

(IV) **RSC_{KW1}**: There is no such thing as meaning anything by any word.

Because there is no fact as to what rule the speaker follows (**BSC_{KW1}**), and because Traditional Realism took for granted the assumption that if the speaker means something by a word, then he follows a rule determining the correct use of the word (**N_{TR1}**), we can reach the Radical Sceptical Conclusion that there is no correct and incorrect use of a word and hence no meaning at all (**RSC_{KW1}**).

1.4.2. The Second Strand of the Sceptical Argument

The first premise in the second strand of the sceptical argument contains our second characterization of Traditional Realism (**N_{TR2}**): if the speaker means *green* by "green", then there is a state of affairs, or a possible meaning fact, that the speaker means *green* by "green", *and not anything else*, such as *grue*, which determines the correct

application of the word for the speaker. The sceptic's Ground Demand (**GD_{KW2}**), in this strand, can be characterized as follows: if there are such meaning facts, then we should be able to cite some facts about the speaker, which rule out any other sceptical hypotheses, for instance, the sceptical hypothesis that it is the state of affairs that "green" means *grue* that determines the correct application of "green" for the speaker. In this regard, we can characterize the second strand of the sceptical argument as follows:

(I) **N_{TR2}**: If S means *green* by his utterance of "green", then there is some possible meaning fact or state of affairs, F, (e.g. that S means *green*, and does not mean *grue*, by "green") that determines the correct application of "green" for S.

(II) **GD_{KW2}**: If there is such a possible meaning fact or state of affairs, F, then there is some fact about S's psychological, behavioral, or social states that constitutes F, rather than F* (e.g. that S means *grue*, and does not mean *green*, by "green") as the one determining the correct application of "green" for S.

(III) **BSC_{KW2}**: There is no such a fact about S constituting F. [Again, this is established by the arguments considered in Sections 1.2.1 – 1.2.10].

As before, the argument generalizes. Hence,

(IV) **RSC_{KW2}**: There is no such thing as meaning anything by any word.

The sceptic argued that no fact about the speaker can constitute one fact, rather than another, as what determines the correct application of "green" for the speaker (**BSC_{KW2}**). If that is true, and if Traditional Realism insists on the idea that if a speaker means something by his words, then there are such meaning facts (**N_{TR2}**), we can conclude that there is no such thing as meaning anything by any word at all (**RSC_{KW2}**).

Therefore, the sceptic, via his sceptical argument, showed that a certain conception of the practice of meaning something by words, which Traditional Realism advocates, leads to the Radical Sceptical Conclusion.

1.5. The Positive Part: Kripke's Wittgenstein's Sceptical Solution

The Radical Sceptical Conclusion of the sceptic's argument, however, is obviously wrong, since we indeed mean something specific by our words every day. As KW emphasizes, the Radical Sceptical Conclusion is "insane and intolerable" (1982, p.60) and "incredible and self-defeating" (1982, p. 71). KW, hence, cannot be in agreement with the sceptic's radical claim that there is no such thing as meaning anything by any word because, in that case, how could the sceptical problem and conclusions themselves have any meaning at all? As Kripke puts it, his Wittgenstein does not "wish to leave us with his problem, but to solve it: the sceptical conclusion is insane and intolerable" (1982, p. 60). However, what sort of solution does KW have in mind?

1.5.1. Straight Solution vs. Sceptical Solution

First of all, Kripke introduces two sorts of solution that can be given in response to a sceptical problem: straight solutions and sceptical solutions. According to Kripke, "call a proposed solution to a sceptical philosophical problem a *straight* solution if it shows that on closer examination the scepticism proves to be unwarranted; an elusive or complex argument proves the thesis the sceptic doubted" (1982, p. 66). In other words, we have a straight solution to the sceptical problem if we can directly rule out the sceptic's sceptical hypotheses by introducing some fact about the speaker, which determines which rule the speaker follows or constitutes the fact that the speaker means *plus* by "plus", and not anything else. We can take the ten candidate facts introduced in the previous section as examples of attempts to provide a straight solution to the sceptical problem. The sceptic, however, already rejected such straight solutions.

Kripke defines a "sceptical" solution to a sceptical problem as follows:

> A *sceptical* solution of a sceptical philosophical problem begins on the contrary by conceding that the sceptic's negative assertions are unanswerable. Nevertheless our ordinary practice or belief is justified because – contrary appearances notwithstanding – it need not require the justification the sceptic has shown to be untenable. And much of the value of the sceptical argument consists precisely in the fact that he has shown that an ordinary practice, if it is to be defended at all, cannot be defended in a certain way. (1982, p. 67)

In the present case, the sceptical solution begins by rejecting the sceptic's Radical Sceptical Conclusion and accepting the Basic Sceptical Conclusion: although it is true that there is no fact as to what a speaker means by his words, it is not the case that there

is no such thing as meaning anything by any word. The success in the practice of meaning something by words cannot be explained in terms of the facts Traditional Realism introduced; rather, we need to search for a different sort of justification for such a practice. KW, hence, agrees with the sceptic that there is no special, superlative fact about meaning because, if there was any, we would be able to find it. As Kripke clarifies,

> [KW's] solution to his own sceptical problem begins by agreeing with the sceptics that there is no 'superlative fact' (§192) about my mind that constitutes my meaning addition by 'plus' and determines in advance what I should do to accord with this meaning. But, he claims (in §§183-93), the appearance that our ordinary concept of meaning demands such a fact is based on a philosophical misconstrual – albeit a natural one – of such ordinary expressions as 'he meant such-and-such', 'the steps are determined by the formula', and the like. (1982, pp. 65-66)

The "philosophical misconstrual", however, is nothing but Traditional Realism's conception of the practice of meaning something by words. As a result, KW believes that the first step towards a solution to the sceptical problem is to reject Traditional Realism, which has brought about the unacceptable Radical Sceptical Conclusion. According to KW,

> any construal that looks for something in my present mental state to differentiate between my meaning addition or quaddition, or that will consequently show that in the future I should say '125' when asked about '68 + 57', is a misconstrual and attributes to the ordinary man a notion of meaning that is refuted by the sceptical argument. (1982, p. 66)

Therefore, we can put the sceptical solution, regarding the *first* strand of the sceptical argument, as follows:

(I) **GD_{KW1}**: If there is a rule, R (e.g. that "green" applies to green things and does not apply to blue things), which determines the correct application of "green" for S, then there is some fact about S's psychological, behavioral, or social states that makes R, rather than R* (e.g. that "green" applies to green things before the time *t*, and applies to blue things after *t*), to be what determines the correct use of "green" for S.

(II) **BSC_{KW1}**: There is no fact about S as to whether he follows R or R*.

(III) **$\sim RSC_{KW1}$**: It is not the case that there is no such thing as meaning anything by any word.

(IV) **~N_{TR1}**: Therefore, it is not the case that if S means *green* by his utterance of "green", then there is a rule, such as R, that determines the correct application of "green" for S.

The sceptical solution starts with accepting the Basic Sceptical Conclusion, that there is no fact about the speaker as to which rule he follows, and rejecting the Radical Sceptical Conclusion. The first characterization of Traditional Realism, as a misconstrual of our ordinary conception of meaning, is hence rejected by KW: it is not the case that meaning something by words can be explained in terms of the assumption that there exist certain rules telling us how we ought to use the words in the future.

Similarly, regarding the *second* strand of the sceptical argument, the sceptical solution can be characterized as follows:

(I) **GD_{KW2}**: If there is a possible meaning fact or state of affairs, F (e.g. that S means *green*, and does not mean *grue*, by "green"), which determines the correct application of "green" for S, then there is some fact about S's psychological, behavioral, or social states that constitutes F, rather than F* (e.g. that S means *grue*, and does not mean *green*, by "green") as the one determining the correct application of "green" for S.

(II) **BSC_{KW2}**: There is no such a fact about S constituting F.

(III) **~RSC_{KW2}**: It is not the case that there is no such thing as meaning anything by any word.

(IV) **~N_{TR2}**: Therefore, it is not the case that if S means *green* by his utterance of "green", then there is some possible meaning fact or state of affairs, such as F, that determines the correct application of "green" for S.

The sceptical solution, again, rejects the Radical Sceptical Conclusion, but accepts the Basic Sceptical Conclusion, i.e. that there is no fact about the speaker constituting the state of affairs that the speaker means one thing rather than another by his words. As Kripke points out,

> [his Wittgenstein] does not give a 'straight' solution, pointing out to the silly sceptic a hidden fact he overlooked, a condition in the world which constitutes my meaning addition by 'plus'. In fact, he agrees with his own hypothetical sceptic that there is no such fact, no such condition in either the 'internal' or the 'external' world. (1982, p. 69)

As a result, KW rejects the second characterization of Traditional Realism too. However, what is KW's sceptical solution?

1.5.2. The Sceptical Solution

The sceptical solution, as we saw above, begins with rejecting Traditional Realism, which, according to Kripke, the early Wittgenstein himself advocated in the *Tractatus*. For Kripke, the main doctrines of the *Tractatus* can be outlined as follows:

> To each sentence there corresponds a (possible) fact. If such a fact, obtains, the sentence is true; if not, false. For atomic sentences, the relation between a sentence and the fact it alleges is one of a simple correspondence or isomorphism. The sentence contains names, corresponding to objects. An atomic sentence is itself a fact, putting the names in a certain relation; and it says that (there is a corresponding fact that) the corresponding objects are in the same relation. Other sentences are (finite or infinite) truth-functions of these. (1982, p. 71)

In this sense, KW can be seen to be denying, and offering an alternative to, his earlier conception of meaning presented in the *Tractatus*, which can be called the "Truth-Conditional Conception of Meaning". This conception, however, is directly related to the way we characterized Traditional Realism. According to Wittgenstein's earlier remarks, sentences are true or false in terms of the certain facts which, if they obtain, make the sentences true. Such facts can be taken to be the sentences' truth-conditions, which provide us with a way to specify their meaning. For instance, the sentence "Snow is white" is true if and only if the state of affairs that snow is white obtains, that is, if and only if snow is white. As a result, to understand what this sentence means, we need to know under what conditions it is true. As Kripke points out, "the simplest, most basic idea of the *Tractatus* can hardly be dismissed: a declarative sentence gets its meaning by virtue of its *truth conditions*, by virtue of its correspondence to facts that must obtain if it is true" (1982, p. 72). The sentence "the cat is on the mat" is understood, according to Kripke's reading of the *Tractatus*, "by those speakers who realize that it is true if and only if a certain cat is on a certain mat; it is false otherwise. The presence of the cat on the mat is a fact or condition-in-the-world that would make the sentence true (express a truth) if it obtained" (1982, p. 72). KW's sceptic argued against such an account of meaning by showing how it leads to the Radical Sceptical Conclusion. According to Kripke, the later Wittgenstein, in his *Philosophical Investigations*, proposes a new picture of the practice of meaning something by words, according to which we must

replace the question "What must be the case for [a] sentence to be true?" by two other questions: first, "Under what conditions may this form of words be appropriately asserted (or denied)?" and, second, "What is the role, and the utility, in our lives of our practice of asserting (or denying) the form of words under these conditions?" (1982, p. 73). KW's conception of meaning concentrates on our ordinary practice of asserting certain sentences under certain conditions, instead of searching for their truth-conditions. As Kripke emphasizes, "we can say that Wittgenstein proposes a picture of language based, not on *truth conditions*, but on *assertability conditions* or *justification conditions*" (1982, p. 74). To understand KW's alternative view, it is important to see what he means by "assertability conditions" or "justification conditions".

1.5.3. Assertability Conditions

Assertability conditions are the conditions under which a speaker can justifiably assert something specific by using a certain sentence. For instance, we may use the sentence "It rained last night" to assert that it rained last night. When are we justified in using this sentence to assert what we did? It seems that we are justified in asserting a sentence if we have enough evidence to do so. For instance, the sentence "It rained last night" can be justifiably asserted if there is enough evidence that it rained last night, such as looking at the street and seeing that it is wet, and the like. Under such conditions, i.e. the assertability conditions, we are justified in asserting that it rained last night. To this extent, Traditional Realism and KW's sceptical solution are in agreement: both agree that it is under specific conditions that certain sentences can be asserted. Traditional Realists, however, do not stop at this point; rather, they offer an explanation of *why* such assertability conditions are as they are. According to their view, the fact that the sentence has the assertability conditions it does is *explained* in terms of the fact that a certain state of affairs, i.e. the sentence's truth-condition, obtains, that is, that it rained last night. Traditional Realists believe that, if the truth-condition obtains, then it is causally responsible for the evidence, such as the street's being wet, on the basis of which we will be justified in asserting that it rained last night. KW's sceptic, however, rejected any view which appeals to the existence of such truth-conditions in order to justify our linguistic practices. Hence, his sceptical solution cannot rely on explaining the existence of such assertability conditions in terms of the existence of truth-conditions or states of affairs.

KW's sceptical solution is concerned with meaning-ascribing sentences, such as "Jones means *green* by 'green'", by which we ascribe meanings to others. The sceptic argued that such sentences cannot be said to have any truth-conditions, since this claim will lead to the Radical Sceptical Conclusion: "Recall Wittgenstein's sceptical conclusion: no facts, no truth conditions, correspond to statements such as "Jones means addition by '+'"" (Kripke, 1982, p. 77). We cannot say that our justification in asserting that John means *green* by "green" has anything to do with the obtaining of certain facts about John's external behaviour or internal states. None of such facts could constitute the fact that he means *green* by "green". There is hence no fact about the speaker, no state of affairs, capable of making the sentence "Jones means *green* by 'green'" true.

Nonetheless, in our ordinary linguistic practices, we ascribe meanings to others by asserting sentences like "Jones means *green* by 'green'", "Michael means *plus* by 'plus'", and the like. According to KW's sceptical solution, "all that is needed to legitimize assertions that someone means something is that there be roughly specifiable circumstances under which they are legitimately assertable, and that the game of asserting them under such conditions has a role in our lives" (1982, pp. 77-78). When can we say, according to the sceptical solution, that someone means something by a word? In KW's view, we first of all need to look at how meaning-ascribing sentences are usually used: "Wittgenstein's sceptical solution concedes to the sceptic that no 'truth conditions' or 'corresponding facts' in the world exist that make a statement like "Jones, like many of us, means addition by '+'" true. Rather we should look at how such assertions are used" (1982, p. 86). As in the case of using the sentence "It rained last night" to assert that it rained last night, here too we need to have certain evidence in order to use the meaning-ascribing sentence "Jones means *green* by 'green'" to assert that Jones means *green* by "green". The evidence, in this case, is given by observation of the linguistic behaviour of Jones, that is, the way he is inclined to use "green" in certain publically observable situations. We can justifiably assert that "Jones means *green* by 'green'" if we can observe, in enough cases, that Jones uses his word as others in his speech community do. For instance, if, in enough cases, we observe that Jones is inclined to apply "green" to the things to which we would be inclined to apply "green", then we can justifiably assert that Jones means *green* by "green", that is, that he means what we mean by the word. Making such assertions, hence, essentially depends on observing the agreement between our responses and the speaker's responses.

It is important to note that the sceptical argument rejected the idea that the speaker

uses a word in a certain way *because* he has already grasped a certain meaning, concept, or rule. One consequence of the sceptical argument was that speakers seem to follow their blind inclinations to use words in one way rather than another. As Kripke emphasizes, "the entire point of the sceptical argument is that ultimately we reach a level where we act without any reason in terms of which we can justify our action. We act unhesitatingly but *blindly*" (1982, p. 87). This means that, in the picture that the sceptical solution proposes, "a speaker may, without ultimately giving any justification, follow his own confident inclination that this way (say, responding '125') is the right way to respond, rather than another way (e.g. responding '5')" (Kripke, 1982, pp. 87-88). The assertability conditions are generally about the conditions under which a speaker feels confident to use a word in a certain way: "The 'assertability conditions' that license an individual to say that, on a given occasion, he ought to follow his rule this way rather than that, are, ultimately, that he does what he is inclined to do" (1982, p. 88). For any speaker, there are conditions under which he thinks a sentence is assertable. The point of the sceptical solution is that, without a community of speakers agreeing in their linguistic practices, e.g. the way they use their words, there would be no assertability conditions available at all. In other words, for a solitary speaker considered in isolation, there would be no distinction between correctness and incorrectness, and hence there would be no meaning at all. In this sense, the assertability conditions essentially involve the shared practices of a community of speakers. The best way to understand this point is to compare the case of a speaker considered in isolation with the case of a speaker as a member of a speech community.

1.5.4. The Role of Agreement and the Private Language

Suppose that Jones, the solitary speaker, is inclined to use "green" in a certain way. He thinks that he ought to apply the word to certain objects, e.g. the object *a*. In this case, we can observe that the assertability conditions for the following two sentences are the same:

(S1) Jones believes that "green" ought to be applied to *a*.

(S2) "Green" ought to be applied to *a*.

What would be the assertability conditions for (S1)? It could be asserted that Jones believes that "green" ought to be applied to *a* only if Jones is inclined to apply "green"

to *a*, no matter what *a* is. The reason is that Jones is entirely alone: Who but Jones himself is there to decide to which object his word ought to be applied? Similarly, the assertability conditions for asserting (S2) would be the same: the correct use of "green", that is, how it ought to be applied, is determined by Jones himself, that is, in terms of the way he is inclined to apply the word. The reason again is that nothing but Jones' inclination to use his word is there to determine the correct application of "green". Hence, whatever seems to Jones, the solitary speaker, to be correct will be correct: there is no distinction between how he is inclined to use the word and the correct way of using the word. As Kripke says, for a solitary person, "there are no circumstances under which we can say that, even if he inclines to say '125', he *should* have said '5', or *vice versa*" (1982, p. 88). The answer "5", "125", or any other answer the solitary person gives is "the answer that strikes him as natural and inevitable" (Kripke, 1982, p. 88). Therefore, for a solitary person, there is no answer to the question "under what circumstances can he be wrong, say, following the wrong rule?" (Kripke, 1982, p. 88).

However, suppose that Jones becomes a member of a speech community and interacts with other speakers of that community. In this case, the assertability conditions for asserting the mentioned two sentences significantly diverge: (S1) can be justifiably asserted if Jones is inclined to apply "green" to *a*, while asserting (S2) is justified if his speech community is inclined to apply "green" to *a*. As a consequence, it would not be the case that whatever the speaker thinks to be correct is always correct. The speaker may be inclined to apply "green" to a blue object, while others are not inclined to do so. As Kripke emphasizes, "others will then have justification conditions for attributing correct or incorrect rule following to the subject, and these will *not* be simply that the subject's own authority is unconditionally to be accepted" (1982, p. 89). Hence, by bringing into play a community of speakers agreeing in their linguistic practices, we can have a legitimate practice of ascribing meanings to each other. The practice of regarding uses of each other's words as correct or incorrect would not have any justification without adverting to the wide agreement in the speech community, that is, the fact that the members of such a community share their inclinations to use words in one way rather than another. According to Kripke, hence, "what do I mean when I say that the teacher judges that, for certain cases, the pupil must give the 'right' answer? I mean that the teacher judges that the child has given the same answer that he himself would give" (1982, p. 90). A member of my speech community observing my responses is in a position to judge whether or not my responses agree with his. On this basis, he can

judge whether I mean the same thing he does: "*Smith* will judge Jones to mean addition by 'plus' only if he judges that Jones's answers to particular addition problems agree with those *he* is inclined to give" (Kripke, 1982, p. 91). Consequently, if I fail to respond as others do, others will have no evidence to judge that I mean the same thing they do: "If Jones consistently fails to give responses in agreement (in this broad sense) with Smith's, Smith will judge that he does not mean addition by 'plus'"(Kripke, 1982, p. 91).

Therefore, it is a consequence of KW's sceptical solution that there can be no "Private Language". There would be no assertability conditions for a solitary person considered in isolation, since there would be no distinction between what he thinks to be correct and what is correct: "As long as we regard [the solitary speaker] as following a rule 'privately', so that we pay attention to his justification conditions alone, all we can say is that he is licensed to follow the rule as it strikes him" (Kripke, 1982, p. 89). As a result, there is no ground to claim that such a solitary speaker means anything by any word. For Kripke, it is the sceptical solution "that contains the argument against 'private language'; for allegedly, the solution will not admit such a language" (1982, p. 60).[19] Although after the sceptic's rejection of Traditional Realism, we were left with nothing but the speaker's blind inclinations to use words, a community of speakers agreeing in their inclinations to use words in certain ways provides the ultimate criterion for the assertability of meaning-ascriptions. In this sense, "the entire 'game' we have described – that the community attributes a concept to an individual so long as he exhibits sufficient conformity, under test circumstances, to the behaviour of the community – would lose its point outside a community that generally agrees in its practices" (Kripke, 1982, p. 96). Any claim about whether someone means something by a word has to be checked and confirmed by others in a speech community: "Any individual who claims to have mastered the concept of addition will be judged by the community to have done so if his particular responses agree with those of the community in enough cases, especially the simple ones" (Kripke, 1982, pp. 91-92). [20] It is membership in a speech

[19] Kripke prefers to talk about the "private model" of rule following, "that the notion of a person following a given rule is to be analyzed simply in terms of facts about the rule follower and the rule follower alone, without reference to his membership in a wider community" (1982, p. 109).

[20] As Kripke summarizes, "An individual who passes such tests is admitted into the community as an adder; an individual who passes such tests in enough other cases is admitted as a normal speaker of the language and member of the community. Those who deviate are corrected and told (usually as children) that they have not grasped the concept of addition. One who is an incorrigible deviant in enough respects simply cannot participate in the life of the community and in communication" (1982, p. 92)

community that provides us with justification for asserting meaning-ascribing sentences: “We say of someone else that he follows a certain rule when his responses agree with our own and deny it when they do not” (Kripke, 1982, p. 92). And, the main point about such a speech community is that they agree in their inclinations to respond in a certain way: “If there was no general agreement in the community responses, the game of attributing concepts to individuals … could not exist” (Kripke, 1982, p. 96). In this sense, the existence of such agreement in the “unhesitating responses we make” can be seen to be essential to the practice of attributing meaning to a speaker: “On Wittgenstein’s conception, such agreement is essential for our game of ascribing rules and concepts to each other” (Kripke, 1982, p. 96).

It is important to note that the sceptical solution in a way reverses the direction of explanation of meaning that Traditional Realism suggested. Traditional Realism claimed that a speaker can be said to use “green” correctly if he can be said to have grasped the rule determining the correct use of the word, or to have grasped the meaning of “green”, that is, if he means *green* by “green”. KW’s sceptical solution, instead, suggests that a speaker can be said to mean *green* by “green” only if others in a speech community can observe that he uses the word as they do, that is, if they accept the speaker as a reliable member of their community. To some extent, hence, Traditional Realism can be taken to be in agreement with the sceptical solution. According to Traditional Realism too, we can assert that Jones means *green* by “green” if there is legitimate evidence available, which is, in the case of asserting meaning-ascribing sentences, the evidence that the speaker applies the word in agreement with the way others do. However, the divergence between this view and the sceptical solution appears at the point in which Traditional Realism tries to offer an *explanation* of why Jones and others agree on the way “green” ought to be used: Jones and others use the word “green” in the way they do *because* they all mean *green* by “green”. As Kripke indicates, Traditional Realism states, although “others may have taught me the concept of addition”, “they acted only as heuristic aids to an achievement – the ‘grasping of the concept’ of addition – that puts me in a special relation to the addition function” (1982, p. 80). According to Traditional Realism, there are certain concepts, meanings, or rules, which we grasp and which determine the correct application of our words for us. This explanation, however, opens the path to the Radical Sceptical Conclusion, since no fact about the speaker could constitute the fact that he means *green*, and not *grue*, by “green”. Hence, as KW points out, “it is essential to our concept of a rule that we maintain some such conditional as

"If Jones means addition by '+', then if he is asked for '68 + 57', he will reply '125'"" (1982, p. 94). We cannot say, if we want to avoid the sceptical problem, that if Jones means *plus* by "plus", then he would respond in a certain way, e.g. by "125". Rather, what we can legitimately say is that, if Jones responds in a certain way, then we can state that he means what we do by "plus", that is, *plus*: "If Jones does not come out with '125' when asked about '68 + 57', we cannot assert that he means addition by '+'" (1982, p. 95).

Therefore, the essential difference between the "Sceptical Solution" and Traditional Realism's "Straight Solution" emerges when we consider that, according to the sceptical solution, there is no explanation of why there is such agreement in a speech community: the fact that we agree on how to use our words is just a brute empirical fact about our speech community. As Kripke states,

> there is no objective fact – that we all mean addition by '+', or even that a given individual does – that explains our agreement in particular cases. Rather our license to say of each other that we mean addition by '+' is part of a 'language game' that sustains itself only because of the brute fact that we generally agree. (Nothing about 'grasping concepts' guarantees that it will not break down tomorrow). (1982, p. 97)

There is no fact of the matter *explaining* the wide agreement there is in our responses in a speech community. Any attempt to explain such agreement in terms of our grasp of some existent concepts, meanings, rules, facts, states of affairs, and the like, would lead to the Radical Sceptical Conclusion. For this reason, "we cannot say that we all respond as we do to '68+ 57' *because* we all grasp the concept of addition in the same way, that we share common responses to particular addition problems *because* we share a common concept of addition" (Kripke, 1982, p. 97). Rather, it is just a brute empirical fact about our community that we generally agree in our practices. Our inclinations in a speech community "are to be regarded as primitive" (Kripke, 1982, p. 91).

1.5.5. The Role and Utility of Our Linguistic Practices

However, if we can assert the sentences like "Jones means *plus* by 'plus'" under certain assertability conditions, that is, when Jones' usage of the word agrees with ours, and if there is no explanation of the fact that we agree to use the word in this way, then what is the reason for having this form of the practice of attributing meaning to each other? Why does such a practice have the form it does? To answer to this question, KW draws

our attention to the role that the practice of attributing meaning to others under certain assertability conditions plays in our everyday lives. Although there is no explanation of agreement in terms of shared grasp of meaning, there is still a reason justifying our participation in such practices. We assert that Jones means *plus* by "plus" if he responds to enough arithmetic queries in the way we do. This form of the practice of attributing meaning to each other has many practical and pragmatic benefits. For instance, if a speaker can be accepted as a reliable member of a speech community, then he can receive the benefit of living in that community and interacting with others in many different ways, which are essential in our everyday lives. Kripke gives a simple example to clarify this point. Suppose that someone goes to a grocery store to buy some apples:

> The customer, when he … asks for five apples, expects the grocer to count as he does, not according to some bizarre non-standard rule; … he expects the grocer's responses to agree with his own. … Of course the grocer may make mistakes in addition; ... But as long as the customer attributes to him a grasp of the concept of addition, he expects that at least the grocer will not behave bizarrely, as he would if he were to follow a quus-like rule; and one can even expect that, in many cases, he will come up with the same answer the customer would have given himself. (1982, pp. 92-93)

If we consider the pragmatic role and utility of making attributions of meaning under the relevant conditions of agreement, then we can find out why we continue exercising such a practice as we do. Trusting each other, communicating our thoughts with other members, discussing issues, asking for help, buying and selling things, and so forth, are just simple examples of the vast benefits of living in a community in which its members attribute meaning to each other in the way KW described: "Our entire lives depend on countless such interactions, and on the 'game' of attributing to others the mastery of certain concepts or rules, thereby showing that we expect them to behave as we do" (1982, p. 93).[21]

The sceptical solution, hence, offered an alternative picture of the practice of meaning something by words. In this new view of meaning, "no supposition that 'facts correspond' to [meaning-ascribing] assertions is needed" (Kripke, 1982, p. 78). Whereas the Traditional Realist would have attempted to explain why our meaning ascriptions have the assertability conditions they do by referring to their truth-conditions, the sceptical solution attempts to explain this in terms of the practical utility of this practice.

[21] KW calls this way of responding to each other the "Form of Life": "The set of responses in which we agree, and the way they interweave with our activities, is our *form of life*. Beings who agreed in consistently giving bizarre quus-like responses would share in another form of life" (1982, p. 96).

Therefore, we can regard the sceptical solution as a “non-factualist” view about meaning, since it accepts the Basic Sceptical Conclusion that there is no fact about a speaker constituting the fact that he means *plus* by “plus”. The sceptical solution can be called a “communitarian” view as well, as it regards the existence of a community of speakers sharing their linguistic practices as essential to the existence of the practice of meaning something by words: a speaker can be said to mean something specific by a word if others can judge that the speaker uses the word in the way they would on similar occasions. Hence, we can take the sceptical solution to be a non-factualist communitarian view about meaning.

1.6. Conclusion

In this chapter, we explored Kripke’s Wittgenstein’s sceptical argument and sceptical solution. We started with KW’s sceptic’s sceptical argument, which aimed at disclosing the implausibility of Traditional Realism about meaning. The sceptic took for granted the normativity of meaning thesis, which was characterized by Traditional Realism in two ways, with respect to the rule-following picture of meaning and with respect to the existence of meaning facts. The sceptic, through criticizing and rejecting various suggestions, argued that there can be found no fact of the matter as to which rule a speaker follows, or what he means by his words. This established the sceptic’s Basic Sceptical Conclusion. This conclusion together with Traditional Realism and the arguments outlined in Sections 1.2.1 – 1.2.10 led to the Radical Sceptical Conclusion, that there is no such thing as meaning anything by any word.

KW started his sceptical solution by accepting the sceptic’s Basic Sceptical Conclusion: he agreed with the sceptic that there is no fact of the matter as to what a speaker means by his words. However, he rejected the Radical Sceptical Conclusion, since it is obviously unacceptable and self-defeating. In this regard, Traditional Realism, which brought about the radical scepticism about meaning, was ruled out by KW as a misconstrual of our ordinary practice of meaning something by words. As his sceptical solution, KW suggests that we need to concentrate on the conditions under which meaning-ascribing sentences can be justifiably asserted and the utility which a practice of asserting those sentences under those conditions has in our lives. According to KW, we can justifiably assert that a speaker means something specific by his words if we can observe that the speaker uses them in agreement with the way we do. Consequently, for

a solitary speaker considered in isolation, there would be no appropriate assertability conditions: what seems right to such a speaker is right. The agreement in a speech community, however, cannot be explained in terms of its members' shared grasp of meaning, since, as the sceptical argument showed, there is no fact of the matter as to what they mean by their words. For KW, it is rather a brute empirical fact that the community is uniform in its practices. Nonetheless, there is still a reason why the members of that community attribute meanings to each other in the way they do: asserting meaning ascribing-sentences under the relevant assertability conditions brings about endless utilities and benefits in their lives.

In the next chapter, we will start our discussion of Davidson's view about meaning and linguistic understanding, which will provide the material we need to make a comparison between the views of these two philosophers in Chapter Four.

2. Chapter 2: The Later Davidson

Introduction

Davidson's works on meaning and linguistic understanding are usually divided into two parts: his earlier writings and his later ones. Davidson's earlier writings concentrate on how a formal theory of truth, in the line of Tarski's, can be constructed and formally and empirically constrained to generate a correct interpretation of a speaker's utterances. His later writings, however, display a sort of shift in focus: now, he is more concerned with the necessary and sufficient conditions for successful communication in general, especially when two speakers seem to be speaking a familiar language.[22] In this chapter, we will focus on Davidson's later views of meaning and linguistic understanding, since it is in these later writings that Davidson starts criticizing Kripke's Wittgenstein. Nonetheless, as this chapter presupposes a knowledge of Davidson's early account of meaning, we have provided an exposition of his earlier views in the Appendix to this chapter.

Davidson's later works aim to show how interpretation can be a ubiquitous linguistic phenomenon, which any adequate view of meaning must take into account. For Davidson, what can best explain success in the practice of meaning something by an utterance is success in mutual interpretation: a speaker can be said to be successful in meaning something by her words if she intends her utterance of the words to be interpreted in a particular way and the interpreter interprets her utterance in that way. Meanings, hence, are treated as the products of successful communication.[23] In this

[22] Indeed, whether or not there really is a considerable shift from his earlier works to his later ones is a matter of controversy, since the essentiality of interpretation and the significant role of theories of truth still seem to be central to his later account of meaning. Radical interpretation (i.e. interpreting a speaker speaking an unknown language) and non-radical interpretation (i.e. interpreting a speaker speaking a familiar language) can be taken to be two different aspects of Davidson's general idea that "all understanding of the speech of another involves radical interpretation" (1973a, p. 125), together with another claim from Davidson, according to which "the problem of interpretation is domestic as well as foreign: it surfaces for speakers of the same language in the form of the question, how can it be determined that the language is the same?" (1973a, p. 125). Davidson's later works are concerned with interpretation within a speech community. Moreover, formal theories of truth are still responsible for modeling or describing speakers' and interpreters' linguistic abilities to speak and understand, though this time they supposedly speak a similar language. Hence, it is better to take this alleged shift as a shift in focus, not in the original ideas he has always advocated. For more discussion of this, see (Glüer, 2013, 2011), (Ludwig & Lepore, 2005), (Heal, 1997), and (Hacking, 1986).

[23] As Glüer (2001) puts it, Davidson suggests that we need to focus on the concept of "*communication by language*", rather than that of "*a language*" (2001, p. 54). As a result, the meanings of words and

picture, speakers' propositional attitudes, more particularly, their intentions, play a very important role: what the speaker means by her words is a matter of the way she intends her words to be understood by her interpreter. If that is the case, then the main thing we need to explain is what it takes for the interpreter to interpret the speaker's utterance, that is, to grasp the meaning the speaker intends her words to have. If such "mutual interpretation" (Davidson, 1997a, p. 83) takes place, then following certain fixed rules and conventions would not be essential to the practice of meaning something by words: the speaker is understood, no matter what rule she follows. In this way, Davidson criticizes and rejects any view which treats conforming to certain conventions as an essential feature of successful communication.

We will split this chapter into two parts: the first part deals with Davidson's negative argument against a certain conception of meaning (i.e. Conventionalism) and, in the second part, we will introduce his positive, or alternative, view of meaning and linguistic understanding (that is, his Interpersonal View).

2.1. The Negative Part

Davidson has argued against a certain conception of language and linguistic understanding, or more specifically, a certain account of the practice of meaning something by an utterance. He has introduced this view differently in different papers (see e.g. 1984b, p. 265; 1986, p. 93; 1991b, p. 143; and 1994, p. 110). We will call this view the "Common View", since Davidson takes it to be the view that "many philosophers and linguists have supposed" (1986, p. 107). Briefly speaking, the Common View takes the existence of conventions to be essential to the possibility of successful communication. These conventions are supposed to determine the meaning of the words a speaker utters in one way or another. By criticizing different versions of such a view, Davidson intends to reach the conclusion that conventions can play no essential role in any adequate account of linguistic understanding. If that is true, then any view which takes the existence of such conventions to be necessary will be wrong. Hence, it is important to clarify Davidson's criticisms of the Common View, since, as

sentences are to be treated as theoretical notions since, for Davidson, it is utterances, as the products of successful communication, that are meaningful. On Davidson's claim that the concept of meaning, and other relevant concepts, such as that of a language, reference, and the like, are theoretical concepts, see (Davidson, 1973a, p. 133; 1977, p. 223; 1992, p. 108); see also (Ludwig & Lepore, 2005, especially chapters 15 and 18).

we will see in Chapter Four, Davidson's rejection of this view has a role to play in Davidson's criticisms of, and his alternative solution to, KW's sceptical problem. First of all, what is the Common View?

2.1.1. The Common View

Davidson's attack on this view and, hence, his first characterization of it, appears in his paper "Communication and Convention" (hereafter "CC") (1984b), followed by "A Nice Derangement of Epitaphs" (hereafter "ND") (1986) and "The Social Aspect of Language" (hereafter "SAL") (1994). In CC, the Common View is introduced as the view that treats conventions as "necessary to the existence of communication by language" (1984b, p. 265). In this regard, we can call this view "Conventionalism" as well, as Glüer does (see e.g. Glüer, 2013, 2001), since it takes the existence of conventions determining the meaning of words to be essential to the existence of successful communication. As we will see in the following sections, Davidson casts doubt on the essentiality of the notion of convention and the adequacy of any view which assigns a similar role to rules, norms, standards, customs, institutions, and the like.

The Common View is again introduced in ND, when he starts criticizing convention-based views of meaning and linguistic understanding. In this paper, the Common View is characterized as one according to which the meaning of our words must be treated as (1) systematically related, (2) shared by speakers of a linguistic community, and (3) prepared in advance of a particular conversation in terms of some specific conventions (see Davidson, 1986, p. 93). Davidson, after some modifications, accepts the first two principles, but he sees no reason to accept any version of the third, that is, the claim that the meaning of our words must be viewed as conventional and determinate in advance. Putting it in terms of theories of meaning modeling the linguistic abilities of a speaker to speak and an interpreter to interpret, the Common View implies that "each interpreter comes to a successful linguistic exchange prepared with a "theory" which constitutes his basic linguistic competence, and which he shares with those with whom he communicates … [and] that the knowledge of abilities that constitute the theory may be called conventions" (Davidson, 1986, p. 102). As we will see in the next section, Davidson believes that, although there must be theories characterizing the speaker's and the interpreter's linguistic abilities, it is neither necessary nor sufficient for the

communicators to share such a theory in order to have successful communication, that is, a theory in which the meanings of their words are specified in advance of the particular conversation they have.

In SAL, Davidson pursues the main criticisms he made in ND and tries to resolve some confusion that his remarks in ND have brought about. This time, he mainly concentrates on the Common View's conception of what a language is and the problems with such a conception. According to the Common View, as presented in SAL, "in learning a language, a person acquires the ability to operate in accord with a precise and specifiable set of syntactic and semantic rules; verbal communication depends on speaker and bearer sharing such ability, and it requires no more than this" (1994, p. 110). Hence, the Common View's main claim is still the same as before: the role of rules and conventions is essential to the practice of successful communication. As another example of Davidson's characterization of this view, we can mention his paper "James Joyce and Humpty Dumpty" (1991b), in which he again points to a similar tension between his favored view and the Common View: "A tension between the thought that what a speaker intends by what he says determines what he means and the thought that what a speaker means depends on the history of the uses to which the language has been put in the past" (1991b, p. 143).

Therefore, in all the above mentioned papers, Davidson is attacking the view implying that the meaning of words must be taken to be determined in advance in terms of some shared conventions or rules. However, what is Davidson's problem with this view?

2.1.2. Davidson's Argument against the Common View

To investigate the main criticisms Davidson makes of the Common View, we need to start with his remarks in CC, in which he spelled out three versions of this view and criticized them by citing different sorts of critiques.

In CC, Davidson begins with drawing our attention to the different practices in which conventions are usually taken to have a role to play. He specifically contrasts the practice of playing and winning a game with the practice of speaking a language. Games are the clearest instances of activities in which rules and conventions play an essential role, while, for instance, in explaining how to eat, talking of rules is dispensable: Following one instruction for eating (e.g. using forks), rather than another

(e.g. using hands), is not essential to the purpose of satisfying the hunger. In the case of playing a game, such as chess, it is essential to move the chess pieces in specific ways, that is, in accordance with certain rules. Now, the question is whether the case of linguistic communication is similar to the case of playing a game or to the case of explaining how to eat. Davidson thinks that the role of conventions in explaining linguistic practices is analogous to the role they play in explaining how to eat: they are not essential to the success of these practices. The general reason Davidson introduces for this claim is that playing and winning a game has a combination of features that speaking a language lacks: (1) those who play a game usually want, or at least, "*represent* themselves as wanting to win"; (2) "one can win only by making moves defined by the rules of the game, and winning is wholly defined by the rules"; and (3) "winning can be, and often is, an end in itself" (Davidson, 1984b, p. 267). According to Davidson, "no linguistic behaviour has this combination of features" (1984b, p. 267). Davidson, to support this claim, introduces and then criticizes three versions of the Common View, which take our linguistic practices to be essentially rule-governed.

Briefly speaking, the three versions of the Common View or Conventionalism are introduced by Davidson as follows: (I) the first sort of theory claims that "there is a convention connecting sentences in one or another grammatical mood … with illocutionary intentions, or some broader purpose"; (II) according to the second version of the Common View, there is "a conventional use for each sentence"; finally, (III) the third kind of theory implies that "there is a convention that ties individual words to an extension or intension" (1984b, p. 266). In the following sections, we characterize each version of the Common View and investigate the main criticisms Davidson makes of them.

(a) Criticizing the First Version of the Common View

According to Davidson, the first sort of view is advocated by Dummett (1973, 1959/1978). As Davidson states, Dummett claims that (a) "there is a conventional connection between uttering a declarative sentence and using it to make an assertion" and (b) there is a conventional "connection between making an assertion and the intention to say what is true" (1984b, p. 266). In other words, Dummett believes that it is because of conforming to a specific linguistic convention that we thereby count using

a declarative sentence as making an assertion, and it is a conventional matter that we suppose an assertion is made with the intention of saying what the speaker believes is true. For instance, when I use the declarative sentence "Your room is messy", it is a matter of following my speech community's relevant conventions that others take me as asserting that your room is messy and that I am saying what I believe to be true by making this assertion. This view, thus, models the practice of speaking a language as analogous to the practice of playing a game, an analogy that Davidson strongly rejects. His specific argument against this version of the Common View is based on the claim that "unlike winning, what constitutes the making of an assertion is not governed by agreed rules or conventions" (Davidson, 1984b, p. 268). According to Davidson, making an assertion is not a rule-governed activity, since (a) there are no "conventions governing assertion" and (b) there is "[no] convention linking assertion to what is believed true" (1984b, p. 268).

Regarding the Common View's claim that there are conventions governing the practice of making an assertion, Davidson argues that, first of all, we can see that there is actually no such convention or conventional sign in our language, the use of which makes an utterance a case of an assertion: "There is no word or sign in language whose function is simply to assert something" (1984b, p. 269). Indeed, "there is no known, agreed upon, publicly recognizable convention for making assertions" (Davidson, 1984b, p. 270). Davidson thinks that the reason why such a convention is absent in language is that it could be used in real situations as well as unreal or fake ones: it would thereby be of no use. For instance, such a convention could be simply used by an actor to make his act more real and touching, and there is no way to distinguish between these unreal situations and the real ones, as Davidson's example clarifies:

> Imagine this: the actor is acting a scene in which there is supposed to be a fire. … It is his role to imitate as persuasively as he can a man who is trying to warn others of a fire. 'Fire!' he screams. And perhaps he adds, at the behest of the author, 'I mean it! Look at the smoke!' etc. And now a real fire breaks out, and the actor tries vainly to warn the real audience. 'Fire!' he screams, 'I mean it! Look at the smoke!' etc. (1984b, pp. 269-270)

Regarding the conventionalists' claim that there is a convention linking assertion to what is believed to be true, Davidson points out that, although it is true that in making an assertion one is often representing himself as believing what he says, there can be no convention or "conventional sign that shows that one is saying what one believes", simply because if there was such a convention, it would be entirely fruitless: "Every liar

would use it" (1984b, p. 270). And there can be no convention helping us to find out who is lying and who is not: "There is no convention of sincerity" (1984b, p. 270). There cannot be such a convention for sincerity because conventions cannot "connect what may always be secret – the intention to say what is true – with what must be public – making an assertion" (1984b, p. 270). How can there be a convention following which discloses what we indeed think in our mind? Making an assertion is a public practice, while intending to say what one believes to be true is entirely subjective.

Therefore, Davidson rules out the first version of the Common View by arguing that, in our linguistic practices, such as making an assertion, we cannot find the combination of features possessed by the practice of playing and winning a game.[24]

(b) Criticizing the Second Version of the Common View

The second sort of theory states that we can derive sentences' literal meanings "from the non-linguistic purposes their utterances serve" (Davidson, 1984b, p. 271). According to this view, when we use different sentences of our language, we use them with specific purposes: we always have a variety of further intentions and purposes in mind when we utter a sentence. If that is the case, then the second sort of theory claims that there is a conventional link between the specific purposes we have in using a sentence and the meaning that sentence has for us. The hope is to extract the sentence's literal meaning from these purposes. For example, as Davidson says, "getting someone to eat his eggplant is, you might say, what the English sentence 'Eat your eggplant' was made to do" (1984b, p. 271). If we can say that "eat your eggplant" is used by speakers with the purpose of getting someone to eat his/her eggplant, then the relevant convention states that the utterance's literal meaning is *eat your eggplant*!

[24] In Davidson's view, not only is there no general assumption "that someone who utters a declarative sentence wants or intends to speak the truth", but also "speaking the truth, in the sense of uttering a true sentence, is never an end in itself" (1984b, p. 268). For instance, I may use the declarative sentence "Your room is messy", not to make an assertion, but to get my child to tidy his room, that is, to make an order or warning. Also, I may use the sentence "The weather is nice" with the intention of asserting that the weather is nice and with the intention of saying what I believe is true, but for the additional purpose of starting a conversation with someone and, then, with the further purpose of getting more familiar with him, and with the further purpose of finding a friend in the neighborhood, etc. Hence, the practice of making an assertion is not like the practice of playing and winning a game. While the latter is a rule-governed activity, the former is not.

Davidson concedes that "each use of language has an ulterior purpose" (1984b, p. 272). The ulterior purposes are indeed non-linguistic purposes: for instance, I may use the sentence "The weather is cold" to mean that *the weather is cold* with the non-linguistic or ulterior purpose of getting my friend to shut the window. If that is true, "then one must always intend to produce some non-linguistic effect through having one's words interpreted" (Davidson, 1984b, p. 272). Indeed, if my friend does not understand that by my utterance I meant that *the weather is cold*, he would not be able to satisfy my desire of getting him to close the window. Thus, my utterance must be understood by my friend in a specific way, i.e. as meaning that *the weather is cold*, if I want to produce that non-linguistic effect on my friend.

According to Davidson, however, the attempt to link non-linguistic purposes with the meaning of sentences is desperate. The first problem is that there can be endless ulterior purposes in using a sentence, and if a convention is to be relevant here, it must be able to "pick out … those cases in which the ulterior purpose directly yields the literal meaning" (Davidson, 1984b, p. 273). For Davidson, however, conventions cannot do this. If it is the specific ulterior purpose in using a sentence that is supposed to deliver the meaning of my utterance of that sentence, then the convention must clarify the situations in which I really use that sentence with that purpose, not with any other purposes. For instance, regarding Davidson's example of "Eat the eggplant", the relevant conception must clarify that using this sentence with such-and-such meaning is related to my having the purpose of getting my friend to eat his eggplant. However, the question is whether getting my friend to eat his eggplant is the only purpose I have in mind when I utter the sentence. According to Davidson, with the sentence meaning the same thing, one may have many different ulterior purposes in mind. By uttering the sentence "Eat the eggplant" with the literal meaning of *eat the eggplant*, I may just *represent* myself as wanting you to eat your eggplant, while what I really want is to warn you to finish your food sooner if you want to get to your meeting on time, or to distract you so that my friend can steal your wallet, or to prove that this much eggplant was too much to order, and the like. How is a convention supposed to pick out the relevant situation in which I just want you to eat your eggplant? This means that, even if there were such conventions, they would require the speaker to be sincere, that he really wants what he *represents* himself as wanting. But, as Davidson emphasized before, "nothing is more obvious than that there cannot be a convention that signals sincerity" (1984b, p. 274). The reason, again, is that such a convention would be entirely useless,

since every liar would use it. Therefore, a sentence can be used with the same meaning in standard as well as non-standard cases: "The literal meaning operates as well when the [standard] use is absent as when it is present" (1984b, p. 275). Ulterior purposes have nothing to do with the literal meanings of sentences; rather, according to Davidson, the ulterior purpose of an utterance and its literal meaning are independent so that the literal meaning cannot be derived from the purpose. Davidson calls this phenomenon "the autonomy of meaning" (1984b, p. 274).

(c) Criticizing the Third Version of the Common View

According to the third kind of theory, "the meaning of a word is conventional, that is, that it is a convention that we assign the meaning we do to individual words and sentences when they are uttered or written" (Davidson, 1984b, p. 276). We can properly call this version of the Common View "Conventionalism", which takes the meaning of words to be determined by specific conventions.[25] Davidson's main question is: What exactly is the necessary convention? Whether "convention in any reasonable sense helps us understand linguistic communication"? (1984b, p. 276).

The Common View might offer an answer to Davidson's question: for successful communication, two people, as a matter of following some conventions, should always "mean the same thing by uttering the same sentences" (Davidson, 1984b, p. 276). In other words, even if there are only two people communicating with each other, they still need to construct some conventions regarding the words they use so that they can be said to mean the same thing by the same word. Davidson's target is now to challenge all

[25] Davidson cites David Lewis's (Lewis, 1975) definition of convention to illustrate this view. For Lewis, as Davidson states, "convention is a *regularity* R in action, or action and belief, a regularity in which more than one person must be involved" (Davidson, 1984b, p. 276). R is counted as a convention if the following conditions hold: "(1) Everyone involved conforms to R and (2) believes that others also conform. (3) The belief that others conform to R gives all involved a good reason to conform to R. (4) All concerned prefer that there should be conformity to R. (5) R is not the only possible regularity meeting the last two conditions. (6) Finally, everyone involved knows (1)-(5) and knows that everyone else knows (1)-(5), etc. (1984b, p. 276). Davidson does not intend to challenge these conditions in detail; rather, for the sake of his argument, he just presupposes Lewis's definition of a convention. According to this definition, a convention can be held between two people communicating with each other: "Lewis's analysis clearly requires that there be at least two people involved, since convention depends on a mutually understood practice. But nothing in the analysis requires more than two people. Two people could have conventions, and could share a language" (1984b, p. 276). To this extent, Davidson has no serious problem with this definition. Davidson's problem is with the role such a convention may be taken to be playing in our account of the practice of meaning something by a word.

such responses. Is meaning the same thing by the same word a necessary condition on successful communication? Davidson does not believe that to have a proper account of success in communication we necessarily need the requirement that two speakers mean the same thing by the same word, or even speak the same language. He even goes further: neither do we really need the requirement that a speaker herself means the same thing by the same word over time. He discusses his general argument for this claim in CC, but the detailed discussion of this argument has been put forward in ND and further supported in SAL. We can take this version of the Common View as the most important one, with which Davidson deals in his later works. Indeed, it is this version of the Common View that Davidson attributes to KW too. Thus, clarifying his criticism of this view will be vital for our future investigation of Davidson's grasp of KW's sceptical argument and solution. To see how Davidson's negative argument against this version of the Common View goes, we need to introduce his remarks on the phenomenon of malapropism, the notion of first meaning, and his discussion of prior and passing theories of meaning.

2.1.3. The Argument from Malapropism

If the Common View is true, then there must be at least some necessary convention, conforming to which can be treated as essential to our success in communication. Davidson aims to reject this claim. For Davidson, following some fixed rules or conventions regarding the meaning of words – shared in advance of a particular conversation between two people – is neither necessary nor sufficient for them to have successful communication. For a better exposition of Davidson's main argument against this sort of view, it is helpful to consider his more recent characterization of the Common View. According to this characterization, the Common View implies that the literal or "first" meaning of words and sentences must be systematically related, shared by at least two people, and fixed previously in terms of some convention (see Davidson, 1986, p. 93). Davidson's problem is particularly with the third condition specified above, i.e. the claim that first meanings have to be essentially fixed in advance of actual episodes of communication. Before looking at Davidson's criticisms of this claim, we need to explicate Davidson's notion of "first meaning".

(a) First Meaning vs. Conventional Meaning

The notion of meaning which Davidson, especially in his recent works, employs is the notion of literal meaning. However, he thinks that this "term is too incrusted with philosophical and other extras to do much work" (1986, p. 91). Instead, he prefers to speak of the "first meaning" that words have when they are used by a speaker. Briefly speaking, the first meaning of a word is the meaning that the word has when it is "uttered by a particular speaker on a particular occasion" (1986, p. 91). Davidson's main point in introducing this notion is that the literal meaning of words uttered by a speaker may not always be the same as their conventional meaning. We can take the conventional meaning of words to be what can be found in good dictionaries, as Davidson states: "If the occasion, the speaker, and the audience are 'normal' or 'standard' …, then the [conventional] meaning of an utterance will be what should be found by consulting a dictionary based on actual usage (such as *Webster's Third*)" (1986, p. 91). For instance, suppose that the conventional meaning of "grave" is *grave*, that is, "A hole dug in the ground to receive a coffin or dead body, typically marked by a stone or mound" (Oxford Dictionaries Online). In ordinary or standard cases, I use this word to talk about graves, e.g. the normal graves in a cemetery. However, it is possible that I use this word in a non-standard way: I may utter "That's the best onion grave I have ever tried" at a dinner party in my friend's home. This is a non-standard use of "grave". What I intended to mean by my utterance of "grave" is not what standard dictionaries suggest: there is no grave or gravestone in the dinner party! Rather, I intended to mean *gravy* by this word, and "gravy", according to standard dictionaries, means "A sauce made by mixing the fat and juices exuded by meat during cooking with stock and other ingredients" (Oxford Dictionaries Online). However, why does Davidson call this meaning the "first" meaning?

According to Davidson, when we use language, we do so with specific intentions and purposes. I might utter "That's the best onion grave I have ever tried" to impress my friend and appreciate her kind hospitality, or I might make that utterance with the purpose of teasing my friend's cooking ability. The point, however, is that, for my hearer to make sense of any further intentions and purposes I have when I utter that sentence, she should be able to understand what I intended to mean by those words on that particular occasion. Otherwise, my words just mean nothing to her. In this regard, "first meaning is the first meaning referred to" (Davidson, 1986, p. 92). In the order of interpretations of my utterance, first meaning comes first. Suppose that the intention

behind my utterance is to get my friend to invite me again whenever she decides to cook the same meal. In this situation, she should assign a correct meaning to my words, that is, she should understand what I intended to mean by the words on that particular occasion, *if* she is to grasp what purpose I had in mind by uttering those words. If she fails – for instance, if she sticks to the dictionary-based or conventional meaning of the words – then what my utterance meant on that occasion would be entirely confusing. The meaning that she should have attributed to my utterance is something like *That's the best onion gravy I have ever tried*. While my communication with her is successful only if she interprets my utterance of "grave" as meaning *gravy*, *gravy* is not the conventional meaning of "grave"; it is not the way my community or even my friend uses the word. Rather, it is the *first* meaning that my hearer must attribute to my utterance in order to successfully communicate with me, i.e. to understand what other meanings, intentions, or purposes I have in uttering those words. Therefore, successful communication depends on correct attribution of first meanings. In most standard situations, first meanings are just conventional meanings. However, they may diverge dramatically.

With this clarification of first meaning in hand, we can go on to characterize Davidson's argument against the Common View.

(b) The Common View Re-Characterized

Davidson's question was: What is the necessary convention that two people must conform to in order to have successful communication with each other? The Common View, according to Davidson, claimed that in order to have successful communication, "first meaning must be systematic, shared, and prepared" (Davidson, 1986, p. 93). What is Davidson's criticism of these conditions?

First of all, he accepts that first meanings can be treated as systematically related and shared. According to Davidson, we can embrace the requirement of systematicity by emphasizing the fact that interpretation is possible because "the interpreter can learn the semantic role of each of a finite number of words or phrases and can learn the semantic consequences of a finite number of modes of composition" (1986, p. 95). As explained in the Appendix, Davidson's learnability argument implies that in order to account for our ability to generate and understand novel sentences, we need to take a language as

consisting of a finite number of primitive words and a finite number of rules regarding the way these parts can be employed to form well-formed sentences. In this regard, "the interpreter thus has a system for interpreting what he hears or says" (1986, p. 95). The model for such a system of interpretation is a theory of truth, which Davidson developed in his earlier works. This theory produces the truth-conditions of all the possible utterances a speaker can make by breaking them down into their parts and specifying the semantic features of the parts, that is, the references of names and the satisfaction conditions for predicates. Having seen matters in this way, Davidson thinks that the first principle is still acceptable, since the meanings that such a theory produces are systematically related: "All that matters in the present discussion is that the theory has a finite base and is recursive" (1986, p. 95).[26] As we will see, this theory, however, must deal with the first meanings of utterances, not necessarily their conventional meanings.

Regarding the second principle, that is, the claim that first meanings must be shared, Davidson believes that we can accept this principle by regarding it as the claim that "what must be shared is the interpreter's and the speaker's understanding of the speaker's words" (1986, p. 96). In other words, for understanding what the speaker utters, the interpreter must understand what the speaker means by her words, though the interpreter himself may use them in a different way. Hence, instead of saying that the speaker and the hearer must share the same words meaning the same thing, we can say that the hearer must understand the first meaning of the speaker's utterance, and, in this sense, they must share what the *speaker* means by her words. Regarding the theories modelling the speaker's and the interpreter's linguistic abilities to speak to and interpret each other, Davidson's claim turns into the claim that the speaker and the hearer must share the theory which produces what the *speaker* means by her words on a particular occasion, i.e. the first meaning of her words.[27]

Nonetheless, Davidson does not believe that first meanings can be treated as essentially conventional and fixed in advance of the particular conversation that the speaker has with her interpreter. Before looking at the reasons Davidson cites to support

[26] For an exposition of Davidson's earlier ideas on the compositionality of natural languages and his theory of truth, see the Appendix to this chapter: The early Davidson.

[27] We will discuss these remarks further in our discussion of Davidson's distinction between prior and passing theories.

this claim, we can characterize Davidson's general argument against this version of the Common View as follows:

(1) The Common View about meaning and linguistic ability requires the first meaning of words to be systematic, shared, and prepared in advance in terms of some conventions.

(2) First meanings, and the theories producing such meanings, can be systematic and shared, but cannot be prepared in advance in terms of conventions.

(3) Therefore, the Common View is wrong.

The main line in this argument is the premise (2), in which Davidson claims that first meanings cannot be regarded as essentially conventional. He believes that sharing and following some fixed rules or conventions is *neither necessary nor sufficient* for the purpose of giving an adequate explanation of success in communication. His argument for this claim appeals to the phenomenon of "malapropism".

(c) The Phenomenon of Malapropism

According to Davidson, there are varieties of commonplace and ubiquitous linguistic phenomena, such as malapropisms, slips of the tongue, and the like, which we normally face in ordinary cases of communication and which we can smoothly understand – or as Davidson sometimes says, we can get away with them all the time (see e.g. Davidson, 1986, p. 98). Davidson thinks that we need to take into account such phenomena if we wish to provide an adequate and comprehensive account of meaning. Malapropisms, in particular, belong to the widespread linguistic phenomenon of giving the meaning of one word to another word, or giving a word a non-usual meaning, or even a new one (see e.g. Davidson, 1986, pp. 90, 94-95; 1994, p.115).[28] For example, consider Davidson's Mrs. Malaprop who has specific verbal habits, such as using some words in a mistaken way. She may simply use "grave" to mean *gravy*, "epitaph" to mean *epithet*, or "derangement" to mean *arrangement*. Suppose that she utters "That's a nice derangement of epitaphs" when she is in her office looking at the title of a book. The

[28] Indeed, we are already familiar with some examples of this phenomenon from our discussion of the Davidsonian distinction between first meaning and conventional meaning (see Section 2.1.3.a).

question is: What is the meaning of this utterance for her hearer? According to the conventional meaning of the uttered words, the utterance means that *that's a nice derangement of epitaphs*. The problem, however, is that this utterance with this conventional meaning does not make any sense, especially when we regard the situation in which it is made, since there are no deranged epitaphs, rather just nicely arranged epithets. Nonetheless, Mrs. Malaprop's utterance is easily understood by her hearer – as meaning something specific which is, of course, non-standard. Her utterance does not mean that *that's a nice derangement of epitaphs*, rather that *that's a nice arrangement of epithets*. It is enough that the hearer utilizes the evidence and clues there are to reach the correct interpretation of Mrs. Malaprop's utterance, for instance, the occasion of speech, the attitudes of Mrs. Malaprop, her character, her habits, her way of living, and so on. In this way, the interpreter gets to the meaning Mrs. Malaprop intended her utterance to have on that occasion, which may not be the conventional meaning of the words. "Derangement" meaning *arrangement* and "epitaphs" meaning *epithets* are not the conventional meanings of these words: They are not the meaning we can find in ordinary dictionaries. Rather, they are the meanings that Mrs. Malaprop intends the hearer to attribute to her words. Therefore, in the case of malapropism, "the hearer realizes that the "standard" interpretation cannot be the intended interpretation" (Davidson, 1986, p. 90). The hearer realizes this tension because, if he takes the malapropian expression to have its standard meaning, the meaning that he would come up with would be absurd considering the situation in which it is uttered. What is crucial for the communication with Mrs. Malaprop to be successful is the interpreter's arriving at this meaning, i.e. the intended meaning of Mrs. Malaprop's utterance. This meaning is indeed the first meaning of her utterance, i.e. the meaning that the hearer must grasp if he is to be able to understand the intentions and purposes Mrs. Malaprop had in mind when she uttered those words.[29] Therefore, successful communication depends on the speaker's utterance being interpreted by the interpreter in the way the speaker intended. However, what are the consequences of such a philosophical investigation of the phenomenon of malapropism for the Common View?

According to Davidson, malapropisms raise a serious challenge for the Common View, since "malapropisms introduce expressions not covered by prior learning" (1986, p. 94). Such a prior learning includes the conventional meanings of words. Indeed, not

[29] For more discussion of malapropism see (Pietroski, 1994), (Whiting, 2007), and (Reimer, 2004)

only does the phenomenon of malapropism raise such a challenge, but also "our ability to perceive a well-formed sentence when the actual utterance was incomplete or grammatically garbled, our ability to interpret words we have never heard before, to correct slips of tongue, or to cope with new idiolects … threaten standard descriptions of linguistic competence" (1986, p. 95). As Davidson emphasizes, for instance, it is more than obvious that "almost no two people share all words" (1994, p. 115), that different speakers "have different stocks of proper names, different vocabularies, and attach somewhat different meanings to words" (1984b, p. 277), that new words and names can always be introduced during a particular conversation, and so on. These phenomena point to the fact that communication can successfully take place even if the communicators use their words in different, non-standard or non-conventional ways. In this regard, the Common View's claim that first meanings are essentially conventional is in conflict with the fact that we can understand malapropisms. Following some fixed conventions or rules cannot be treated as essential to the success in the practice of meaning something by words. We can, hence, characterize the Argument from Malapropism as follows:

(I) If the Common View was correct, i.e. if first meanings were essentially conventional, then malapropisms would not be smoothly understandable.

(II) Malapropisms are smoothly understandable.

(III) Therefore, the Common View is wrong: Although first meanings can be systematic and shared, they cannot be conventional and fixed in advance.

Successful communication does not require the communicators to speak the same language, or mean the same thing by the same word.[30] Two speakers can have their own unique idiolects and still successfully communicate with each other. Indeed, what seems to be necessary for successful communication to occur is that "speaker and hearer must assign the same meaning to the speaker's words" (Davidson, 1984b, p. 277). Mrs. Malaprop and her hearer only need to agree on the meaning Mrs. Malaprop attaches to her utterance on that particular occasion: they do not need to mean the same thing by the same words.

[30]Similarly, Davidson believes that "if the concept of following a rule is not quite appropriate to describe meaning something by saying something, it is also questionable whether … we should accept without question the idea that meaning something demands (as opposed to sometimes involving) a convention, custom, or institution" (1992, p. 114).

However, as a final opportunity for the Common View to find a role for conventions in our explanation of successful communication, it may be claimed that there must be some sort of agreement *over time* between the speaker and the interpreter on what the speaker means by a word. The advocates of this view may insist on the idea that, even if we accept Davidson's claim that for successful communication to occur, it is enough that the speaker and the interpreter agree on what the speaker means by her words, we still need the requirement that this agreement is repeated over time: Mrs. Malaprop must use "epitaph" to mean *epithet* next time, or at least during her conversation with the interpreter. As a result, the existence of some recurrence over time is essential to successful communication – for instance "the interpretation of sound patterns" in conformity with some rules or conventions, or the requirement that "speaker and hearer must repeatedly, and with mutual agreement, interpret relevantly similar sound patterns of the speaker in the same way (or ways related by rules that can be made explicit in advance)" (Davidson, 1984b, p. 277). This means that Mrs. Malaprop's meaning *epithet* by "epitaph" must be treated as a convention that she follows during her conversation with her hearer. Does, however, Mrs. Malaprop need to mean the same thing by the same word over time in order to have successful communication with her hearer?

Davidson disagrees with this claim too. For Davidson, a speaker is not required to mean the same thing by the same word over time. The reason again is that if the evidence and clues are sufficient for the interpreter to reach the speaker's intended meaning, it does not really matter what the speaker meant by those words before, or what she is going to mean by them in the future: "You can change the meaning provided you believe (and perhaps are justified in believing) that the interpreter has adequate clues for the new interpretation" (Davidson, 1986, p. 98). Mrs. Malaprop may intend to mean something different by "epitaph" next time. All that matters is that the interpreter can successfully understand what she means by her word. These remarks are sometimes put in a different terminology by Davidson, that is, in terms of the theories that can be used to model the speaker's and the interpreter's linguistic abilities to speak and understand: "prior" theories and "passing" theories.

2.1.4. Prior and Passing Theories

As Davidson argued, what the speaker and the interpreter must agree on is what the

speaker intends her words to mean on a particular occasion. Successful communication does not essentially depend on the speaker and the interpreter using the same word to mean the same thing. Davidson puts these ideas in terms of the theories of meaning, which the speaker or the interpreter can be taken to be using when they communicate with each other. These theories are indeed employed by theorists or philosophers to model the speaker's and the interpreter's linguistic abilities to speak and interpret. As explained in the Appendix, the Davidsonian Tarski-style theory of truth produces the truth-conditions of sentences in its theorems, by specifying the references of words and the satisfaction conditions for predicates in its axioms. In this regard, Davidson thinks that,

> it is possible, of course, that most interpreters could be brought to acknowledge that they know some of the axioms of a theory of truth; for example, that a conjunction is true if and only if each of the conjuncts is true. And perhaps they also know theorems of the form 'An utterance of the sentence "There is life on Mars" is true if and only if there is life on Mars at the time of the utterance'. (1986, pp. 95-96)

It is, however, important to note that Davidson's claim is not that the speaker or the hearer explicitly knows such theories. Rather, they are the theories that *we* use to systematically theorize about their linguistic abilities. As Davidson states, "to say that an explicit theory for interpreting a speaker is a model of the interpreter's linguistic competence is not to suggest that the interpreter knows any such theory" (1986, p. 95).[31]

Davidson believes that if the speaker's ability to speak in an understandable way and the interpreter's ability to interpret the speaker's utterances in that way is not necessarily limited to what they have learnt before, e.g. the fixed conventional meanings of words, then the theories modeling such abilities should not be treated as fixed in advance of the particular conversation they have and constrained to generate just the conventional meanings of their words. That is to say, the theory producing what the speaker means by her words on a particular occasion cannot just contain what the words conventionally mean, or even what they meant before as used by the speaker in the past. The reason, again, is that the speaker may intend to mean something different by her words next time, and if her utterance is understood by the hearer, we can say that the speaker has

[31] As Davidson emphasizes, "the point is not that speaker or hearer has a theory, but that they speak and understand in accord with a theory – a theory that is needed only when we want to describe their abilities and performance" (1994, p. 113). More generally, "no one now has explicit knowledge of a fully satisfactory theory for interpreting the speakers of any natural language" (1986, p. 96).

been successful in meaning something specific by them. To deal with this situation, Davidson distinguishes between what he calls "prior theories" and "passing theories", and characterizes them separately in the case of the speaker and the hearer:

> For the hearer, the prior theory expresses how he is prepared in advance to interpret an utterance of the speaker, while the passing theory is how he *does* interpret the utterance. For the speaker, the prior theory is what he *believes* the interpreter's prior theory to be, while the passing theory is the theory that he *intends* the interpreter to use. (1986, p. 101)

Prior theories contain what the speaker or the interpreter knows before the beginning of their particular conversation. Such theories can contain the conventional meanings of words as well as any other general knowledge and information that each of them has so far learnt from the other – for instance, the way the speaker used a word in previous conversations, and the like. When the interpreter starts his conversation with the speaker, he is equipped with such a theory: he comes with certain expectations. The speaker too comes to the conversation with her own theory, regarding the expectations she has of the interpreter's abilities. According to Davidson, sharing such prior theories is neither necessary nor sufficient for guaranteeing success in communication, since what matters is what the speaker now intends to mean by her words, which may not be compatible with what the interpreter is prepared for. Passing theories, however, contain what the speaker presently intends to mean by her words on this particular occasion. For instance, regarding the example of Mrs. Malaprop, according to the interpreter's prior theory, "epitaph" means *epitaph*, that is, it just means what others in his speech community mean by it. According to Mrs. Malaprop's prior and passing theory, however, "epitaph" means *epithet*, since using "epitaph" in this way is one of her verbal habits. This means that the interpreter's and the speaker's prior theories may not be the same. Nonetheless, if the hearer understands what Mrs. Malaprop intends to mean by her word, then what they actually converge on is a passing theory specifying "epitaph" meaning *epithet*, or "derangement" meaning *arrangement*. As Davidson indicates,

> Mrs. Malaprop's theory, prior and passing, is that 'A nice derangement of epitaphs' means a nice arrangement of epithets. An interpreter who, as we say, knows English, but does not know the verbal habits of Mrs. Malaprop, has a prior theory according to which 'A nice derangement of epitaphs' means a nice derangement of epitaphs; but his passing theory agrees with that of Mrs. Malaprop if he understands her words. (1986, pp. 103-104)

However, the question is whether "epitaph" meaning *epithet* has to be considered as a fixed convention determining the meaning of "epitaph" for Mrs. Malaprop hereafter.

Suppose that, no long after Mrs. Malaprop's last utterance mentioned above, she utters "Newton had an intense epitaph". On this occasion and at this time, however, "epitaph" is not intended by Mrs. Malaprop to mean *epitaph* or even *epithet*. What she intended to mean by "epitaph" is rather *epiphany*. The interpreter's *prior* theory contains "epitaph" meaning *epithet*, since it is the meaning that Mrs. Malaprop attached to her word in the previous conversation. However, his prior theory should be improved in a way to produce "epitaph" meaning *epiphany* if he is to have successful communication with Mrs. Malaprop on this particular occasion. In other words, the interpreter's passing theory will contain "epitaph" meaning *epiphany*, which is what Mrs. Malaprop's passing theory on this occasion contains as well. This means that prior theories are not supposed to contain only the conventional meanings of words; rather, any meanings that have been attached to the words in advance of this particular conversation are within such theories. As Davidson indicates, "the interpreter's theory has been adjusted to the evidence so far available to him: knowledge of the character, dress, role, sex, of the speaker, and whatever else has been gained by observing the speaker's behaviour, linguistic or otherwise" (1986, p. 100). However, according to Davidson, communication is successful simply if their passing theories coincide, that is, if the interpreter interprets the speaker in the way the speaker intended her utterance to be interpreted. Their passing theory produces what the speaker now means by her words, regardless of what she meant before by the same words or what her speech community means by them. This is, therefore, another manifestation of Davidson's argument against the Common View. Following some specific conventions or rules is *neither necessary nor sufficient* for successful communication:

(I) It is not necessary because we can understand what Mrs. Malaprop means by "epitaph" on this specific occasion without appealing to the conventional meanings of her words.

(II) It is not sufficient because in order to understand what Mrs. Malaprop intends to mean by "epitaph", even if she means what "epitaph" conventionally does, we need knowledge and information over and above just the knowledge of what "epitaph" conventionally means.

In other words, even if Mrs. Malaprop intends to use her words in the standard way, knowing the conventional or standard meaning of "epitaph" is not by itself sufficient to understand what she means by her words because, in addition to this knowledge, the

interpreter must also understand that Mrs. Malaprop intends her words to be understood as meaning what they conventionally mean. In order to reach such a conclusion the interpreter has to use the available evidence and clues – e.g. that Mrs. Malaprop does not seem to follow her verbal habits this time, that she uses her words in the standard way when she is happy, or non-standardly when she is drunk, and so on. Even if a speaker speaks as others do, she *intends* to do so, that is, she intends her words to be interpreted as such: "Even when a speaker is speaking in accord with a socially acceptable theory he speaks with the intention of being understood in a certain way, and this intention depends on his beliefs about his audience, in particular how he believes or assumes they will understand him" (Davidson, 1994, p. 122). This knowledge is what the mere knowledge of the conventional meaning of words does not include: there is more to successful communication than just sharing what the words conventionally mean. In this regard, although Davidson accepts that in order to have successful communication, the interpreter must understand what the speaker means by her words – and, in this sense, this meaning must be shared by the speaker and the interpreter – he denies that it is necessary or sufficient to take such meanings to be prepared in advance and conventional: "What I denied was that such sharing is sufficient to explain our actual communicative achievements, and more important, I denied that even such limited sharing is necessary" (1994, p. 110).

Therefore, according to Davidson, what leads the interpreter and the speaker to have successful communication is converging on or sharing their passing theory, not the prior theories: "It is not a condition on successful communication that prior theories be shared" (1986, p. 103). Prior theories are formed in advance of a particular conversation, and hence they are not the sort of theories that must be shared in order to produce the correct interpretation of the particular utterance the speaker makes on a particular occasion. Moreover, reaching the agreement on what the speaker intends to mean by her words, i.e. converging on passing theories, is not something that can be formalized or characterized in terms of some previously fixed conventions. The abilities that the communicators possess leading them to the convergent passing theory and the process through which they come to such convergence cannot be explained in terms of their following some conventions or rules. As Davidson emphasizes, for successful communication, what two people need is "the ability to converge on passing theories from utterance to utterance" (1986, p. 106), but such abilities can be acquired in different ways so that the communicators' strategies to reach such convergence would

be different. In other words, they have different backgrounds of knowledge, information, social status, family, personality, habits, and so on. If that is true, then we should conclude that "there are no rules for arriving at passing theories" (1986, p. 107). There are no such rules because to converge on a passing theory much non-linguistic knowledge and general information is required: "A passing theory really is like a theory at least in this, that it is derived by wit, luck, and wisdom from a private vocabulary and grammar, knowledge of the ways people get their point across, and rules of thumb for figuring out what deviations from the dictionary are most likely" (1986, p. 107). For this reason, Davidson believes that "there is no more chance of regularizing, or teaching, this process than there is of regularizing or teaching the process of creating new theories to cope with new data in any field" (1986, p. 107). What must be shared in order to have successful communication (i.e. the passing theory) is not necessarily learnt in advance of the particular conversation the speakers have, and what is learnt in advance (i.e. the prior theory) does not have to be shared in order to have successful communication. [32]

2.1.5. Passing Theories as Systematic Theories

Davidson criticized the three principles that the Common View imposed on first meanings, that is, that they must be systematic, shared, and conventional. Since it is the passing theory that is supposed to generate the first meanings, we can take the Common

[32] It is, however, important to note that Davidson does not claim that prior theories are entirely irrelevant to success in communication: "Of course things previously learned were essential to arriving at the passing theory" (1986, p.103). The way we can speak and understand is somehow limited to the sort of general knowledge we have of each other and of the world. In other words, the speaker, when she starts speaking, takes into account the interpreter's ability to understand her speech and the sort of knowledge he has of the speaker's behaviour, attitudes, life, and the like: "The speaker's view of the interpreter's theory is not irrelevant to what he says, nor to what he means by his words" (1986, p. 101). Not only does a hearer interpret "(normally without thought or pause) on the basis of many clues", but also the speaker "must bear many of these things in mind when he speaks, since it is up to him to try to be understood" (1984a, p. 12). The point, however, is that sharing what we have learnt before is not by itself necessary or sufficient to have successful communication with the speaker. We may gain a better understanding of the role of prior theories by considering Davidson's idea that "a speaker or hearer uses what he knows in advance plus present data to produce a passing theory" (1986, p. 106). Prior theories are improved after every particular conversation: "As speaker and interpreter talk, their prior theories become more alike; so do their passing theories" (1986, p. 102). But, as we saw, what is essential to successful communication is sharing the passing theory: Each conversation may present new ways of speaking and new meanings of words, which are not present even in the best-improved prior theories. Therefore, according to Davidson's view, the linguistic ability to speak and understand cannot be explained in terms of sharing prior theories: "Linguistic ability is the ability to converge on a passing theory from time to time – this is what I have suggested, and I have no better proposal" (1986, p. 107).

View's claim to be that passing theories must be systematic, shared, and conventional. While Davidson accepts that passing theories are systematic and shared, he strongly rejects the requirement that they are essentially fixed and conventional. The passing theories, in this Davidsonian picture, are the theories that change from time to time, from utterance to utterance, and from occasion to occasion. How can such a theory be systematic? How can such a theory possess the sort of characteristics that Davidson, in his earlier works, attributed to Tarski-style theories of meaning? We can best see the apparent difficulty that the passing theories face by looking at again the passing theory that the interpreter of Mrs. Malaprop's utterance uses: according to Mrs. Malaprop's and the interpreter's passing theory, "epitaph" means *epithet*. And there is no guarantee that next time Mrs. Malaprop uses "epitaph" to mean the same thing; she may mean something different and the interpreter's passing theory must produce Mrs. Malaprop's new intended meaning. However, if such a passing theory is still supposed to be a Tarski-style theory of meaning, in what sense can it be regarded as systematic, that is, compositional and recursive? Recall that a Tarski-style theory of meaning was a compositional theory generating the truth-conditions of all the potential utterances a speaker speaking a language could produce.[33] Each word, however, can appear in indefinitely many different sentences, and the meaning of those sentences depends on the semantic features of their parts and the way they contribute to the meaning of the sentences as a whole. For instance, if "epitaph" means *epithet*, then the meaning of the utterance "There are no epitaphs" will be that *there are no epithets*, and the like.

Davidson indeed believes that passing theories are still Tarski-style theories of truth with all the characteristics he attributed to them in his earlier works, and, in this sense, Davidson does not tend to give up on any of his main earlier ideas on these theories. For Davidson, we still need to "account for the ability to interpret utterances of novel sentences" and "this is possible because the interpreter can learn the semantic role of each of a finite number of words and phrases and can learn the semantic consequences of a finite number of modes of composition" (1986, p. 95). Also, "since the modes of composition can be iterated, there is no clear upper limit to the number of sentences utterances of which can be interpreted" (1986, p. 95). As before, this theory is to be taken as a device producing interpretations when it is fed by utterances. In this regard, Davidson emphasizes that "one model for such a machine is a theory of truth, more or

[33] For an exposition of Davidson's earlier view, see the Appendix: The Early Davidson.

less along the lines of a Tarski truth definition. It provides a recursive characterization of the truth-conditions of all possible utterances in terms of sentences made up from the finite vocabulary and the finite stock of modes of composition" (1986, p. 95). Nonetheless, as Davidson himself asks, "why should a passing theory be called a theory at all?" (1986, p. 103). Indeed, Davidson himself is aware of the contrast between the sort of rigidity that his Tarski-style theory of truth was taken to have in his earlier works and the sort of flexibility that passing theories are now supposed to have. Nonetheless, Davidson thinks that passing theories can still be taken as theories for an entire language, "even though its expected field of application is vanishingly small" (1986, p. 103).

I think that, to understand in what sense a passing theory is a theory of a language, we need to concentrate on the idea that such a theory is designed to generate an interpretation of all the possible utterances a speaker *would* make if she *continued* speaking that language. In other words, if Mrs. Malaprop intended to mean *epithet* by "epitaph", and if she intended to speak in this way from now on, the passing theory would become a Tarski-style theory of truth producing the meaning of all the potential utterances Mrs. Malaprop would make by speaking this language. For instance, this theory would contain the following theorem:

> (T) For the speaker S, at time t, "That's a nice derangement of epitaphs" is true if and only if that is a nice arrangement of epithets.

This theorem is inferred from the following axioms:

> (A1) For the speaker S, at time t, "epitaph" refers to epithets.
>
> (A2) For the speaker S, at time t, "*x* is a nice derangement of something" is satisfied by *x* if and only if *x* is a nice arrangement of something.

This theory would specify the meaning of all the possible utterances in which these parts could occur, together with, of course, theorems capable of specifying the meaning of all other potential utterances Mrs. Malaprop would make in general. Indeed, this theory can be taken as producing an interpretation of all the possible utterances that such a language might contain, *though there is no guarantee, no requirement, that the speaker continues speaking this particular language*. In this sense, what a passing theory, as used by an interpreter in a particular conversation with a speaker, *actually*

produces is an interpretation of a very limited number of, or perhaps just one of, the actual utterances of such a language. Mrs. Malaprop may intend to change the language she speaks by intending to mean *epiphany* by "epitaph" this time. This spoken language is not the same language as the one Mrs. Malaprop would have spoken if she had continued speaking the way she did before. In this regard, the new passing theory would change accordingly:

> (T) For the speaker S, at time t, "Newton had an intense epitaph" is true if and only if Newton had an intense epiphany.
>
> (A1) For the speaker S, at time t, "Newton" refers to Newton.
>
> (A2) For the speaker S, at time t, "*x* had an intense epitaph" is satisfied by *x* if and only if *x* had an intense epiphany.

Hence, passing theories are still Tarski-style theories of truth. As Davidson says,

> when a word temporarily or locally takes over the role of some other word or phrase (as treated in a prior theory, perhaps), the entire burden of that role, with all its implications for logical relations to other words, phrases, and sentences, must be carried along by the passing theory. (1986, p. 103)

In this regard, "someone who grasps the fact that Mrs. Malaprop means 'epithet' when she says 'epitaph' must give 'epithet' all the powers 'epitaph' has for many other people" (Davidson, 1986, p. 103). The same is true in the case of introducing new words with new meanings or giving an entirely new meaning to an old word. As a result, according to Davidson, "only a full recursive theory can do justice to these powers" (1986, p. 103).

In this way, Davidson holds onto the most important feature of his earlier works: we must use a theory of truth to describe the speakers' linguistic abilities, though these abilities are no longer regarded (or presupposed) to be fixed in advance of particular conversations. While in the *radical* interpretation scenario the interpreter takes it for granted that the meanings of the speaker's utterances are stable in general, here Davidson treats meanings and the speakers' linguistic abilities to be less stable and apt to change occasion by occasion.[34] Therefore, the account of the process of interpretation

[34] Glüer (2013) compares Radical Interpretation with Davidson's later remarks on interpretation. For her, radical interpretation involves cases in which other alternative ways of interpreting are not available; hence the only option for successful communication is to appeal to the regular application of words. See e.g. (Heal, 1997) for more discussion of this comparison.

and the idea of a theory of truth as a theory of interpretation are the two main doctrines of Davidson's earlier works which remain completely alive and at work in his later account of meaning.

2.1.6. "There Is No Language" Claim

However, if passing theories are theories of whole languages, and if passing theories are theories which can change all the time, then the conception of a language that the Davidsonian view is introducing does not seem to be compatible with the one some philosophers usually talk about. Indeed, a consequence of Davidson's criticisms of the Common View is a rejection of the conception of a language that the Common View advocates, i.e. language as a rigid set of semantical and syntactical rules and conventions, which are fixed in advance and sharing which is necessary and sufficient to account for success in communication. Davidson argued that communicators do not need to speak the same language, or in this sense, a common language at all. In other words, even if there is such a language, sharing it is neither necessary nor sufficient for successful communication to take place. Davidson's remarks on prior and passing theories led to such an idea: neither passing theories, nor prior theories, appear to be theories about such a commonly shared and fixed language. Passing theories contained a changing list of vocabularies and divergent uses of words. They are not stable in their nature: "Every deviation from ordinary usage, as long as it is agreed on for the moment … is in the passing theory as a feature of what the words mean on that occasion" (Davidson, 1986, p. 102). Because of this feature, a passing theory cannot really be a theory of a "language", i.e. a language as a set of fixed conventions and rules, which is capable of being learned or mastered in advance of a particular conversation: "A passing theory is not a theory of what anyone (except perhaps a philosopher) would call an actual natural language" (Davidson, 1986, p. 102). Not only cannot such a language be learnt, but also the "'Mastery" of such a language would be useless, since knowing a passing theory is only knowing how to interpret a particular utterance on a particular occasion" (1986, p. 102).

Regarding prior theories, although they are established in advance of actual communicative exchanges, they are usually different from one another: "The prior theory has in it all the features special to the idiolect of the speaker that the interpreter is

in a position to take into account before the utterance begins" (Davidson, 1986, p.104). In this way, an interpreter would have different prior theories for different speakers. For instance, as we saw in the example of Mrs. Malaprop, the interpreter's prior theory for his first conversation with Mrs. Malaprop contained an axiom specifying "epitaph" meaning *epitaph*; his theory for the next conversation, when he found out about Mrs. Malaprop's verbal habits, contained "epitaph" meaning *epithet*. Our interpreter's prior theory for another speaker would probably be different. Thus, even prior theories cannot be theories of a natural language as the Common View defines it.[35] If that is the case, then, as Davidson says, "neither the prior theory nor the passing theory describes what we would call the language a person knows, and neither theory characterizes a speaker's or interpreter's linguistic competence" (1986, p. 104).[36]

As a consequence of these remarks, Davidson concludes that "there is no such a thing as a language" (Davidson, 1994, p. 109). This claim would be entirely astonishing and confusing if the reader does not take into account the conception of a language that Davidson is now casting doubt on, since there really seems to be a language that we speak everyday. The point that is sometimes missed by those surprised by these remarks from Davidson is the proviso following the above claim: "Not if a language is anything like what many philosophers and linguists have supposed" (Davidson, 1994, pp. 109-110). As Davidson emphasizes, "there is no such thing as what some philosophers … have *called* a language" (1994, p. 110). The conception of a language which many

[35] According to Stroud (1998), "Standing" or "Prior" Theory is "a general characterization of the meanings of expressions in the language which we might possess in advance of being presented with any "occasions of interpretation" from Mrs. Malaprop" (1998, p. 10). Passing theory, however, is "something that works at the moment for the particular utterance [Mrs. Malaprop] has made. That is, we assign meanings to some of her words as used on that occasion which differ from the meanings assigned to those words by the standing or prior theory of English" (1998, p. 10). The way Stroud characterizes prior theory, however, does not seem to be quite compatible with the way Davidson does: a prior theory is not necessarily supposed to be a "*general* characterization of the meanings of expressions" in English, or that we are equipped with such a theory "in advance of being presented with *any*" communicative exchange with Mrs. Malaprop. Rather, a prior theory specifies the meanings that the expressions are given, no matter whether by Mrs. Malaprop herself in the past or by her speech community, *prior* to the new utterance Mrs. Malaprop makes using those expressions. Our example about Mrs. Malaprop's different ways of using "epitaph" clarifies this point.

[36] As Davidson summarizes this idea, "what interpreter and speaker share, to the extent that communication succeeds, is not learnt and so is not a language governed by rules or conventions known to speaker and interpreter in advance; but what the speaker and interpreter know in advance is not (necessarily) shared, and so is not a language governed by shared rules or conventions" (1986, pp. 105-106).

philosophers have advocated was indeed the one the Common View suggested,[37] that is, the view that "in learning a language, a person acquires the ability to operate in accord with a precise and specifiable set of syntactic and semantic rules; verbal communication depends on speaker and hearer sharing such ability, and it requires no more than this" (Davidson, 1994, p. 110). The Common View claimed that because of the mastery every speaker acquires of such a set of shared rules – a common language – they will become capable of understanding each other. Davidson, however, rejected such a view: there is no such thing as a language in the sense which the Common View described.

One important point that we need to take into consideration is that Davidson's target is a *philosophical* clarification of "what is necessary and sufficient for such understanding" (Davidson, 1994, p. 112). As a matter of fact, we mostly intend to speak more or less as others do or follow the same rules they follow, since we want to be more easily understood and to have smoother communication with others: "I did not deny that in practice people usually depend on a supply of words and syntactic devices which they learned to employ in similar ways" (Davidson, 1994, p. 110). However, as a philosophical explanation of the necessary and sufficient conditions for successful communication, "sharing such a previously mastered ability was neither necessary nor sufficient for successful linguistic communication" (Davidson, 1994, p. 110). Hence, although people may possess different linguistic abilities when they come to different occasions of speech, mutual understanding is still achievable because converging on passing theories is all we need, and achieving such a convergence does not depend on sharing what we learnt before; rather, it needs "the exercise of imagination, appeal to general knowledge of the world, and awareness of human interests and attitudes" (Davidson, 1994, p. 110). Although the speaker and the interpreter may have their own unique idiolect, mutual understanding can be gained if the interpreter can be led by the speaker to the intended meaning of the words she utters. This is why, for Davidson, the notion of the idiolect, i.e. the special way each speaker intends to speak, is of more philosophical importance than the notion of a common language. As Davidson specifies, hence, "there is no such thing as a language apart from the sounds and marks people

[37] When we considered Davidson's criticism of the Common View, we became familiar with some of the supporters of the Common View, such as Michael Dummett and David Lewis (see Sections 2.1.2.a and 2.1.2.c, fn. 25). Davidson even counts himself as advocating the Common View in his earlier works (see Davidson, 1994, p. 110). He also attributes this view to Tyler Burge and Kripke, as he states: "If I am right, then the important claims by Tyler Burge, Saul Kripke, and perhaps Wittgenstein and Dummett must be false" (1994, p. 119).

make, and the habits and expectations that go with them" (1997b, p. 131).

2.1.7. Linguistic Error

Davidson rejects the Common View, together with its accompanying notions and conceptions, such as its conception of a language (as a fixed set of rules) or its conception of meaning (as essentially conventional). Davidson also rejects the conception of linguistic error that the Common View implied. According to the Common View, failing to follow the conventional way of using words is counted as an error or mistake. Davidson's idea is that this conception of error is problematic too: "Error or mistake of this kind, with its associated notion of correct usage, is not philosophically interesting. We want a deeper notion of what words, when spoken in context, mean" (1986, p. 91). This deeper notion of meaning was the notion of first meaning, and so the notion of linguistic error must be defined in conformity with this notion. A speaker is allowed to change the meaning of her words and thereby to use them in a different way, provided that it is understandable for her hearer. In this regard, the speaker *correctly* or successfully means something specific by her words if the interpreter can understand what she intends to mean by those words. Failing to do so, that is, failing to be understood – and not failing to use words as others do – is to be counted as making a mistake. As Davidson clarifies, two speakers "could both have been wrong and things would have gone as smoothly" (1986, p. 90). They are "wrong" in the sense that the Common View takes linguistic mistakes to be so. But it seems unimportant to Davidson "who makes a [standard or conventional] mistake, or whether there is one" (1986, p. 90). Putting it in terms of theories of meaning, since the correct theory is the theory generating the first meanings of words, i.e. the passing theory, the interpreter cannot be taken to have a correct theory *in advance of* his particular conversation with the speaker; rather, she should try to adjust her theory "so that it yields the speaker's intended interpretation" (Davidson, 1986, p. 99). In this way, the Common View's conception of linguistic error is rejected by Davidson and replaced by a conception of error, according to which making a mistake is a matter of failing to be understood in a particular way by the hearer.

2.1.8. Davidson's General Argument against the Common View

In the above sections, we introduced Davidson's criticisms of the Common View or Conventionalism in detail. In this part, our aim is to characterize this argument in a more general way, that is, in a way similar to the way we characterized KW's sceptic's sceptical argument. Davidson used the phenomenon of malapropism to argue against the Common View with the purpose of showing how implausible the consequences of this view are for our explanation of success in linguistic communication. We can take this argument to be a sceptical argument which Davidson's alleged "sceptic" proposes to rule out a certain conception of meaning, as KW's sceptic proposed a sceptical argument to cast doubt on Traditional Realism's conception of meaning. Therefore, we can put Davidson's sceptic's negative argument as follows. If we take the Common View for granted, then we will come up with a sceptical conclusion about the linguistic practice of Mrs. Malaprop:

(I) **CV** (The Common View): If Mrs. Malaprop means something by "epitaph", then there has to be a fixed convention or rule, which determines the correct use of "epitaph" for her and in terms of shared grasp of which her ability to communicate with others is to be explained.

(II) **BSC_D** (The Basic Sceptical Conclusion): There is no rigid set of rules or conventions, which determines the correct use of "epitaph" for Mrs. Malaprop and in terms of shared grasp of which her successful communication with others employing "epitaph" can be explained.

(III) **RSC_D** (The Radical Sceptical Conclusion): Mrs. Malaprop means nothing by her utterance of "epitaph".

If the Common View is true, then Mrs. Malaprop's utterance fails to make any sense. Davidson believes that what is argued for in the case of Mrs. Malaprop can be properly generalized: "Mrs. Malaprop … make[s] the case general. There is no word or construction that cannot be converted to a new use by an ingenious or ignorant speaker" (1994, p. 100). In this regard, the argument can be put in more general terms:

(I) **CV**: If a speaker means something by a word, then there has to be a fixed convention or rule, which governs the correct use of that word for her and in terms of shared grasp of which her ability to communicate with others is to

be explained.

(II) **BSC_D**: There is no rigid set of rules or conventions, which determines the correct use of a word for a speaker and in terms of shared grasp of which her successful communication with others employing that word can be explained.

(III) **RSC_D**: There is no such thing as meaning anything by any word.

Therefore, this sceptical argument shows how the Common View can lead to an entirely implausible conclusion, i.e. (RSC_D). As already mentioned, Davidson does not believe that what Mrs. Malaprop means by her words is absurd or that she means nothing by her words. Rather, he rejects the view resulting in the Radical Sceptical Conclusion. We will see in detail how Davidson's alternative view arises from these criticisms of the Common View.

2.2. The Positive Part: Davidson's Alternative View

In this section, we will introduce Davidson's alternative account of the practice of meaning something by words. This alternative proposal is intended by Davidson to provide a more plausible and comprehensive explanation of the necessary and sufficient conditions for successful communication.

2.2.1. Davidson's Alternative View of Meaning

Davidson's argument against the Common View resulted in the Radical Sceptical Conclusion, according to which if the Common View is accepted, then Mrs. Malaprop means nothing by her utterance of "epitaph". According to Davidson, however, this conclusion is obviously false. For him, malapropisms are not meaningless: it is a "fact that in all these cases the hearer has no trouble understanding the speaker in the way the speaker intends" (1986, p. 90). This means that Davidson rejects the Radical Sceptical Conclusion, (RSC_D). Indeed, Davidson's alternative view starts with a rejection of the Radical Sceptical Conclusion and an acceptance of the Basic Sceptical Conclusion, (BSC_D), according to which success in communication does not essentially depend on the speaker and hearer following some fixed set of rules or conventions, just as KW's

sceptical solution began with a rejection of KW's sceptic's Radical Sceptical Conclusion and an acceptance of his Basic Sceptical Conclusion. KW thought that the Radical Sceptical Conclusion is a result of presupposing a mistaken conception of meaning, that is, Traditional Realism. Similarly, Davidson can be taken to claim that the implausibility of the Radical Sceptical Conclusion of his alleged argument reveals the implausibility of the Common View, presupposing which led to such an intolerable conclusion. Hence, the Common View has to be rejected. In this way, we can characterize Davidson's argument for his alternative view as follows:

(I) **BSC_D**: There is no rigid set of rules or conventions, which determines the correct use of words for a speaker and in terms of shared grasp of which her successful communication with others regarding those words can be explained.

(II) **$\neg RSC_D$**: It is not the case that there is no such thing as meaning anything by any word. (We can smoothly understand what Mrs. Malaprop means by her words.)

(III) **¬CV**: Therefore, it is not the case that if a speaker means something by a word, then there is a fixed convention or rule, which governs the correct use of that word for her and in terms of shared grasp of which her ability to communicate with others is to be explained.

According to Davidson, all we need to explain the success in communication between Mrs. Malaprop and the hearer, and in general between any two speakers, is to explain how the hearer can successfully understand the speaker in the way the speaker intended to be understood. For Davidson, such mutual understanding can take place simply if there is enough evidence to enable the interpreter to interpret the speaker's utterances in the way the speaker intended them to be interpreted. This idea is the main clue for understanding Davidson's alternative solution: "What a speaker intends by what he says determines what he means" (Davidson, 1991b, p. 143). Davidson's alternative view takes the role of propositional attitudes, especially intentions, to be essential to the practice of meaning something by words. However, what exactly is the relation between the speaker's intentions and her success in meaning something by her words?

(a) Intention and Mutual Interpretation

As explained in the Appendix, according to Davidson's earlier remarks on meaning and propositional attitudes, meanings and beliefs, or more generally, propositional attitudes, are deeply interconnected so that the radical interpreter takes the speaker's utterances to be expressing what the speaker believes. Davidson's later view still heavily relies on the interdependence of meaning and propositional attitudes. This time, intentions play the main role in his alternative account. As Davidson states, "an interpreter (correctly) interprets an utterance of a speaker only if he knows that the speaker intends the interpreter to assign certain truth conditions to his (the speaker's) utterance" (1992, pp. 111-112). Success in communication depends on the speaker's intending to mean something that the interpreter can understand. For this reason, the speaker has freedom to intend to mean anything she may, *provided that* the evidence and clues suffice to lead the interpreter to a grasp of the speaker's intended meaning. Indeed, as Davidson explains, "if a speaker wishes to be understood, he must intend his words to be interpreted in a certain way, and so must intend to provide his audience with the clues they need to arrive at the intended interpretation" (1987, p. 28). This claim is directly related to Davidson's distinction between first meaning and conventional meaning: the first meaning of a speaker's utterance is the meaning she intended to attach to her words and, when communication is successful, is correctly understood by her interpreter. It does not matter whether this meaning is conventional or not. However, Davidson believed that when a speaker utters a sentence, she always has further intentions and purposes in mind. For instance, she may intend to assert something, with the intention of being ironical, sarcastic, and so on. How can Davidson's view account for such notions?

(b) First, Conventional, and Speaker Meaning, and Ulterior Purposes

Not only does Davidson believe that his alternative view can preserve the notion of literal or first meaning, but he also thinks that his view can accommodate all the further meanings and intentions that a speaker may have in her communication with the interpreter. For Davidson, there are different sorts of intentions and purposes that a speaker has in mind when she utters a sentence, and the interpreter must grasp such intentions and purposes if he is to have successful communication with the speaker:

> (1) In uttering a sentence, the speaker intends to utter words that in the circumstances will be interpreted as having a certain literal meaning; (2) through the recognition of this meaning by an audience, the speaker intends to be understood as making a particular assertion, asking a particular question, … etc.; (3) by means of uttering words with the intended interpretation and with the intended force (also correctly interpreted), the speaker intends to accomplish some ulterior (non-linguistic) purpose. (1984c, p. 21)

Davidson called the literal meaning of words their first meaning because it was the meaning which has to be grasped first by the interpreter if he is to understand the speaker's further meanings, intentions, and purposes. As we saw earlier, according to Davidson, the first meanings and the conventional meanings do not have to coincide all the time; nor does it have to be the case that grasping the speaker's further intentions and meanings necessarily depends on a grasp of the conventional meanings of her words. For example, when Mrs. Malaprop with her mentioned verbal habits uttered "A nice derangement of epitaphs", not only could she be interpreted as intending to mean something specific by her words, but she could also be taken to be performing some specific speech act (such as making an assertion, asking a question, and the like), having some particular speaker meaning (e.g. being ironical, sarcastic, and the like), and having certain ulterior (non-linguistic) purposes. Suppose that she utters the above sentence with the intention of making an assertion. According to Davidson's account, we will have the following distinctions:

(I) Conventional Meaning: That's a nice derangement of epitaphs.

(II) First Meaning: That's a nice arrangement of epithets.

(III) Illocutionary Force: Asserting that that's a nice arrangement of epithets.

Moreover, we can employ Davidson's example again to show how Mrs. Malaprop's utterance can be used with different purposes and further intentions. For instance, the sentence, "A nice derangement of epitaphs", with the literal (i.e. first) meaning that it has can be ironically used by Mrs. Malaprop's sister:

(IV) Speaker Meaning: That's not a nice arrangement of epithets.

Mrs. Malaprop's sister used the same sentence with the meaning that Mrs. Malaprop intended this utterance to have to make the ironical point that the relevant epithets are not nicely arranged. Finally, Mrs. Malaprop or her sister will have specific ulterior (i.e. non-linguistic) purposes in mind when they utter that sentence with that first meaning,

such as:

> (V) Ulterior Purposes: Impressing the hearer, making the audience happy, teasing the hearer, etc.

In this picture, the first meanings of the speaker's words are not necessarily the conventional meanings of the words, but they are the meanings that must be interpreted by the interpreter in the way the speaker intended them to be understood if the interpreter is to understand what other meanings and intentions the speaker has in mind by her utterance.[38]

Moreover, we can see that speaker meanings are not really *meanings* in the ordinary sense of the term, though they are so called. They are rather the further effects which the speaker intends to have on her audience in virtue of her utterance being understood by the audience in a certain way. These effects are understood only if the speaker's words are interpreted in the way she intended them to be. There are, of course, cases in which the conventional meanings of words and their first meanings coincide, e.g. when the speaker intends to mean what the words conventionally mean. However, it was reaching an agreement on what the speaker intends to mean by her words that played the main role in Davidson's alternative view: it is converging on passing theories that is essential, not prior theories.

Davidson's alternative view, hence, has an important individualistic feature: it assigns a considerable role to the individual speaker and her intentions to use words in certain ways. It is ultimately up to the speaker to decide what she wants to mean by her words. Nonetheless, this view is not purely individualistic; rather, it embraces a fundamental social element. What is the social aspect of Davidson's account?

(c) Davidson's Interpersonal View

Although Davidson allowed the individual speaker to follow the linguistic practice of her own, that is, to use words in a way she thinks she should, he puts a social constraint on such a practice. The fundamental idea in Davidson's account is the demand that the

[38] Our example clarifies what Davidson meant by "first meaning". As Ludwig and Lepore (2005) state, "first meaning is what a speaker intends his words to be understood to have so as to form the basis for subsequent effects achieved by his using the words he does" (2005, p. 266).

speaker must speak in an understandable way. This requirement puts a restriction on the way the speaker is allowed to speak and the meanings she is allowed to intend. In other words, Davidson's individualism about meaning is thereby constrained by a certain social limitation. To clarify this constraint, Davidson contrasts the case of Mrs. Malaprop with the case of Humpty Dumpty, i.e. the case of a sheer innovator who does not care about the linguistic abilities of his hearer to understand his words. As Humpty Dumpty says, "when *I* use a word it means just what I choose it to mean" (Davidson, 1991b, p. 143). While Davidson generally accepts this claim, he insists that this can be treated as a plausible claim only if the hearer can be in a position to understand what Humpty Dumpty means by his words, that is, if the hearer is somehow in a position to interpret what Humpty Dumpty says. Davidson gives an example to clarify his point. Imagine that Humpty Dumpty, at the end of his speech with Alice, "says to Alice, "There's glory for you!" Alice says she doesn't know what he means. Humpty Dumpty replies, "Of course you don't – till I tell you. I meant 'There's a nice knock-down argument for you!'" (1991b, p. 147). According to Davidson, this position is just wrong. His alternative view does not permit such an unconstrained freedom in speech. Rather, Davidson has always insisted on the idea that the speaker is required to take into account the abilities, knowledge, and general background information of her hearer: "In speaking or writing we intend to be understood. We cannot intend what we know to be impossible" (1991b, p. 147). Thus, although Davidson's view rejects the idea that following fixed conventions, or speaking as others do, is essential for two people to have successful communication, it does not entail a Humpty-Dumpty sort of theory: if the speaker fails to speak in a way capable of being understood by her hearer, she fails to mean anything by her words. It is, therefore, understanding that is fundamental, rather than following rules determining the correct use of words: "Where understanding matches intent we can, if we please, speak of 'the' meaning; but it is understanding that gives life to meaning, not the other way around" (1994, p.121).[39] Meanings appear as the result of successful communication: the speaker means something by her words if she intends to mean that thing and her utterance is understood by her hearer as having that meaning.

Hence, we can see that Davidson's later works on meaning were mostly dedicated to providing an explanation of the extent to which a social element must enter into our

[39] As Davidson says elsewhere, "someone can't mean something by his words that can't be correctly deciphered by another" (1987, p 28).

account of meaning and linguistic understanding. Indeed, we can take Davidson to be distinguishing between two different conceptions of such a social element: a "wider" conception and a "narrower" one. In its wider sense, the social element involves a wide agreement within a community of speakers about the meaning of words. The use of words and their meanings are fixed in advance of particular conversations on the basis of certain conventions. This wider notion of social setting was offered by the Common View suggesting that speaking as others do is essential to the existence of successful communication. Davidson, however, believed that "we could get along without it" (1994, p. 114). Understanding others does not necessarily depend on our mastery of some rigid set of rules. Rather, it depends on the interpreter's ability to use the available evidence and clues to interpret the speaker's utterances and on the speaker's ability to speak in an understandable way. This is, for Davidson, the correct conception of a social setting, which is more fundamental than the Common View's suggested conception. As Davidson states,

> I thought I saw (and see) clear reasons to doubt that language, if language is taken to imply shared ways of speaking, is essential. The same doubts apply to the notion of following a rule, engaging in a practice, or conforming to conventions, *if these are taken to imply such sharing*. (Please note the proviso). (1994, p. 115)

As a matter of theoretical and philosophical concern, there is no reason why someone should speak as others do if understanding is still achievable without conforming to the way others speak: "Linguistic communication does not require … rule-governed repetition" at all (1984b, pp. 279-280).[40] What is basic to linguistic communication is mutual interpretation, and if the point of following rules is to explain success in communication, then the fact that we can successfully communicate with others without following such rules shows that they are not essential to our success in communication. Hence, "philosophers who make convention a necessary element in language have the matter backwards" (Davidson, 1984b, p. 280). As Davidson concludes, "what matters, the point of language or speech or whatever you want to call it, is communication,

[40] Indeed, as Davidson emphasizes, it is only after mutual understanding takes place that the notions of convention and rule have the opportunity to emerge: "When we talk of rules of language, we usually have in mind grammarians' or linguists' descriptions (generalized and idealized) of actual practice, or prescriptions grammarians wish we would follow. Rules can be a help in learning a language, but their aid is available only in the acquisition of a second language. Most learning of how to use words is accomplished without explicitly learning any rules at all" (1992, p. 113). In this regard, the role of conventions cannot be essential as "convention does not help explain what is basic to linguistic communication" (1984b, p. 280).

getting across to someone else what you have in mind by means of words that they interpret (understand) as you want them to" (1994, p. 120). This can be achieved without conforming to any fixed convention about the correct use of words.

Davidson's view, hence, can be properly called an "interpersonal view", i.e. the view that "meaning something requires that by and large one follows a practice of one's own, a practice that can be understood by others" (Davidson, 1994, p. 125). This view is in conflict with the Common View, or a Communitarian view of meaning. As Verheggen (2006) explains, an interpersonal view implies that "having a (first) language essentially depends on having used (at least some of) one's words, whatever one means by them, to communicate with others", while the communitarian view demands that "having a (first) language essentially depends on meaning by one's words what members of some community mean by them" (2006, p. 203). Davidson's view implies the first sort of view, not the second: "communication without convention" is what Davidson advocates in his later account of meaning and linguistic understanding.

2.2.2. Davidson's Responses to Objections

Davidson thinks that there are at least three important objections to his alternative view, which need to be discussed and responded to. In the following sections, we introduce these objections and consider Davidson's responses to them, since they lead to a very significant part of Davidson's later works in which he introduces the notion of "triangulation".

(a) Being Obliged to Semantic Norms

The first criticism that Davidson discusses is the claim that his alternative view ignores the fact that speakers of a language are actually "responsible to a social norm even if they do not hold to it" (Davidson, 1994, p. 116), i.e. that they are obligated to "the norm constituted by the 'accepted' meanings of words" (1994, p. 117). In this sense, actual speakers of a speech community seem to feel obligated to conform to the way others in their speech community speak, e.g. to use their words in the way others in their linguistic society do. Davidson, however, thinks that this idea is not philosophically important. For him, it is implausible to think that speakers must be taken to be feeling

obliged to speak as others do. The reason is that someone might be taught to speak in a certain way, for instance, to speak as her teacher does, without feeling any obligation to speak in that way. In other words, while she continues speaking in that particular way, she does not feel obliged to do so; rather, for her it is just the natural way of speaking. As Davidson says, "suppose someone learns to talk as others do, but feels no obligation whatever to do so. For this speaker obligation doesn't enter" (1994, p. 117). If someone does not feel obliged to speak in the way she was taught to speak, we are not entitled to accuse her of being unintelligible, or not to belong to our speech community. She is still a member of this community, whether or not she feels such obligations. Hence, generally speaking, speakers speak in the way they were taught to speak, and it seems natural to them to speak in that way: "What magic ingredient does holding oneself responsible to the usual way of speaking add to the usual way of speaking?" (Davidson, 1994, p.117)

However, one might claim that this response is not plausible, since it neglects the fact that speakers always try to correct themselves; this indicates the sort of obligation they feel about their linguistic practices. What else but feeling obliged to follow certain semantic conventions can explain the phenomenon of correcting mistakes in speaking? Davidson, however, does not think that speakers' correcting themselves has anything to do with their feeling obliged to follow specific semantic norms or conventions. The fact that a speaker corrects herself when she makes a mistake does not reveal the existence and the pressure of semantic obligations; rather, the pressure that we feel is social stemming from our ordinary life in a society. "These pressures are social and they are very real", as are those we feel for correcting ourselves when we are "using the wrong fork at a dinner party" (Davidson, 1994, p. 117). We do not tend to use the wrong fork at a dinner party because it is embarrassing: we want others to see that we know how to behave at such social events. As we saw (Section 2.1.2), Davidson argued that the practice of speaking a language is more similar to the practice of eating at a dinner party than to that of playing and winning a game. Although it is important to use the fork and knife in a certain (socially accepted) way at a dinner party, it is not *essential* if our hunger is to be satisfied: we can eat in whatever way we like. The only obligation we may be committed to when we engage in linguistic communication is the obligation of remaining understandable: "If we can make ourselves understood while deviating from the social norm, any further obligation has nothing to do with meaning or successful communication" (Davidson, 1994, p. 118). As a result, holding ourselves responsible, or

feeling obliged, to socially accepted ways of using words is irrelevant to the success or failure of communication, simply because this requirement is not true in all the situations in which "we will be better understood if we deviate from the 'socially accepted' use" (1994, p. 118).

(b) Practical Possibility of Communication without Convention

The second sort of objection which Davidson discusses has been put forward by those who think "communication without shared practices may be theoretically possible, but argue that this is pointless speculation given that it never occurs in a pure form and probably couldn't" (Davidson, 1994, p. 116). We already know that this objection is unimportant for Davidson because his concern has been philosophical and theoretical. He thinks that speculating on "the remote possibility of there being speakers who, though they express themselves in distinct idiolects, understand one another" is philosophically fruitful, since such an investigation "shows that such sharing cannot be an essential constituent in meaning and communication" (1994, p. 119). Such a possibility indicates the falsehood of any philosophical view which implies that speaking in the socially accepted way is essential to the practice of meaning something by words. In addition, a philosophical investigation of the possibility of communication without convention will reveal the fact that in order to have successful communication we need more than just knowing the conventional meaning of words. As Davidson indicates, if "much successful communing goes on that does not depend on previously learned practices", then we can appreciate the fundamental importance of sharing "general information and familiarity with non-linguistic institutions (a 'way of life')" (1994, p. 119) for understanding even the literal meaning of a speaker's utterance.[41]

(c) The Wittgensteinian Seems Right/Is Right Distinction

The last challenge which Davidson engages with is the objection that his alternative proposal does not provide a satisfactory criterion for evaluating speakers' success or

[41] This is again another manifestation of Davidson's claim that to converge on passing theories, i.e. to understand each other, we need much more knowledge and information than the knowledge of conventional meaning of words.

failure in their practice of meaning something by words. When can we say, based on the Davidsonian account, that the speaker has made a mistake in her use of words? Indeed, Davidson deals with two (related) issues here.

(I) According to Davidson's view, two speakers intending to use their words in different ways can have successful communication with each other. If that is the case, then how can we say of such speakers that they make a mistake in their use of words? If a speaker can intend to mean whatever she may by her words, then it is difficult to provide a norm for assessing the speaker's success or failure in meaning something by the words, especially given that Davidson has already rejected the Common View's criterion, that is, that it is conforming to, or failing to conform to, the previously fixed conventions that determines the speaker's success, or failure, in meaning something by her words. Hence, Davidson is asked to give an alternative criterion for correctness, which is highlighted by the Wittgensteinian idea that there needs to be a distinction between what seems right to a speaker and what is actually right. As initially discussed before (Section 2.1.7), Davidson proposed a criterion, according to which the speaker is successful in her practice of meaning something by an utterance if she is successfully interpreted by at least another speaker, i.e. an interpreter, in the way the speaker intended her utterance to be interpreted. As we will discuss in more detail in the following sections, it is success in mutual interpretation that determines success in the practice of meaning something by an utterance. To this extent, Davidson agrees that his suggested norm must presuppose the existence of a social setting. Such a norm, i.e. a speaker's utterances' "being interpreted as intended", however, presupposes the notions of meaning and intention. Although this norm can provide us with a criterion for assessing the speaker's success in her practice of meaning something by her words, it does not explain how such a speaker has come up with the notion of error or mistake, that is, how such a speaker has mastered the Wittgensteinian seems right/is right distinction through her interaction with another person.

(II) Davidson, however, tends to give an explanation of how a speaker can command the Wittgensteinian seems right/is right distinction as a result of her mutual interaction with another person. He wants to give an argument for the claim that commanding the Wittgensteinian distinction, or as Davidson puts it, acquiring the concept of objective truth, is essential for a creature to possess any language and propositional attitudes at all and that this concept can be acquired only if the creature linguistically interacts with another similar creature. Hence, although it is not possible for one person considered in

isolation to grasp the concept of objective truth and thereby to have any language and thoughts at all, acquiring the concept of objectivity does not require there to be more than two people linguistically communicating with each other. As a result, we do not need to appeal to the shared practice of a speech community to provide a criterion for correctness: actual mutual interpretation between two people is both necessary and sufficient for them to grasp the concept of objectivity and hence to be capable of meaningful responses.

To have a clear picture of Davidson's idea on these two issues, we must first clarify Davidson's suggested norm for evaluating speakers' linguistic behaviour and then investigate his remarks on the second issue, that is, the notion of triangulation.

2.2.3. Davidson's Criterion for Success in Communication

The Common View could provide us with a straightforward criterion for assessing Mrs. Malaprop's success or failure in her linguistic practice of meaning something by her words: she uses the word "epitaph" incorrectly, or in a mistaken way, since she failed to use the word in conformity to the relevant conventions governing the correct use of "epitaph". Davidson, however, rejected the essentiality of such conventional norms. If that is true, then "it may be urged that no alternative answer to Wittgenstein's query has been offered, the query being: what is the difference between using words correctly and merely thinking that one is using them correctly?" (Davidson, 1994, p. 116). Does Mrs. Malaprop merely think that she has used her words correctly?

Davidson agrees that he has to introduce a norm or a (public) criterion if the speaker's behaviour is to be assessable by others, and this norm must be introduced "without appeal to the test of common usage" (1994, p. 119). Davidson introduces his norm for evaluating the speaker's success in her linguistic behaviour as follows: "The intention to be taken to mean what one wants to be taken to mean" is what "is common to all verbal behaviour" and thus "constitutes a norm against which speakers and others can measure the success of verbal behavior" (1994, p. 120). We are already familiar with this criterion via our discussion of Davidson's alternative view and his account for linguistic error: success in the practice of meaning something by words is not a matter of following some rules; rather, it is a matter being understood by another person in a way the speaker intended her words to be understood. If the speaker is interpreted as she

intended to be, then we can say that she successfully meant something by her word, whether or not she follows, or fails to follow, certain conventions. In this regard, the Wittgensteinian distinction between what seems right to a speaker and what is actually right can be drawn without any appeal to fixed community-wide agreements about the use of words. In this Davidsonian picture of meaning and linguistic understanding, there is still a sense in which the speaker may fail to use a word correctly: the speaker fails to use her words correctly if she fails to use the words in an understandable way, that is, if she fails to lead the interpreter to grasp what she means by the words.

Therefore, Davidson's criterion takes the notion of successful communication as basic and makes the notion of meaning dependent on the notion of success in communication: without successful communication, there is no meaning. As Davidson emphasizes, "meaning … gets its life from those situations in which someone intends (or assumes or expects) that his words will be understood in a certain way, and they are" (1994, p. 120). How a speaker "intended to be understood, and was understood, is what he, and his words, literally meant on that occasion" (1994, p. 120). The relevant intentions of the speaker are those which are directly related to the speaker's "beliefs about how an audience will interpret his utterances" (1994, p. 122). It is indeed success in carrying out such intentions that is introduced as the norm for evaluating the speaker's linguistic behaviour. Therefore, Davidson's account seems to successfully accommodate the Wittgensteinian seems right/is right distinction by appealing to success in communication: it is not the case that whatever seems right to the speaker is right. If the speaker's interpreter fails to interpret the speaker's utterances in the way she intended them to be, then it just seemed to the speaker that she has been successful in meaning something by her utterance; she actually failed to mean anything by her words.[42]

[42] To this extent, hence, Davidson's alternative view is Wittgensteinian, as Davidson himself mentions: "There is, I believe, a direction in which to look for a solution, and that direction has been pointed out by Wittgenstein. What is needed is something that can provide a standard against which an individual can check his or her reactions, and only other individuals can do this" (2001d, p. 141). Davidson appealed to Wittgenstein's remarks to give his suggested norm or standard: "The answer Wittgenstein seems to offer … is: without an interpreter no substance can be given to the claim that the speaker has gone wrong—that he has failed to go on in the same way" (1992, p. 116). Davidson's suggested norm, as we saw, started with treating the success in mutual interpretation as fundamental: "Only communication with another can supply an objective check. … In communication, what a speaker and the speaker's interpreter must share is an understanding of what the speaker means by what he says" (1991a, pp. 209-210).

(a) An Initial Problem

However, this criterion for correctness seems to have a problem to deal with. Davidson was asked to provide an explanation of the fact that there is a distinction between one's actually meaning something by one's words and one's just thinking that one means something by the words. The criterion Davidson introduced was that the speaker may intend to mean something by her words, while she actually fails to do so. The speaker may fail to carry out her intention and, hence, it is not always the case that whatever the speaker intends to mean by her words is what her words actually mean. To be successful, the speaker must provide enough clues and evidence for the interpreter to reach the speaker's intended interpretation. Nonetheless, the problem is whether applying this criterion can really determine whether or not the speaker has succeeded, or failed, in her linguistic practices. The question is: if the speaker provides enough clues, but the interpreter fails to properly utilize them to reach the intended interpretation of the speaker's utterance, then should we still accuse the speaker for failing to mean something specific by her words? Suppose that Mrs. Malaprop intends to mean *epiphany* by "epitaph" when she utters "Newton had an intense epitaph". She thinks that there are enough clues for her hearer to understand what she meant by her words; for instance, she points to an apple falling from a tree in the garden. Nonetheless, the interpreter fails to understand what the speaker meant by her words. The question is whether this is a failure on the speaker's part or a failure on the interpreter's part. It seems that it is possible for the interpreter to interpret Mrs. Malaprop's utterance in the way she intended it to be understood. The same is true in the case of Mrs. Malaprop's using the word "epiphany" to mean *epiphany*. Even in this case, the interpreter may still fail to understand what Mrs. Malaprop means by her utterance.[43] In all these cases, although the interpreter failed to interpret the speaker as she intended, the speaker has not really made a mistake. Davidson's account, however, seems to imply that the speaker has made a mistake, since she has used her words in a way that the interpreter failed to understand.

In this regard, it seems that Davidson has to explain what the ability of the interpreter is supposed to be. Davidson did not give any clear explanation of this matter. Is the interpreter a normal human being with his normal skills and abilities to understand such

[43] For Davidson, it does not really matter whether the speaker intends to use her words in the standard way or in a non-standard one. What matters is the speaker's utterance being understood in the way the speaker intended it to be understood.

linguistic phenomena, such as malapropism? If that is the case, then it seems that we should count any failure in their communication as the result of the speaker's mistakes. The speaker alone is responsible for all the mistakes made during the conversation. But it does not seem to be a plausible account of error: intuitively, Mrs. Malaprop meant something understandable by her words, and it is the interpreter that failed to understand that meaning. It follows that Davidson's criterion does not take into account the errors or mistakes which are made by the interpreter, that is, the failures in communication that occur because of the interpreter's failure to interpret the speaker. This problem becomes more important when we note that Davidson's account seems to take *mutual* interpretation or *interpersonal* communication to be fundamental. Mutual or interpersonal communication is supposed to be a two-way conversation, in which both the speaker and the interpreter have their own role to play.[44]

One might reply to this objection by claiming that Davidson could take the interpreter to be immune to such mistakes so that we can simply assume that if the evidence and clues are indeed sufficient, then the interpreter would interpret the speaker as she intended her utterance to be interpreted. Hence, the interpreter's communication with Mrs. Malaprop would not fail if the evidence and clues are sufficient. This response, however, is not compatible with the way Davidson introduces his alterative view. In the case of the relation between a writer and a reader, the issue turns out to be more vivid: there is no speaker to be accused of using words in a non-understandable way. Rather, the author just assumes that her readers will have the ability to understand what she has written. The writer publishes the book and the readers are alone in their attempt to understand the book. Can we claim that, if a reader fails to understand what the author has written, then the author failed to mean anything by her words? Davidson (1991b) does not think so, especially regarding his appreciation of James Joyce's novels. Joyce's novels, or more particularly, many of his innovative sentences and phrases, may not be easily understandable for many readers. Davidson generally agrees that using language freely may bring about difficulties in understanding, just as the case of Humpty Dumpty. Because of that, according to Davidson, in using language it seems that the writer cannot "ignore what his readers know or assume about the words he uses, and such knowledge and expectations can come only from the reader's exposure to past usage" (1991b, p. 147). It is so because "in speaking or writing we intend to be understood. We

[44] For a different criticism from Dummett regarding the way Davidson treats communication between the speaker and the interpreter, see (Dummett, 1986, p. 462)

cannot intend what we know to be impossible; people can only understand words they are somehow prepared in advance to understand" (1991b, p. 147). Nonetheless, Davidson thinks that the writer can provide enough information to lead the audience to understand his writing by using "every resource his readers command (or that he hopes they command, or thinks they should command), every linguistic resource, knowledge of history, past writers, and styles" (1991b, p. 147). Joyce too, according to Davidson, relies on the knowledge of his readers and takes into account their ability to understand his written sentences. If that is true, then the readers' failure to understand Joyce's sentences seems to be a failure on their part, not Joyce's. Hence, it seems that Davidson does not think that the speaker (or the writer) is totally responsible for any mistake that may happen in her communication with her hearer or reader: "Mistakes on the part of speaker or interpreter are to be expected" (1998a, p. 89). If that is the case, then the speaker may successfully mean something specific by her utterance, while the interpreter fails to interpret the speaker's utterance correctly, in which case it is not clear how "being interpreted as intended" can indeed provide us with a criterion for correctness and accommodate the Wittgensteinian seems right/is right distinction.

Nonetheless, there is another, yet more serious, problem with Davidson's suggested criterion or norm. Davidson's criterion for correctness presupposes the notion of intention: meaning appears to be the result of the speaker's utterance's "being interpreted as intended". However, these two notions, i.e. the notion of meaning and that of intention, are interrelated: the notion of intention somehow presupposes the notion of meaning, since what the speaker means by her words depends on what she intends her words to mean. Especially for Davidson who thinks that meaning and propositional attitudes are deeply interdependent, this seems to be a serious problem: it makes his norm or criterion circular. If what the speaker means by her words depends on what she intends to mean by them, then the notion of meaning is already in play in this account. There is a meaning that the speaker intended to attach to her words. If Davidson really wants to provide a convincing criterion for success in the practice of meaning something by words, he cannot simply appeal to the notion of intention, that is, to the notion of meaning again. Davidson himself is aware of this problem: "I can easily understand why this observation can seem too true to be interesting, given that it assumes the notion of meaning" (1994, p. 120). However, Davidson thinks that, to this extent, his view and the Common View are in a similar position. What he was asked to do was to provide a norm distinguishing between a speaker's success and failure in her

practice of meaning something by words, that is, between using words correctly and using them incorrectly. His suggested norm did not explain how the speakers could come up with such meanings and intentions, how two people can acquire a language and propositional attitudes at all. His account does not explain why the existence of a second person interpreting the speaker is necessary for the speaker to draw the seems right/is right distinction. To this extent, however, Davidson's account is in the same position with the Common View: the Common View too does not explain these issues. Davidson mentions this point in his discussion of Dummett's adherence to the Common View[45]: "Michael avails himself of a notion of meaning [i.e. the Common View's notion of meaning] that he does not explain, while I avail myself of a concept of understanding that I don't explain" (1994, p. 123). Davidson thinks that Dummett too has not "given an argument to show that a shared way of speaking, a practice or convention, is essential to meaning something by what one says" (1994, p. 123). Although Davidson appreciates the fundamental idea behind the Wittgensteinian distinction between what seems right to a speaker and what is actually right, he thinks that a different account of such a distinction can be given, which does not require the speaker to go on as others do. As Davidson clarifies,

> [Dummett] has available an argument that purports to show that a shared practice is required in order to answer Wittgenstein's question, while I have only contended that a common practice isn't necessary for communication if each speaker goes on more or less as before. I have given no answer to the question what it is to go on as before. As a corollary, neither have I given any reason to think meaning is an essentially social phenomenon. … Agreeing with Dummett and Kripke, and perhaps with Wittgenstein, I hold that the answer to the question what it is to go on as before demands reference to social interaction. Where I disagree is on how this demand can be met. (1994, pp. 123-124)

What is it for a speaker to go on as before, based on Davidson's account? How can a speaker be said to fail to go on as before? Providing an answer to these questions is Davidson's main concern in his discussion of the notion of triangulation. Davidson's account did not require the speaker to speak as others do, or even speak as she did in the past. If that is the case, then his account of what it is to go on as before cannot be reliant on the requirement that the speaker has to speak as she did before. Indeed, it seems that what Davidson means by "going on as before" is the requirement that the speaker speak in an understandable way, that is, in a way that the interpreter can understand. As Davidson explains, "speaking a language … does not depend on two or more speakers

[45] Davidson discusses Dummett's view as presented in (Dummett, 1978, 1986).

speaking in the same way; it merely requires that each speaker intentionally make himself interpretable to the other (the speaker must 'go on' more or less as the other expects, or at least is equipped to interpret)" (1992, p. 115). The speaker can be said to be going on as before if she continues speaking in an understandable way, that is, if her interpreter can interpret her utterance successfully: it does not "matter whether the speaker speaks his 'native' tongue, since his past social situation is irrelevant. … The best the speaker can do is to be *interpretable*, that is, to use a finite supply of distinguishable sounds applied consistently to objects and situations he believes are apparent to his hearer" (1984a, p. 13).

Generally speaking, Davidson's remarks on the notion of triangulation aim at pinning down the idea that, although it is up to a speaker alone to decide what she intends to mean by her utterance, this practice is essentially social in that the actual presence of, and interaction with, a second person is necessary for the speaker to have any meaningful utterances and propositional attitudes with determinate content. In the following sections, we will investigate Davidson's use of the notion of triangulation to support his alternative account of meaning. In "Rational Animals" (1982), Davidson employs the notion of triangulation to clarify the necessary and sufficient conditions for propositional attitudes to emerge. In "The Social Aspect of Language" (1994), "The Second Person" (1992), "The Emergence of Thought" (1999c), "What Thought Requires" (2001d), and "Three Varieties of Knowledge" (1991a), his use of the notion of triangulation is more concerned with the necessary and sufficient conditions for a language to emerge. We start with Davidson's arguments offered in "Rational Animals" (hereafter RA) (1982), which aim to show that having any propositional attitudes essentially depends on two creatures' having a specific sort of interaction, i.e. linguistic triangulation. We then go to investigate Davidson's claim that having a language also depends on communicating propositional content.

2.2.4. Rational Animals: The Interdependence of Language and Thought

In RA, Davidson tries to argue for the claim that to be rational, a creature must be in linguistic interaction with at least another person and, in this sense, having any intention or belief essentially depends on having successful linguistic communication with others.

According to Davidson, a creature is rational if it has propositional attitudes, such as

beliefs, intentions, desires, hopes, wishes, fears, and the like: “By rationality I mean whatever involves propositional thought” (2001b, p. xiv). For Davidson, propositional attitudes have an intrinsically holistic character: having a belief requires having many other beliefs, together with many other basic attitudes such as intentions and desires, so that “the intrinsically holistic character of the propositional attitudes makes the distinction between having any and having none dramatic” (1982, p. 96). This holistic feature of propositional attitudes can provide us with a criterion for rationality: a creature can be said to be rational only if it has many interrelated propositional attitudes. If we agree with Davidson on the idea that propositional attitudes are extensively interrelated, then “obviously a rich pattern of beliefs, desires, and intentions suffices for rationality” (1982, p. 96). Davidson is going to argue that it is difficult to make sense of the situation in which a creature possesses only one propositional attitude, such as one belief or one desire, without having many other general and particular beliefs, desires, hopes and the like. He then goes on to argue that only creatures with a language can be said to have such a set of propositional attitudes.

According to Davidson, if the evidence which we use to justify our attributions of propositional attitudes to a creature is the creature’s behaviour, then such behaviour must be sufficiently complex that we can attribute many other propositional attitudes to the creature. In other words, only linguistic behaviour is capable of displaying the complexity that can justify our attribution of a rich set of propositional attitudes to the creature. Because of that, Davidson rejects the claim Norman Malcolm (1972-3) makes, according to which we can be justified in attributing a belief to a dog. Davidson introduces Malcolm’s example as follows:

> Suppose our dog is chasing the neighbor’s cat. The latter runs full tilt toward the oak tree, but suddenly swerves at the last moment and disappears up a nearby maple. The dog doesn’t see this maneuver and on arriving at the oak tree he rears up on his hind feet, paws at the trunk as if trying to scale it, and barks excitedly into the branches above. We who observe this whole episode from a window say, “He thinks that the cat went up that oak tree”. (Davidson, 1982, p. 97)

The problem with this claim is that, if such behavioural evidence can justify the attribution of that particular thought to the dog, then it seems that it must enable us to attribute many other relevant general and particular beliefs to it as well. Nonetheless, the dog’s behaviour is not sufficiently rich and complex to allow for such attributions:

> [H]ow about the dog’s supposed belief that the cat went up that oak tree? That

> oak tree, as it happens, is the oldest tree in sight. Does the dog think that the cat went up the oldest tree in sight? Or that the cat went up the same tree it went up the last time the dog chased it? It is hard to make sense of the questions. (Davidson, 1982, p. 97)

When we attribute a belief to a normal human being, we can easily make sense of attributing many other beliefs and propositional attitudes to that person, which are related to that particular attributed belief. For instance, if we attribute the particular belief that the cat went up that oak tree to someone, then we can expect that she has many other general beliefs about what cats are, what trees are, what animals are, what sort of general characteristics they have, what other things exist in the world, and the like, together with many particular beliefs about that black cat in the yard, that oak tree, that house, that day, and so on. As Davidson states, "there is no fixed list of things someone with the concept of a tree must believe, but without many general beliefs, there would be no reason to identify a belief as a belief about a tree, much less an oak tree" (1982, p. 98). For these reasons, Davidson concludes that "there may be no fixed list of beliefs on which any particular thought depends" (1982, p. 99).[46] Hence, the holistic feature of propositional attitudes makes them interdependent with each other so that a belief should be counted as the belief it is because of its relation to many other beliefs and propositional attitudes possessed by the creature.

However, although these remarks hint at the conclusion Davidson wishes to draw, that is, that having a rich set of propositional attitudes essentially requires the creature to have a very complex pattern of linguistic behaviour, they do not provide the sort of argument that Davidson has in mind for this conclusion. According to Davidson, "it is clear that a very complex pattern of behavior must be observed to justify the attribution of a single thought. And unless there is actually such a complex pattern of behavior, there is no thought. I think there is such a pattern only if the agent has language" (1982, p. 100). These remarks just show that the behaviour of the creature must be complex enough to allow us to attribute a rich set of propositional attitudes to that creature. But Davidson needs to argue for why this behaviour has to be linguistic. The conclusion that

[46] Davidson makes an even stronger claim about the interdependence of propositional attitudes: "Thoughts, like propositions, have logical relations. Since the identity of a thought cannot be divorced from its place in the logical network of other thoughts, it cannot be relocated in the network without becoming a different thought. Radical incoherence in belief is therefore impossible. To have a single propositional attitude is to have a largely correct logic, in the sense of having a pattern of beliefs that logically cohere" (1982, p. 99). What makes a belief the belief it is, is its location in the network of beliefs and propositional attitudes.

Davidson wants to argue for is "that a creature cannot have a thought unless it has language. In order to be a thinking, rational creature, the creature must be able to express many thoughts, and above all, be able to interpret the speech and thoughts of others" (1982, p. 100). Davidson's general argument for this claim can be put as follows:

(I) To have a belief, it is necessary to have the concept of belief.

(II) To have the concept of belief, the creature must have language.

(III) Therefore, to have a belief, the creature must have language.

Since the discussion of the notion of triangulation is brought into the picture in Davidson's argument for the second premise, we will concentrate on Davidson's argument for this premise, though we have to first briefly look at Davidson's idea about why the possession of a belief requires the possession of the concept of belief.

(a) The Argument from Surprise

Davidson begins his argument by considering the concept of surprise. He claims that to have a belief, it is necessary that the creature has the concept of a belief, that is, a belief about another belief. These beliefs can be called higher-level or second-order beliefs, or as Davidson calls them, "reflective" beliefs (1982, p. 103). As Davidson states, "my claim is … this: in order to have any propositional attitude at all, it is necessary to have the concept of a belief, to have a belief about some belief" (1982, p. 104). The general idea is that, if a creature can be said to have a belief about another belief, then we can conclude that the creature already possesses the concept of a belief. Without having such a concept, it is hard to make sense of the claim that the creature is aware of the fact that it has certain beliefs, that it knows that it has beliefs that can be true or false. Davidson thinks that, if a creature is capable of being surprised, then we can say that the creature has beliefs, since "surprise requires the concept of a belief" (1982, p. 104). We can characterize Davidson's argument from surprise as follows:

(1) To have a belief, the creature must be able to be surprised.

(2) To be able to be surprised, the creature must have the concept of a belief.

(3) Therefore, to have a belief, the creature must have the concept of a belief.

However, what is it to be surprised?

According to Davidson, a creature is surprised if it conceives some contrast in its beliefs. In other words, a creature is capable of being surprised if it can be aware of a contrast between what it did believe and what it comes to believe. When things do not turn out to be in the way the creature expected them to be, the creature must be able to understand that its previous beliefs or expectations were indeed wrong. This understanding or awareness is itself a belief, that is, a belief about another belief. Generally speaking, hence, if I am surprised, then I know that I had a belief that turns out to be false. Davidson gives an example to clarify this point:

> Suppose I believe there is a coin in my pocket. I empty my pocket and find no coin. I am surprised. … It is not enough that I first believe there is a coin in my pocket, and after emptying my pocket I no longer have this belief. Surprise requires that I be aware of a contrast between what I did believe and what I come to believe. Such awareness, however, is a belief about a belief: if I am surprised, then among other things I come to believe that my original belief was false. (1982, p. 104)

Davidson's reasons for both the premises (1) and (2) of his argument from surprise are supposed to be given in the above passage. However, Davidson's reason for the first premise is not really clear. Why is it the case that if a creature has a belief, then it must be able to be surprised? Davidson says, "it is not enough that I first believe there is a coin in my pocket, and after emptying my pocket I no longer have this belief" (1982, p. 104). This can be taken to be a reason for the claim that to have a belief, the creature must be capable of being surprised, since if I have a belief, and if beliefs, for Davidson, are states that can be true or false, then I must be able to realize whether my belief is true or false. In other words, if I have a belief, and if I form another belief which is in conflict with my previous belief, I must be able to grasp the contrast between these two beliefs. This means that I must be capable of being surprised. Regarding premise (2), however, Davidson's reason was clearer. Suppose that I believe that my friend will knock the door soon. The door is knocked and I see my brother, instead of my friend. According to Davidson, now I am surprised. I have found out that my previous belief was false, and this is a new belief about my previous belief: I now believe that my prior belief that my friend is knocking the door was false. Hence, if I am surprised, then I have a belief about another belief, that is, I have the concept of belief. Therefore, according to Davidson's argument from surprise, having a belief depends on having the

concept of a belief, i.e. to have a belief about another belief's being true or false, right or wrong.

This conclusion points to premise (II) of Davidson's main argument, that is, the claim that having the concept of a belief requires having a language. According to Davidson's argument from surprise, if a creature is to have a rich set of beliefs, it has to have the capacity of being surprised. If a creature has the concept of a belief, i.e. a belief about another belief, it is thereby capable of realizing that there can be a tension between its belief and what actually happens in the world, that there is a distinction between thinking that a belief is true and the belief actually being true. Beliefs are apt to be true or false: "Much of the point of the concept of belief is that it is the concept of a state of an organism which can be true or false, correct or incorrect" (Davidson, 1982, p. 104). If after opening the door I see my friend, then my belief that it was my friend knocking the door turns out to be true. But, if instead of my friend I see my brother, then I learn that my previous belief was false. In this regard, the claim that to have beliefs, the creature must be able to be surprised leads to the claim that the creature must have the concept of objective truth as well, that is, that things may turn out to be different from what the creature expected. As Davidson clarifies,

> to have the concept of belief is therefore to have the concept of objective truth. If I believe there is a coin in my pocket, I may be right or wrong; I'm right only if there is a coin in my pocket. If I am surprised to find there is no coin in my pocket, I come to believe that my former belief did not correspond with the state of my finances. I have the idea of an objective reality which is independent of my belief. (1982, p. 104)

Therefore, it seems that having the concept of a belief essentially involves having the concept of truth, objective truth, or objectivity. The requirement of having these concepts points to an important distinction that a creature with propositional attitudes must be aware of: the distinction between what seems right to a creature and what is actually right, that there is a world which exists independently of the creature and its beliefs, that there is a difference between what the creature thinks to be the case and what is actually the case. The creature must be aware of the fact that it is not always the case that whatever it thinks to be true is necessarily true.

Davidson's next target, hence, is to argue for premise (II) of his original argument, that is, the claim that to have any such concepts, including the concept of a belief, the concept of truth, the concept of a cat, and the like, the creature must have a language: it

must linguistically communicate with at least another creature. Why does having the concept of a belief, or the concept of truth, require having a language? Why is linguistic interaction with another person necessary for having such concepts?

2.2.5. Triangulation, Objectivity, and Cause-Determination

First of all, Davidson argues that a solitary speaker isolated from birth cannot be said to have the concept of truth, or to master the distinction between what seems right to her and what is actually right. Davidson wants to argue that "there may not be some other way of drawing the distinction … that does not depend on a social environment" (1992, p. 117).

According to Davidson, there is no opportunity for a solitary person to grasp the seems right/is right distinction, since, for her, there would be no objective criterion for error. To have a better view of Davidson's reasons for this claim, we need to consider his externalism about the content of our thoughts and utterances. According to Davidson's externalism, "what a person's words mean depends in the most basic cases on the kinds of objects and events that have caused the person to hold the words to be applicable" (1987, p. 37). The meaning of our utterances and the content of our thoughts are partly determined by their typical external causes in the world. If that is true, then if we want to determine the content of our thoughts and the meaning of our utterances, then we need to determine the actual causes of those utterances and thoughts. For instance, if I utter that "That apple is green", or if I have the belief that that apple is green, then the meaning of my utterance or the content of my belief is at least partly determined by the specific object (i.e. that apple in view) which caused me to utter that sentence, or to form that belief. My friend, hence, will be able to understand my utterance if, first of all, he can determine the actual cause of my utterance, that is, if he understands that my utterance is about that specific apple in view.[47] Therefore, according to Davidson's externalism about content, "the stimuli that cause our most basic verbal responses also determine what those verbal responses mean, and the content of the beliefs that accompany them" (1991a, p. 213). Davidson's externalism

[47] As Davidson clarifies, "in the simplest and most basic cases, words and sentences derive their meaning from the objects and circumstances in whose presence they were learned. A sentence which one has been conditioned by the learning process to be caused to hold true by the presence of fires will (usually) be true when there is a fire present" (1988, pp. 44-45).

implies that there is a causal relation between our thoughts and the goings-on in the world so that "concepts, and the sentences and thoughts that employ them, are in part individuated by their causal relations to the world" (2001d, p. 138).[48] However, the question is whether a solitary person can determine the actual cause of her utterances and thoughts, and whether she can master the objective distinction between what she thinks to be the case and what is actually the case.

Davidson's answer is negative to the above questions. According to Davidson, there are two general problems regarding the practices of a solitary person: the solitary speaker would not be able to acquire the concept of objectivity and to determine the actual cause of her utterances and thoughts. These problems are the problems the solitary person faces regarding two important aspects of her thoughts, that is, "the objectivity of thought and the empirical content of thoughts about the external world" (Davidson, 1999c, p. 129). For a creature to have objective thoughts, it must be aware of the distinction between believing that something is the case and that thing being the case independently of what the creature believes, that is, it must have the concept of objectivity. Davidson will argue that a solitaire cannot command such a distinction. Moreover, Davidson thinks that if the content of a thought is partly determined by its external cause, then for thoughts to have any determinate content, their external causes must be determined; but, for the solitary speaker, the "cause is doubly indeterminate: with respect to width, and with respect to distance" (Davidson, 1999c, p. 129). For instance, suppose that the solitary person sees a cow in view and utters "Cow". Tomorrow, she sees the cow again and utters "Cow". Davidson thinks that, if this person is a solitary person isolated from birth, she would not be able to determine whether it is an external object a certain distance from her or just some proximal stimulus on her eyes' retina that causes her to utter that word. This is the problem with respect to distance, that is, the problem with determining whether the actual cause is a distal cause in the world or a proximal one on the surface of the speaker's skin. The next problem, i.e. the problem with respect to width, is the problem of determining which aspect of the object actually caused the speaker to produce that response. As Davidson says, this ambiguity concerns "how much of the total cause of a belief is relevant to content" (1999c, pp. 129-130). For instance, suppose that the speaker can somehow determine

[48] For more remarks from Davidson on his externalism, see (Davidson, 1997a, 1995a, 1991c, 1987). For more discussion, see (Ludwig & Lepore, 2005, section 19), (Glüer, 2011, 2006), and (Verheggen, 2007, 2006, 1997).

that the cause of her utterance is an external object, e.g. the cow in view, not some proximal stimulus on the surface of her skin. Now, is it the cow as a whole that caused the speaker to utter “Cow”, or part of the cow, or its colour, or the cow and the environment, etc.? If the actual cause of the speaker’s thoughts and utterances is to be determined, then both problems must be solved: the speaker’s response must be caused by a specific aspect of a certain external cause in the world.

Davidson by appealing to Wittgenstein’s considerations on private language argues that, for a solitary person, there would be no determinate and objective cause at all, since Wittgenstein has shown us that “without an interpreter no substance can be given to the claim that the speaker has gone wrong – that he has failed to go on in the same way” (Davidson, 1992, p. 116). For the solitaire, there is no opportunity to draw the distinction “between thinking one means something and actually meaning it” (Davidson, 1992, p. 117). The reason is that whatever seems right to the solitary speaker is right. If such a speaker is to have the concept of a belief, and thereby any belief at all, she has to grasp the distinction between what just appears to be the case for her and what is actually the case. For the solitary speaker, however, it is entirely up to her to decide whether something is the cause of her thoughts, and hence whatever she takes to be the actual cause of her utterance is the actual cause of her utterance: if she responds to something by “Cow” today because of thinking that this is the same thing to which she responded by “Cow” yesterday, then there is no objective criterion for her to evaluate the correctness of her response. There is no objective ground for her to assess her thinking that this is the same cow as she saw yesterday; perhaps, there is even no cow in the vicinity. Hence, she cannot draw the desired seems right/is right distinction, and hence she cannot grasp the concept of objectivity or objective truth: “As Wittgenstein says, by yourself you can’t tell the difference between the situations seeming the same and being the same” (Davidson, 1994, p. 124). In the same way, the cause of the solitary speaker’s utterances and thoughts would remain entirely indeterminate, since what seems to be the actual cause of her response may not be the actual cause of her utterance. There is no opportunity for her to objectively determine the actual, external cause of her responses. As a result, she fails to have any belief or propositional attitudes whatsoever, since the possession of any propositional attitude depends on the possession of the concept of objective truth. And, the solitary speaker cannot possess such a concept. For Davidson, the same is true in the case of a child at the earliest stages of learning his first language: “In the beginning the child lacks

awareness of the distinction between what is believed and what is the case, what is asked or demanded and what is answered or done" (1998a, p. 86). However, what does it take for a creature, e.g. for the child, to acquire such a concept, i.e. to master the Wittgensteinian seems right/is right distinction? The answer, according to Davidson, is triangulation:

> There is a prelinguistic, precognitive situation which seems to me to constitute a necessary condition for thought and language … . Both in the case of nonhuman animals and in the case of small children, it is a condition that can be observed to obtain. The basic situation is one that involves two or more creatures simultaneously in interaction with each other and with the world they share; it is what I call *triangulation*. (1999c, p. 128)

Davidson's idea is that if the solitary speaker can interact with another person, she will be in a position to grasp the concept of objectivity and to determine the actual cause of her responses. Interacting with another person and simultaneously with a certain object in the world, to which both are responding, is what Davidson calls "triangulating". It is a situation in which two creatures respond to a certain stimulus in the world and, at the same time, to each other's responses. To understand the role of the notion of triangulation in Davidson's argument, it is important to distinguish between two sorts of triangulation: "primitive" or "prelinguistic" triangulation and "linguistic" triangulation.

(a) Primitive Triangulation

Davidson thinks that the necessary condition for a creature to make sense of the concept of objectivity is to engage in the primitive triangulation. This kind of triangulation is the one in which two creatures similarly respond to the same stimulus in the world as well as to each other's responses to that stimulus. As Davidson states, such a triangulation "is the result of a threefold interaction, an interaction which is twofold from the point of view of each of the two agents: each is interacting simultaneously with the world and with the other agent" (1999c, p. 128). There are different examples of this situation. When two non-linguistic animals interact with each other to hunt another animal is an example of such prelinguistic triangulation, e.g. when two lionesses cooperate to hunt a gazelle. For example, consider this description of cooperative hunting in lions:

> Group hunts generally involved a formation whereby some lionesses ("wings") circled prey while others ("centres") waited for prey to move towards them. Those lionesses that occupied "wing" stalking roles frequently initiated an attack

> on the prey, while lionesses in "centre" roles moved relatively small distances and most often captured prey in flight from other lionesses. Each lioness in a given pride repeatedly occupied the same position in a hunting formation. Hunts where most lionesses present occupied their preferred positions had a high probability of success. (Stander, 1992, p. 445)

This description of group hunting illuminates the sort of complex interaction involved in primitive triangulation between non-linguistic animals. Also, when a child starts learning his first language, he is in such a primitive triangulation, since the child has no language: "For the learner ostension is not learning something already there. The learner is in at a meaning baptism" (Davidson, 1997b, p. 140). According to Davidson, engaging in this sort of triangulation is necessary to acquire the concept of objectivity and to determine the actual cause of responses, since, generally speaking, it creates a space for the triangulators to realize disagreement between their responses.

Obviously, such a triangulation can be established between the creatures that are similar in their grouping things together and in their responses to the world: "It is clear that for triangulation to work, the creatures involved must be very much alike. They must class together the same distal stimuli, among them each other's reactions to those stimuli" (Davidson, 2001d, p. 143). If each of us sorts things in a radically different way, then no real similarity or correlation between our responses to different items in the world would be observable. In this regard, primitive triangulation is not a situation in which only human beings can engage; rather, it can take place via the interaction of any two or more similar animals.[49] In these situations, the process of conditioning and the phenomenon of generalization have their roles to play. A child is conditioned by his parents or teachers to react to certain things in the world in certain ways, for instance, to respond to certain things by "table" and to certain other things by "chair". Also, the child has the capacity to generalize these primitive cases of learning to further cases. For example, as Davidson states, "the dog hears a bell and is fed; presently it salivates when it hears the bell. The child babbles, and when it produces a sound like 'table' in the evident presence of a table, it is rewarded; the process is repeated and presently the child says 'table' in the presence of tables" (1992, p. 117).The child and the dog can generalize what they have learnt: "The phenomenon of generalization, of perceived similarity, plays an essential role in the process. One ring of the bell is enough like

[49] As Davidson points out, "one sees this in its simplest form in a school of fish, where each fish reacts almost instantaneously to the motions of the others. This is apparently a reaction that is wired in. A learned reaction can be observed in certain monkeys which make three distinguishable sounds depending on whether they see a snake, an eagle, or a lion approaching" (1999c, p. 128).

another to provoke similar responses in the dog, just as one presentation of food is enough like another to engender salivation" (1992, p. 117).

However, putting these details to one side, what is crucial in this picture is the triangular situation in which the creatures find the opportunity to correlate their similar responses to a specific stimulus and perceive the *failure* of such correlations. This situation helps them to tackle two problems: it helps the triangulators to determine the actual cause of their responses and, at the same time, to make sense of the concept of objectivity. As Davidson describes such a situation:

> Involved in our picture there are now … three similarity patterns. The child finds tables similar; we find tables similar; and we find the child's responses in the presence of tables similar. … Given these three patterns of response we can assign a location to the stimuli that elicit the child's responses. The relevant stimuli are the objects or events we naturally find similar (tables) which are correlated with responses of the child we find similar. (1992, p. 119)

The child responds to a certain thing in the world by uttering "table", his parents respond by uttering "table" to the same thing, and they all perceive the similarity between their responses to that stimulus. The actual distal cause of the child's response is located in the intersection of the lines from their responses to that specific stimulus:

> It is a form of triangulation: one line goes from the child in the direction of the table, one line goes from us in the direction of the table, and the third line goes between us and the child. Where the lines from child to table and us to table converge, 'the' stimulus is located. Given our view of child and world, we can pick out 'the' cause of the child's responses. It is the common cause of our response and the child's response. (1992, p. 119)

The actual cause of the child's response, hence, is the "common cause" prompting the similar responses of the child and his parents. Unless a creature interacts with at least another creature, there would be no answer to the question to which specific cause it is responding: "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 than, say, on its skin" (Davidson, 1992, p.119). Engaging in primitive triangulation is thus necessary for a creature to determine the actual, distal cause of its utterances and thoughts.[50] And if it is the external cause that at least partly determines

[50] This was, however, the distance problem in cause-determination. Does primitive triangulation help to solve the aspect problem too? It does not seem so. Although primitive triangulation may help to solve the distance problem, that is, the problem of determining whether the cause is distal or proximal, it is not helpful in dealing with the problem of determining which aspect of the cause determines the content of the creature's responses: mere interaction between two creatures, that is, the correlation between their

the content of the creature's thoughts and the meaning of its utterances, then, without such triangulating, the creature would not have any thought and language at all.

At the same time, being in primitive triangulation is necessary to make sense of the concept of objectivity, or to conceive the Wittgensteinian seems right/is right distinction, because it is only through engaging in such a triangulation that the creature can find out what it is to make a mistake, what it is to be in a genuine disagreement with others. Suppose that our speaker utters "Cow" when she sees a cow in view. Her friend also utters "Cow" on that occasion. In this situation, they have correlated their similar responses to what they think to be the cause of their responses. Tomorrow she sees the same animal and utters "Cow". This time, however, her friend utters "Sheep". The correlation between their previously similar responses is now broken: something went wrong with their responses.[51] They now confront a disagreement between their responses which needs to be settled. The difference between this triangular situation and the situation of the solitary speaker is that, in the case of the solitaire, it was totally up to her to decide whether or not the cause of her present response is the same cause which prompted her previous response yesterday; in the case of the two triangulators, however, it is not up to one of them alone to make such a decision. As Davidson mentions, "once these correlations are set up, each creature is in a position to expect the external phenomenon when it perceives the associated reaction of the other. What introduces the

---

similar responses to a certain object, does nothing to determine which *aspect* of the cause prompts their similar responses. The reason is that, no matter how similar the responses of the two creatures are and how many times such similar responses have been repeated, there are always different aspects of the object that can prompt the similar responses in the creatures, and there is nothing by appealing to which we can show that the creatures respond to the *same* aspect of the object. It is not the case that if a creature always responds with "table" in the presence of a whole table, then it takes the whole table to be the actual cause of its response; rather, it may be responding to an undetached table part, for instance (recall Quine's "Gavagai" example). Hence, while one creature is responding to an undetached table part and the other is responding to the table as a whole, they both have similar responses, i.e. they both utter "table". However, can linguistic triangulation help? We will discuss Davidson's remarks on the notion of linguistic triangulation in the next section, but, even without knowing the details about linguistic triangulation, we can roughly guess that the creatures' responses becoming linguistic would not *by itself* help to solve this problem. If, according to Davidson, it is the *similarity* of responses that is supposed to determine to what (aspect of an) object the creatures are responding, it does not really matter whether the responses are linguistic or otherwise. It would matter, however, if we make an assumption about them in advance, i.e. that linguistic responses are already meaningful, i.e. that the actual cause of the responses is fixed. But, doesn't making such an assumption presuppose what was supposed to be explained in the beginning? We will discuss these problems further in Chapter Four.

[51] As Davidson says, "If you and I can each correlate the other's responses with the occurrence of a shared stimulus, however, an entirely new element is introduced. Once the correlation is established it provides each of us with a ground for distinguishing the cases in which it fails" (1994, p. 124).

possibility of error is the occasional failure of the expectation; the reactions do not correlate" (1999c, p. 129). In this way, this primitive triangulation helps the creatures to make sense of error, the seems right/is right distinction, or the concept of objectivity.

Nonetheless, while this kind of triangulation, i.e. the primitive triangulation, is necessary, it is not sufficient for the creatures to be said to have actually acquired the concept of objectivity, or that of a cow, a table, and the like: "The kind of triangulation I have described, while not sufficient to establish that a creature has a concept of a particular object or kind of object, is necessary if there is to be any answer at all to the question what its concepts are concepts of" (Davidson, 1992, p. 119). The reason is that primitive triangulation can be established between animals with no language and propositional attitudes at all, such as fishes, lions, and the like:

> The triangular relationship between agents and an environment to which they mutually react is, I have argued, necessary to thought. It is not sufficient, as is shown by the fact that it can exist in animals we do not credit with judgement. For this reason we are in a position to say something about a situation that must exist if thought does, but it is a situation that can exist independently. (1999c, p. 130)

As non-linguistic animals can establish this sort of primitive triangulation, the mere correlation between the responses of the two triangulators, or complex responses of a solitary person, cannot be seen as sufficient to bestow on them a language or thoughts with determinate content.[52] However, Davidson's point is that, without engaging in such triangulation, there would be no answer to the question to what objects or events a creature is responding. And, as Davidson emphasizes, without an answer to this question, "there is no answer to the question what language a creature speaks, since to designate a language as one being spoken requires that utterances be matched up with objects and events in the world (and not, in general, events on the surface of the skin)" (1992, pp. 119-120).[53]

---

[52] As Davidson says, "a creature may interact with the world in complex ways without entertaining any propositions. It may discriminate among colors, tastes, sounds, and shapes. … Yet none of this, no matter how successful by my standards, shows that the creature commands the contrast between what is believed and what is the case, as required by belief" (1982, pp. 104-105).

[53] Thus, it is important to note that what Davidson is after in his discussion of the notion of primitive triangulation is just to clarify the "necessary" condition, which must be satisfied if the creature is to be said to have the opportunity to possess the concept of objectivity and have responses with determinate content: primitive triangulation "does no more than indicate the sort of conditions in which the idea of error could arise. Thus, it suggests necessary (though certainly not sufficient) conditions for conceptualization" (2001d, p. 142).

What else must be added to primitive triangulation to allow the triangulators to acquire the concept of truth or objectivity and to have thoughts with determinate content?

(b) Linguistic Triangulation

According to Davidson, "the answer is language" (1999c, p. 130). To acquire and apply the concept of objectivity and to solve the cause-determination problem, "linguistic communication suffices" (1982, p. 105). The reason is that linguistic interaction, that is, speaking with others and understanding the speech of others, is possible only if each triangulator knows to which stimulus the other is responding. Otherwise, no linguistic communication between them can take place. Each of them must know what the other has in mind when she utters certain words: "To understand the speech of another, I must be able to think of the same things she does; I must share her world. I don't have to agree with her in all matters, but in order to disagree we must entertain the same propositions, with the same subject matter, and the same concept of truth" (1982, p. 105). If two creatures have linguistic communication with each other, then it is obvious that they have the concept of objective truth and thoughts with determinate content. In this sense, it is clear that linguistic triangulation is *sufficient* for the creatures to be said to have the concept of objectivity. However, this claim by itself is not helpful, since it is circular and indeed presupposes what was required to be explained. Davidson is aware of this problem: "This is not much help, since it is obvious that a creature that has a language can think; language is an instrument for the expression of propositional contents" (1999c, p. 130).[54] A more important question, however, is: Why can't there be any other way for the creatures to acquire the concept of objectivity and fix the actual cause of their responses? Why is linguistic triangulation the only option? In other words, while Davidson thought that linguistic triangulation is *sufficient* for the creatures involving in it to be said to have acquired the concept of objectivity and fixed the cause of their responses, he still needs to argue why linguistic triangulation is *necessary*. Davidson himself confesses the difficulty in answering to this question:

[54] The charge of circularity is a common objection to Davidson's discussion of triangulation. We will not have the space to discuss this objection here as there is a wealth of secondary literature on this issue and disagreement on the importance of this challenge. For discussion of this problem, see e.g. (Ludwig & Lepore, 2005, Section 22), (Verheggen, 2006, 2007, 2016), (Lasonen & Marvan, 2004), and (Glüer, 2006).

> To complete the 'argument', however, I need to show that the *only* way one could come to have the belief-truth contrast is through having the concept of intersubjective truth. I confess I do not know how to show this. But neither do I have any idea how else one could arrive at the concept of an objective truth. (1982, p. 105)

Davidson, however, tries to argue that to acquire and apply the concept of objectivity, together with any other concepts, such as that of a table, a tree, and the like, the two creatures must know of each other that they are responding to specific things in the world, and to do so they must know what object the other has in mind, that is, they must know which concept the other is applying or expressing by her responses. Knowing such complex things, however, is possible *only* through linguistic triangulation, that is, through interpreting each other and being interpreted by each other.

First of all, having any concept to apply to the world essentially depends on the creature's possession of the concept of objectivity. The reason is that to have a concept is to sort things in the world in a specific way, "to class things under it" (Davidson, 2001d, p. 137). To say that creatures classify objects and aspects of the world in certain ways is to say that "they treat some stimuli as more alike than others" (Davidson, 1991a, p. 212). In this sense, to have the concept of a table is to treat tables as similar things in the world, to sort them under a similar group of things. However, for this classification, or conceptualization, to be possible, the creature needs to have a conception of what it is to make a mistake, since, for instance, putting certain things under the class of "tables" means to put other things, which are not tables, under a different category, *not* under the category of tables. As Davidson states, "to apply a concept is to make a judgment, to classify or characterize an object or event or situation in a certain way, and this requires application of the concept of *truth*, since it is always possible to classify or characterize something wrongly" (1995b, p. 9). If that is true, then "a creature that has a concept knows that the concept applies to things independently of what it believes. A creature that cannot entertain the thought that it may be wrong has no concepts, no thoughts" (Davidson, 2001d, p. 141). Hence, if a creature is to have any language and thought at all, it must acquire the concept of objectivity, that is, "the creature itself must be able to recognize error" (Davidson, 2001d, p. 141). It is not enough that the creature merely discriminates between different objects or different aspects of the world, since non-linguistic animals can do this. Rather, "having a belief demands in addition appreciating the contrast between true belief and false, between appearance and reality, mere seeming and being. … Someone

who has a belief about the world … must grasp the concept of objective truth, of what is the case independent of what he or she thinks" (Davidson, 1991a, p. 209).[55]

However, when can a creature understand what concept the other creature is applying to the world? When can a creature know to what specific object in the world the other is responding? Davidson believes that if each triangulator is to know of the other what concepts it has in mind when it produces specific responses, they must linguistically triangulate with each other. As Davidson states,

> the only way of knowing that the second apex of the triangle—the second creature or person—is reacting to the same object as oneself is to know that the other person has the same object in mind. … For two people to know of each other that they are so related, that their thoughts are so related, requires that they be in communication. Each of them must speak to the other and be understood by the other. They don't, as I said, have to mean the same thing by the same words, but they must each be an *interpreter* of the other. (1992, pp. 120-121, emphasis added)

In this way, Davidson concludes that linguistic triangulation is necessary for a creature to have thoughts with determinate content and utterances with determinate meaning: it is the only way in which the creatures can acquire and apply the concept of objectivity and fix the content of their thoughts and utterances.

> [U]nless the base line of the triangle, the line between the two agents, is strengthened to the point where it can implement the communication of propositional contents, there is no way the agents can make use of the triangular situation to form judgements about the world. Only when language is in place can creatures appreciate the concept of objective truth. (Davidson, 1999c, p. 130)

Indeed, unless the creatures linguistically respond to each other, there would be no answer to the question whether they are responding to the same thing in the world, that there is a world which exists independently of their beliefs. Therefore, the cause-determination problems and the problem of the objectivity of thought are solved only through engaging in linguistic triangulation. Linguistic triangulation can be said to be the most basic situation of communication, or mutual interpretation: each triangulator must be the interpreter of the other. To have a language, hence, creatures must

[55] As Davidson states, "I should therefore like to reserve the word 'concept' for cases where it makes clear sense to speak of a mistake, … as seen from the creature's point of view. … It may seem that one could have the concept of, say, a tree, without being able to think that, or wonder whether, something is a tree, or desire that there be a tree. Such conceptualization would, however, amount to no more than being able to discriminate trees—to act in some specific way in the presence of trees—and this, as I said, is not what I would call having a concept" (1995b, pp. 8-9)

communicate the propositional content of their thoughts, and to be able to communicate the propositional content of their thoughts, that is, to know of each other to what specific thing the other is responding, they must linguistically triangulate. This means that having any belief and intention requires a social interaction, the most basic example of which is linguistic triangulation.[56] In this way, Davidson completes his original argument for the conclusion that to have any propositional attitudes, the creature must have a language, since, to have the concept of a belief and, similarly, to have the concept of truth, objectivity, or objective truth, it is both necessary and sufficient to be in linguistic triangulation. As Davidson himself concludes, "the conclusion of these considerations is that rationality is a social trait. Only communicators have it" (1982, p. 105).

We now can come back to Davidson's suggested norm for assessing the verbal behaviour of a speaker and conclude this section: there is a difference between the speaker's thinking that she means something by her words and the speaker's actually meaning that thing *only if* she can be said to have been in actual communication with another person and acquired the concept of objective truth. The speaker's communication with the second person, or the interpreter, is successful only if the interpreter can understand what the speaker means by her words, that is, only if the interpreter understands to what objects in the word the speaker is applying her words. The speaker and the interpreter need not mean the same thing by the same word, or follow the same rule or convention; rather, they must be able to understand each other. It is success in this basic social interaction that is fundamental, not speaking as others do, or conforming to previously shared conventions. [57] In this way, his discussion of the notion of triangulation leads us back to his original idea that it is idiolect, i.e. the way an individual speaker (intends to) use her words, that is philosophically fundamental, not the existence of a common "language" as a rigid set of rules. Nonetheless, the individual

[56] As Davidson says, "what gives each the concept of the way things are objectively is the base line formed between the creatures by language. The fact that they share a concept of truth alone makes sense of the claim that they have beliefs, that they are able to assign objects a place in the public world" (1982, p. 105).

[57] Davidson summarizes his remarks as follows: "Those who insist that shared practices are essential to mean are half right: there must be an interacting group for meaning – even propositional thought … – to emerge. Interaction of the needed sort demands that each individual perceives others as reacting to the shared environment much as he does; … It follows that meaning something requires that by and large one follows a practice of one's own, a practice that can be understood by others. But there is no fundamental reason why practices must be shared (1994, p. 125).

speaker's linguistic practices are argued to be essentially social: "Language is necessarily social. … to have thoughts, and so to mean anything in speaking, it is necessary to understand, and be understood by, a second person" (Davidson, 2001b, p. xv). The best conclusion for our discussion is given by Davidson himself summarizing his view as follows:

> If Wittgenstein held that language is necessarily social, then [my] central thesis … is Wittgensteinian. But it is denied that communication requires that one person speak as others do. Rather, the objectivity which thought and language demand depends on the mutual and simultaneous responses of two or more creatures to common distal stimuli and to one another's responses. This three-way relation among two speakers and a common world I call 'triangulation'. (2001b, p. xv)

We will come back to, and critically discuss, Davidson's argument from triangulation in Chapter Four, where we discuss his reading of Kripke's Wittgenstein. The next chapter starts with an investigation of Davidson's grasp of Quine's indeterminacy of translation arguments and an argument to the effect that Davidson's grasp of the significance of Quine's argument appears to be inadequate. It is very important to investigate Davidson's reading of Quine on this matter, since we are going to argue that the way Davidson treats KW's sceptic's argument is similar to his treatment of Quine.

2.2.6. Conclusion

In this chapter, we introduced Davidson's negative argument against the Common View and then his alternative view of meaning and linguistic understanding. Davidson's criticisms started by introducing three versions of the Common View, each of which tried to find a fundamental place for conventions in our linguistic practices. The most important version of the Common View implied the idea that success in communication essentially depends on using words in accordance with some specific convention. Davidson's chief claim was that following such conventions is neither necessary nor sufficient to succeed in communication. He used the phenomenon of malapropism and our smooth understanding of Mrs. Malaprop's utterances to argue that, if the Common View is true, then Mrs. Malaprop's linguistic usage makes no sense: she means nothing by her words.

Davidson's alternative view, however, suggests that success in mutual interpretation is both necessary and sufficient for explaining success in the practice of meaning

something by words. If the speaker is understood in the way she intended to be understood, then it does not matter whether she uses her words in agreement with the way her speech community does. Davidson's account could provide a norm to evaluate the speaker's success in her linguistic practice of meaning something by her words, which could accommodate the Wittgensteinian distinction between what seems right to the speaker and what is actually right. The norm or the criterion was the requirement of the speaker's utterances' "being interpreted as intended". Davidson, however, tried to explain how the existence of such a norm essentially depends on the existence of a certain sort of social setting, that is, triangulating with at least another person. Primitive triangulation, that is, the correlation between two creatures' similar responses to a certain stimulus, was argued by Davidson to be necessary for thoughts and language to emerge, since, for a solitary creature, the actual cause of its response would remain indeterminate: whatever cause the creature takes to be the actual cause of its response would be the actual cause of its response. However, primitive triangulation was not enough for the creatures to fully determine the actual cause of their responses and to acquire the concept of objectivity. Davidson argued that it is linguistic triangulation that is both necessary and sufficient for the creatures to fix the cause of their responses and acquire the concept of objectivity, which involves the process of interpreting others and being interpreted by others.

3. Chapter 3: Davidson on Quine's Indeterminacy of Translation Thesis [58]

Introduction

In this chapter, I will introduce and evaluate Davidson's reading of Quine's indeterminacy of translation thesis. The reason for this investigation is that there is, as we will see in Chapter Four, a very important similarity between Davidson's grasp of KW's sceptic's sceptical argument and his interpretation of Quine's indeterminacy of translation thesis. Therefore, our discussion in this chapter will help our evaluation of Davidson's response to KW's sceptical argument and solution.

Davidson has always been explicit in his faithful adherence to the main doctrines of Quine's philosophy of language, among which the indeterminacy of translation thesis is significant. For Quine, the indeterminacy of translation thesis has important ontological consequences, construed as leading to a sceptical conclusion regarding the existence of fine-grained meaning facts.[59] Davidson's suggested reading of Quine's arguments for the indeterminacy of translation thesis seems to be intended to block any such ontological/sceptical consequences. According to this reading, Quine's arguments at most amount to the conclusion that there are different ways of representing the facts about meaning, rather than the sceptical conclusion that there are no such facts. In this regard, we will start by introducing Quine's arguments for the indeterminacy of translation. After introducing Davidson's interpretation of Quine's arguments, we will go on to discuss whether Davidson's reading of Quine is plausible. I will argue that the way Davidson treats Quine's arguments is puzzling and leaves him with a difficult dilemma: Davidson must either give up on Quine's main doctrines on radical translation by accepting non-reductionism about meaning or remain faithful to Quine's project by giving up his reading of Quine's indeterminacy of translation thesis. Once the problematic character of Davidson's response to Quine's sceptical arguments is

[58] Readers unfamiliar with the early Davidson's account of meaning may wish to read the Appendix before proceeding with this chapter and the next.

[59] Fine-grained meaning is a notion on which, for instance, "rabbit" and "undetached rabbit part" have different meanings. There is, however, another notion of meaning, i.e. stimulus meaning, on which they have the same meaning. I will introduce the notion of stimulus meaning in Section 3.1.2 below (pp 119-120).

clarified, it will also shed light on the general approach Davidson takes towards KW's sceptical argument.

3.1. Quine's Arguments for the Indeterminacy of Translation

Quine famously argued that translation is indeterminate and thereby cast doubt on the existence of the sort of determinate meaning that we intuitively expect each sentence to possess. There are two arguments proposed by Quine for the indeterminacy of translation, "the argument from below" and "the argument from above" (see Quine, 1970, p. 183). We will introduce the two arguments after a brief review of Quine's physicalism, which plays a vital role in the arguments.

3.1.1. Quine's Physicalism

The indeterminacy of translation thesis relies foremost on Quine's physicalism/behaviourism: he says that "the thesis is a consequence of my behavioursim" (1987, p. 5).[60] Putting to one side the controversies about how Quine's physicalism can best be characterized, the following passage from Quine can be taken to roughly capture the main theme of his physicalism: "Nothing happens in the world, not the flutter of an eyelid, not the flicker of a thought, without some redistribution of microphysical states" (1981, p. 98). In other words, nothing changes in the world, unless there are relevant physical changes. Putting it in a different terminology, Quine's physicalism is the view that fixing the relevant physical facts fixes all other facts of the matter. [61] For example, consider the case of temperature. Quine's physicalism implies that if there are facts about temperature, they will be fixed once the relevant physical

[60] Indeed, Quine's behaviourism can be taken to be a manifestation of his physicalism in the case of radical translation, where the physical facts are the facts about speakers' (verbal or otherwise) behaviour.

[61] According to Hookway, Quine's physicalism can be captured in the claim that ''the physical facts are all the facts'' (Hookway 1988, p. 212). In a different interpretation of Quine's physicalism, Soames (1999) characterizes it as follows: "It states that all genuine truths (facts) supervene on the physical truths (facts)" (1999, p. 333). Also, according to Peter Hylton (2007), "Quine's physicalism is: no difference without a physical difference" (2007, p. 316). The way we have characterized this view is closer to the way Scott Soames (2003) does, according to which "all genuine truths (facts) are determined by physical truths (facts)" (2003, p. 247). For more discussion, see e.g. (Miller, 2006), (Kirk, 1986), (Glock, 2003), and (Kemp, 2006, 2012).

facts are fixed, which are, in this case, facts about molecules' kinetic energy at a certain time.

3.1.2. The Argument from Below

Quine's "argument from below" starts with investigating the process of translating words and terms. The argument invites us to imagine a translator trying to translate a radically unknown language spoken by a native. Since the language is entirely unknown, the relevant data or evidence on which our radical translator can rely for his translation is very limited, that is, nothing more than the verbal behaviour of the native, (taking into account Quine's physicalistic view). The particular sort of evidence Quine allows is the translator's observing the native's verbal dispositions to assent to or dissent from a sentence on specific occasions. Hence, based on Quine's physicalism/behaviorusim, "there is nothing in linguistic meaning, then, beyond what is to be gleaned from overt behaviour in observable circumstances" (Quine, 1987, p. 5). Consequently, the sentences the translator initially deals with are the sentences "that do hinge pretty strictly on the concurrent publicly observable situations: sentences like 'It's raining' or 'That's a rabbit', which I call *observation sentences*" (Quine, 1987, p. 6). Suppose that a rabbit scurries by and the native utters "Lo, gavagai". The translator's hypothetical translation for this sentence is "Lo, a rabbit", since this is exactly the situation in which the translator himself would assent to "Lo, a rabbit". The translator continues investigating and checking whether the native assents to "Lo, gavagai" in different situations in which there is similar stimulation. Suppose that the translator has become relatively certain that "Lo, a rabbit" should be taken as the translation of "Lo, gavagai" on the basis of the fact that what prompts the native to assent to "Lo, gavagai" seems to be what would prompt the translator to assent to "Lo, a rabbit" as well. In this sense, both sentences are said to be "stimulus synonymous", or to have similar "stimulus meaning" (see Quine, 1960, pp. 32-33, 51). More specifically, Quine defines "the *affirmative stimulus meaning* of a sentence such as "Gavagai", for a given speaker, as the class of all stimulations … that would prompt his assent" (1960, p. 32). The *negative stimulus meaning* is defined on the basis of all the stimulations which prompt

the native's dissent, and finally, "the stimulus meaning as the ordered pair of the two" (1960, pp. 32-33).[62]

However, sooner or later, the translator has to deal with the sentences that are not directly connected to immediately observable events in the world. To deal with these non-observation sentences, the translator needs to break them down to their parts and find the translation of their components. One such part is the term "gavagai" that may occur in many of these non-observation sentences. The translator, by appealing to the evidence that "Lo, gavagai" and "Lo, a rabbit" were stimulus-synonymous, translates the native's term "gavagai" into "rabbit". However, an important problem emerges at this point: the fact that "Lo, gavagai" and "Lo, a rabbit" had the same stimulus meaning does not show that the terms "gavagai" and "rabbit" have the same reference, or that they are the correct translations of each other. For "rabbit" and "gavagai" to be correct translations of each other, they should have the same extension or referent. But, all the available evidence that the translator has for his claim that "gavagai" and "rabbit" have the same referent is that the native assents to "Lo, gavagai" when a rabbit is present. The evidence that the native assents to "Lo, gavagai" when a rabbit scurries by, however, can be taken to be equally good evidence for other translations of "gavagai": what if, in the native's language, "gavagai" is used to refer, not to a whole enduring rabbit, but to an undetached rabbit part only, or to a temporal stage of rabbit, and the like. For instance, it is possible that another translator comes up with a different list of translations of the native's words, in which "an undetached rabbit part" is suggested as the translation of "gavagai", exactly on the basis of the evidence that the native assents to "Lo, gavagai" whenever an undetached rabbit part is present. In this situation, "Lo, gavagai" and "Lo, an undetached rabbit part" are stimulus-synonymous. As Quine explains, "point to a rabbit and you have pointed to a stage of a rabbit, to an integral part of a rabbit, to the rabbit fusion, and to where rabbithood is manifested" (1960, pp. 52-53). Quine calls such different systems or sets of translations of the native's words, "analytical hypotheses", each of which "equates a native word or construction to a

[62] It is important to note that, for Quine, what prompts the speaker to assent to a sentence, such as "Lo, gavagai", is stimulation, not the object itself, i.e. the rabbit. The reason is that, even if the rabbit itself is replaced by a fake rabbit-looking object, the stimulation will remain the same. Though the angle, lighting, and colour contrast, and the like, will affect the stimulation's strength in prompting the speaker's assent or dissent, Quine believes that "in experimentally equating the uses of 'Gavagai' and 'Rabbit' it is stimulations that must be made to match, not animals' (1960, p. 31). In this regard, Quine defines such stimulation, in the case of visual stimulation, as follows: "A visual stimulation is perhaps best identified, for present purposes, with the pattern of chromatic irradiation of the eye" (1960, p. 31).

hypothetical English equivalent" (1960, p. 70). They are incompatible with each other but fit all the possible data (see, e.g. Quine, 1960, pp. 68-70; 1987, p. 7; 1990, p. 45).

In this way, the reference of the native's term "gavagai" is inscrutable: there are different manuals of translation, different sets of analytical hypotheses, for the native's words, which are incompatible with each other but respect all the relevant physical facts that are, according to Quine's physicalism, the facts about stimulus meaning. Hence, as Quine states, "according to my thesis of indeterminacy of translation, many alternative systems of analytical hypotheses will conform equally to all the facts of stimulus meaning and stimulus synonymy" (1969, p. 313). However, we intuitively expect that the native's expressions have a unique determinate meaning, more fine-grained than their stimulus meanings or extensions. For instance, as discussed above, "Lo, a rabbit" and "Lo, an undetached rabbit part" have the same stimulus meaning, but they intuitively have different fine-grained meanings: "Lo, a rabbit" and "Lo, an undetached rabbit part" are supposed to have their own determinate meaning so that "Lo, a rabbit" seems to mean something different from what "Lo, an undetached rabbit part" does. Granting that two expressions are said to have different intuitive, or fine-grained, meanings if they have the same stimulus meaning but differ in the meaning that we intuitively expect them to have, the translation manuals in question, although all true, intuitively deliver different fine-grained meanings. Therefore, since the facts about stimulus meaning, that is, *all* the acceptable physical facts, fail to discriminate between the translation manuals, which respect all such facts but intuitively deliver different fine-grained meanings, we should conclude that there is no fact about the correct translation of the native's utterances, that is, no fact about their supposedly fine-grained meanings. Therefore, we can characterize Quine's argument from below as follows:

(1) (Quine's Physicalism) Fixing the relevant physical facts fixes all other facts of the matter.

(2) In the case of translation, the relevant physical facts are the facts about stimulus meaning.

(3) (Indeterminacy of Translation through the Inscrutability of Reference) The relevant physical facts, i.e. the facts about stimulus meaning, cannot discriminate between the translation manuals respecting all such facts but delivering different fine-grained meanings, since the facts about stimulus meaning are consistent

with incompatible assignments of referents and extensions to the sub-sentential expressions of the native's language.

(4) Therefore, there are no facts about fine-grained meanings.

In this way, the argument from below can be taken to be an argument with a sceptical conclusion about the existence of the fine-grained meanings that we intuitively expect linguistic expressions to possess.

3.1.3. The Argument from Above

Quine has a second argument for the indeterminacy of translation, which is intended to be "very different, broader and deeper" than the argument from below (Quine, 1970, p. 178). This argument, instead of investigating the process of translating words, relies on a famous distinction which Quine draws between the "underdetermination of theory" and the "indeterminacy of translation".

The underdetermination of theory is the doctrine that there can always be theories incompatible with each other but compatible with all possible data.[63] This doctrine has its roots in Quine's "epistemological holism", according to which "our statements about the external world face the tribunal of sense experience not individually, but only as a corporate body" (Quine, 1951, p. 38). In other words, we are not necessarily bound to entirely dismiss our current theory if we face some experience or observation standing against it; rather, since, based on this holistic view, it is our theory as a whole together with all of its associated hypotheses, presuppositions, and the like, that conflicts with the experience, it is always possible to hold onto any particular statement of the theory, provided that we make proper compensation elsewhere in the theory. Such adjustments can be offered, for instance, by giving up any of the theory's different parts, revising

[63] It is important to note that, for Quine, such theories are empirical theories which help us to predict future events and are testable against evidence. As Quine says, "I cite predictions as the checkpoints of science" (1990, p. 20), and a scientific "theory is tested by prediction. It is the relation of evidential support" (1990, p. 1). For Quine, "the evidential support of science … comes … to be seen as a relation of stimulation to scientific theory" (1990, p. 2), since "from impacts on our sensory surfaces, we … have projected our systematic theory of the external world. Our system is proving successful in predicting subsequent sensory input" (1990, p. 1). In Quine's view, a scientific theory "consists of sentences, or is couched in them" (1990, p. 2). The important part of such theories consists of their observation sentences, which, as discussed in the previous section, are directly linked to the world. These sentences, hence, are "the vehicle of scientific evidence" (1990, p. 5) and "the means of verbalizing the prediction that checks a theory" (1990, p. 4).

any of the associated hypotheses, and so on. Hence, for any set of data, no matter how big it is, we will always have freedom of choice between competing theories. In this sense, as Quine states, "physical theory is underdetermined even by all *possible* observations", such that it "can still vary though all possible observations be fixed. Physical theories can be at odds with each other and yet compatible with all possible data even in the broadest sense. In a word, they can be logically incompatible and empirically equivalent" (1970, p. 179).

For Quine, linguistics is part of natural science and, to this extent, translation manuals naturally inherit the underdetermination of scientific theories: "The totality of possible observations of nature, made or unmade, is compatible with physical theories that are incompatible with one another. Correspondingly, the totality of possible observations of verbal behavior, made and unmade, is compatible with systems of analytical hypotheses of translation that are incompatible with one another" (Quine, 1968, pp. 274-275). However, "the indeterminacy of translation is not just inherited as a special case of the under-determination of our theory of nature. It is parallel but additional" (Quine, 1968, p. 275). To see how the underdetermination of theory and the indeterminacy of translation diverge, we can imagine the radical translator this time trying to translate the native's total theory of physics. This theory is naturally underdetermined by all possible observations; the translator's theory, as the alleged translation of the native's theory, is also underdetermined by all possible observations. Quine's claim is that, to the extent that the native's physical theory is underdetermined by all possible observations, the translation of the native's physical theory is also underdetermined by the translation of its observation sentences. The reason is that the translator starts with translating the observation sentences of the native's theory, that is to say, he tries to match these sentences with the observation sentences of his own theory. But, to find out whether or not the observation sentences match up, he has to find out whether the sentences have the same stimulus meaning, according to Quine's methodology of translation.[64] Suppose that the translator translates all such observation sentences. Since the translator's theory, as well as the native's theory, is underdetermined by all possible observations, the translation of the native's theory is thereby underdetermined by the translation of all the observation sentences of the native's theory. In this sense, the theory that the translator

[64] For a discussion of whether this claim amounts to a presupposition of the conclusion of the argument from below see (Kirk, 1986), and see (Miller, 2006) for the claim that the argument from above need only assume the methodology (as opposed to the conclusion) of the argument from below.

comes up with as the translation of the native's theory is naturally underdetermined by its observation sentences – which are supposed to be the translations of the native's theory's observation sentences. As Quine states,

> insofar as the truth of a physical theory is underdetermined by observables, the translation of the foreigner's physical theory is underdetermined by his observation sentences. If our physical theory can vary through all possible observations be fixed, then our translation of his physical theory can vary through our translations of all possible observations reports on his part be fixed. (1970, p. 179)

Therefore, as the underdetermination of translation implies, "our translation of his observation sentences no more fixes our translation of his physical theory than our own possible observations fix our own physical theory" (Quine, 1970, pp. 179-180).

Nonetheless, the question is: Why does Quine think that translation manuals, and not physical theories, are *indeterminate*, rather than merely underdetermined? To find an answer to this question, it is enough that we combine the mentioned underdetermination of theory, or of translation, with Quine's physicalism. Recall that, according to Quine's physicalism, either the physical facts, which are in the case of translation the facts that the native's and the translator's observation sentences match up because they have the same stimulus meaning, fix the facts about correct translation or there is no fact of the matter. We saw that, even after matching all the observation sentences, the translator still has freedom of choice between different translations of the native's theory, simply because of the underdetermination of physical theory. Fixing the totality of physical facts, i.e. the facts about stimulus meaning, does not fix the facts about correct translation. Therefore, translation is indeterminate, while physical theory is underdetermined. The difference between the underdetermination of theory and the indeterminacy of translation can be further illuminated by coming back again to the case of temperature and comparing it to the case of translation: although our total theory of physics is naturally underdetermined by all possible observations, once we choose our favored theory and hence settle or fix all the truths about the world, we have thereby fixed all the facts about temperature. Once we know an object's molecules' kinetic energy at a certain time, we will know all there is to know about the temperature of the object. However, this is what we lose in the case of translation: if, despite having the totality of physical facts fixed, the translation of the native's theory can still vary, then there is simply no fact of the matter about correct translation. And, as Quine states,

translation is so because "it withstands even all this truth, the whole truth about nature" (1968, p. 275).[65]

A translator might decide, for pragmatic reasons such as simplicity, to choose one of the incompatible translations of the native's total theory of physics; he may attribute to the native one of the competing theories, among the underdetermined ones, as the one the native holds, but it does not mean that other translations, other theories, are not equally well attributable, or that the attributed theory is necessarily the theory that the native *really* believes. The reason is that there is simply no fact of the matter as to which translation of the native's theory is correct: "Where physical theories A and B are both compatible with all possible data, we might adopt A for ourselves and still remain free to translate the foreigner either as believing A or as believing B" (Quine, 1970, p. 180). The totality of physical facts fails to fix the facts about correct translation, that is, the facts about fine-grained meanings.[66] Thus, we can characterize Quine's argument from above as follows:

(1) (Quine's Physicalism) Fixing the relevant physical facts fixes all other facts of the matter.

[65] It is important to note that Quine does not claim that the translation of all physical theories at any level is essentially indeterminate; rather, he is flexible in this regard: to the extent you take your theory to be underdetermined, the translation of your theory is indeterminate at that level. "What degree of indeterminacy of translation you must then recognize … will depend on the amount of empirical slack that you are willing to acknowledge in physics" (Quine, 1970, p. 181). It is not the case that there is widespread agreement about the extent of such empirical slack: "Some will acknowledge such slack only in the highest and most speculative reaches of physical theory, while others see it as extending even to common-sense traits of macroscopic bodies" (Quine, 1970, p. 179). However, Quine himself thinks "the empirical slack in physics extends to ordinary traits of ordinary bodies and hence that the indeterminacy of translation likewise affects that level of discourse" (1970, p. 181).

[66] Quine also clarifies this point by emphasizing that, in physics, theory is the ultimate parameter so that "if you ask a physicist a theoretical question, well out beyond the observation sentences, his answer will be predicted on his theory and not on some unknown and incompatible theory which would have fitted all possible data just as well" (1968, p. 275). Apparently, the same seems to hold in the case of translation: although translation is underdetermined by all possible observations, there are unique translations of the native's theoretical sentences within the translator's favored manuals of translation. "If you ask a linguist 'What did the native say?', where the native's remark was far from the category of observation sentences, the linguist's answer will be predicted on his manual of translation and not on some unknown and incompatible manual which would have fitted all possible linguistic behavior just as well" (Quine, 1968, p. 275). However, the similarity is superficial: it is a mistake if the translator "thinks that the question has a right English answer which is unique up to equivalence transformations of English sentences" (Quine, 1968, p. 275). The translator expects such unique translations, unique meanings of the native's sentences, but "he is mistaken" (Quine, 1968, p. 275). This is the reason why Quine thought that "where indeterminacy of translation applies, there is no real question of right choice; there is no fact of the matter even to *within* the acknowledged under-determination of a theory of nature" (1968, p. 275). There is no fact of the matter about correct translation, about fine-grained meanings.

(2) (Underdetermination of Theory) Physical theory can vary even if all possible observations are fixed; similarly, to the extent that the physical theory varies, the translation of the physical theory varies even if the translation of its observation sentences is fixed.

(3) In the case of translation, the relevant physical facts are the facts about stimulus meaning, (the theories' observation sentences match up in terms of having the same stimulus meaning).

(4) (Indeterminacy of Translation) The translation of physical theory still varies even within an accepted underdetermined theory, that is, even when all physical facts are fixed. [from (1), (2), and (3)]

(5) Therefore, there is no fact about correct translation, about fine-grained meanings. [from (4)]

It is important to note again that Quine's challenge is not epistemological, but rather a challenge with metaphysical/ontological consequences. Quine is very clear on this matter: "The problem is not one of hidden facts, such as might be uncovered by learning more", for instance, about the speaker's environment, his behaviour, or "the brain physiology of thought processes" (1970, p. 180). Rather, for the competing translation manuals, "as to which was right and which wrong there was no fact of the matter" (1978, p. 167).[67] Therefore, as summarized by Quine, "a sentence has a meaning, people thought, and another sentence is its translation if it has the same meaning. This, we see, will not do" (1987, p. 8).

Having introduced Quine's arguments for the indeterminacy of translation, we are now in a position to investigate Davidson's grasp of these arguments.[68]

---

[67] As Quine emphasizes, "the point is not that we cannot be sure whether the analytical hypothesis is right, but that there is not even … an objective matter to be right or wrong about" (1960, p. 73). Again, "if translators disagree on the translation of a Jungle sentence but no behaviour on the part of the Jungle people could bear on the disagreement, then there is simply no fact of the matter" (1987, p. 10). This is the sceptical challenge Quine proposes about what is traditionally conceived as intuitive, or fine-grained, meanings: "What I have challenged is … an ill-conceived notion within traditional semantics, namely, sameness of meaning" (1987, p. 10). And, for Quine, there is "no meaning without sameness of meaning" (1995, pp. 75-76). Hence, "where indeterminacy of translation applies, there is no real question of right choice" (Quine, 1968, p. 275).

[68] For more discussion of Quine's indeterminacy of translation thesis, see (Ricketts, 2011), (Miller, 2006; 2007a, Chapter 4), (Hylton, 2007, Chapter 8), (Soames, 1999, 2003), (Ebbs, 1997, Chapter 2), (Gaudet, 2006, especially Chapter 6), (Gibson, 1994, 2004), and (Kemp, 2006, Chapter 3; 2012, Chapter 4).

3.2. Davidson on Quine's Indeterminacy Arguments

Regarding the arguments sketched above, Davidson endorses almost all of the premises. On Quine's physicalistic approach, although Davidson rejects behaviourism in the sense of implying a reduction of mental concepts to physical concepts,[69] he, at the same time, believes that, in radical interpretation, the ultimate data the radical interpreter can rely on is behavioural: "Perhaps the most important thing he [Quine] taught me was that there can be no more to the communicative content of words than is conveyed by verbal behavior" (Davidson, 1999a, p. 80). In other words, Quine and Davidson share the general theme of their (supervenience) physicalism; as Davidson states,

> [his view] is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect. (1970, p. 214)[70]

More particularly, Davidson describes his agreement with Quine on this matter as follows: "We both assume that the observed range of phenomena which prompt assent and dissent to certain sentences allow us to connect those sentences to events and objects in the world" (1997a, pp. 77-78). Again, "Quine's suggestion, which I shall essentially follow, is to take *prompted assent* as basic, the causal relation between assenting to a sentence and the cause of such assent" (1983, p. 147). Hence, for both, a certain phenomenon in the world (for Davidson, a distal stimulus, and, for Quine, proximal stimulation) causes the speaker to be disposed to assent to or dissent from a certain sentence under specific circumstances.

---

[69] For instance, Davidson says, "Psychological concepts, I have been arguing, cannot be reduced, even nomologically, to others. But they are essential to our understanding of the rest" (1974b, pp. 243-244), and again, "psychological characteristics cannot be reduced to the others" (1973b, p. 253). See also (Davidson, 1999b, p. 599; 1970, p. 214).

[70] Quine also endorses Davidson's "anomalous monism" as "token physicalism": "I acquiesce in what Davidson calls anomalous monism, also known as token physicalism" (1990, p. 72). Davidson describes his anomalous monism as follows: "Finally there is *anomalous monism*, which classifies the position I wish to occupy. Anomalous monism resembles materialism in its claim that all events are physical, but rejects the thesis, usually considered essential to materialism, that mental phenomena can be given purely physical explanations. Anomalous monism shows an ontological bias only in that it allows the possibility that not all events are mental, while insisting that all events are physical" (1970, p. 214).

Moreover, Davidson, in different places in his writings, declares his acceptance of Quine's indeterminacy of translation thesis: "I accept the indeterminacy of *Word and Object*, which particularly affects sentences and the subsequent extension to names and predicates" (1999a, p. 80);[71] he restates his agreement with Quine as follows: "There remain two important kinds of indeterminacy on which we agree: indeterminacy due to what Quine calls the inscrutability of reference, and indeterminacy that results from the blurring of the distinction between the analytic and the synthetic" (1997a, p. 78).[72]

However, he does not believe that Quine's thesis of the indeterminacy of translation leads to any sceptical conclusion about meaning. In his reply to Rorty (Davidson, 1999b), he makes this point clear. Rorty's worry is that "indeterminacy implies that there is something mysterious, second rate, or even not quite real, about the mental, that there are no 'fact of the matter' about meaning or the propositional attitudes" (Davidson, 1999b, p. 595);[73] Davidson continues that, on the contrary, "the mental is no more mysterious than molecular biology or cosmology" (1999b, p. 595). According to Davidson, the indeterminacy of translation should not be considered as a source of worry, since "indeterminacy as I understand it is endemic in all disciplines" (1999b, p. 596). The indeterminacy is not surprising, nor does it result in any scepticism about meaning: "Given the richness of all natural languages, it would be surprising if it were not always possible to describe the facts of any discipline in many ways. Such indeterminacy does not threaten the reality of what is described" (1999b, p. 596). The puzzling issue, however, is how Davidson could endorse the main premises of Quine's arguments and yet reject the resulting sceptical conclusions, that is, claim that "this negative point [Quine's indeterminacy of translation thesis] does not entail that there are no facts of the matter" (Davidson, 1999b, p. 596). The immediate question that arises is: What are the facts of the matter that Davidson refers to? According to Davidson, "the facts are the empirical relations between a speaker, her sentences, and her environment. This pattern is invariant" (1999b, p. 596). In other words, it is a fact, for Davidson, that a speaker, on a specific occasion, assents to a sentence, or holds it true, because of what

[71] See the Appendix for an introduction to Davidson's thesis of the indeterminacy of interpretation.

[72] For Davidson, the indeterminacy appears in the process of translation as well as interpretation: "Here I accept Quine's thesis of the indeterminacy of translation, and extend it to the interpretation of thought generally" (1991a, p. 215, fn. 3).

[73] Rorty reads and criticizes Davidson's view from a pragmatist point of view. For him, neither a factualist nor a non-factualist view is the right way to treat the indeterminacy of translation and related issues about meaning-determinism. See (Rorty, 1999).

she means by the sentence and what she believes about the world; this invariant relation held between the speaker, the sentence (to which she assents or from which she dissents), and the world (the stimulus prompting such assent or dissent), is what Davidson intends by the "facts of the matter".

From this point of view, Davidson reads Quine's indeterminacy of translation thesis as follows:

> [Quine's point was actually that] there is no more to the identification of meanings that is involved in capturing these complex empirical relations. This can sound like a negative thesis, and it is; it is an attack on the idea that meanings can be captured in exactly one way, by pinning Platonic meanings on expressions. (1999b, p. 596)

Based on this reading, the conclusion from Quine's arguments is rather the claim that there is no *unique* way of capturing the same fact about what a speaker means. But, if what Davidson thinks of as the facts about meaning are just Quinean facts about stimulus meaning, then Quine has already shown that these facts cannot be regarded as facts about fine-grained meanings, since they could not discriminate between translation manuals respecting all such facts, but intuitively delivering different fine-grained meanings. In other words, if the invariant pattern Davidson talks about is just what we get from the speaker's behavioural dispositions to assent or dissent, then why does Davidson think that the Quinean indeterminacy does not threaten the reality of meaning facts? To see how his response to Quine's indeterminacy argument goes, we need to consider the main supporting reason Davidson proposes in order to sustain his reading of Quine's thesis of the indeterminacy of translation.

3.2.1. The Measurement Scales Analogy

Davidson employs an analogy to show why, for him, the indeterminacy of translation thesis is not threatening. The analogy he draws is between measuring temperature, weights, and the like, by employing different scales of measurement and translating a speaker's utterances by using different, but equally acceptable, translation manuals. According to Davidson, we can see different translation manuals, which respect all the facts about stimulus meaning, as compatible and equally good ways of capturing the same facts about what a speaker means, just as different scales of measuring

temperature or weight compatibly and equally well capture the same facts about the temperature or weight of an object:

> With weight, an arbitrarily chosen positive number is assigned to some particular object; relative to that assignment, the numbers that measure the weights of all other objects are fixed. You can get an equally good way of keeping track of weights by multiplying the original figures by any positive constant. (1999b, p. 596)

In the same way, "a theory of measurement for temperature leads to the assignment to objects of numbers that measure their temperature"; but, "the *pattern* of assignments is significant. (Fahrenheit and Centigrade temperature are linear transformations of each other; the assignment of numbers is unique up to a linear transformation)" (Davidson, 1977, p. 225). In the case of temperature, there are facts of the matter as to what the temperature of the object is, which can be represented in different ways, and these ways of representation are compatible, since they are simply the linear transformations of each other informing us of the same fact about temperature.

Analogously, Davidson suggests that "the attribution of attitudes is analogous in many ways to the measurement of various magnitudes" (1997a, pp. 74-75). As he puts it,

> each of us can think of his own sentences (or their contents) as like the numbers; they have multiple relations to one another and to the world. … Just as endless sets of numbers allow us to keep track of the same complex structures in the world, so our sentences can be used in endless different ways to keep track of the attitudes of others, and of what they mean. (1999b, p. 596)

It is important to note that the point of the analogy is to show "why different assignments of objects can capture all the relevant information about situation without compromising the truth or 'reality' of the situation" (1997a, p. 75). In other words, the fact that there are different ways of capturing the facts about temperature, according to Davidson, does not lead to any doubt about the existence of the facts about temperature, just as "no one thinks the fact that we can register weight in either pounds or kilograms shows that there is something unreal about the weight of an object: different sets of numbers can be used to keep track of exactly the same facts" (1997a, p. 75). One of the important remarks from Davidson in making this point clear appears in the following passage:

> Because there are many different but equally acceptable ways of interpreting an agent we may say, if we please, that interpretation or translation is indeterminate, or that there is no fact of the matter as to what someone means by his or her words. In the same vein, we could speak of the indeterminacy of weight or temperature. But we normally accentuate the positive by being clear

> about what is invariant from one assignment of numbers to another, for it is what is invariant that is empirically significant. The invariant is the fact of the matter. We can afford to look at translation and the content of mental states in the same light. (1991a, pp. 214-215)

However, how can such an analogy show that the Quinean indeterminacy of translation does not lead to a sceptical conclusion regarding the facts about fine-grained meanings? To run from the sceptical conclusion of Quine's arguments, Davidson seems to be suggesting that Quine's indeterminacy thesis has to be read as essentially having epistemological as opposed to ontological or metaphysical consequences. On this reading, the thesis, and Quine's arguments, questions the existence of a unique way of capturing the meaning facts, rather than questioning the existence of such facts. In the following section, I will consider whether Davidson's reading of Quine's arguments is satisfactory.

3.2.2. The Puzzle

Davidson appealed to the measurement scales analogy in order to justify the way he reads Quine's indeterminacy of translation thesis, i.e. as a thesis with merely epistemological consequences. According to Davidson, "interpretation can proceed because we can accept any of a number of theories of what a man means, provided we make compensating adjustments in the beliefs we attribute to him", while "the remaining indeterminacy should not be judged as a failure of interpretation, but rather as a logical consequence of the nature of theories of meaning (just as it is not a sign of some failure in our ability to measure temperature that the choice of an origin and a unit is arbitrary)" (1973b, p. 257). If Davidson really takes this analogy seriously, his reading of Quine's thesis will be susceptible to the same sort of misunderstanding that Quine (1968) attributes to Chomsky. Chomsky (1968) presents Quine's indeterminacy of translation thesis as a special case of the underdetermination of theory by all possible observations. According to Chomsky,

> it is, to be sure, undeniable that if a system of "analytical hypotheses" goes beyond evidence then it is possible to conceive alternatives compatible with the evidence … Thus the situation in the case of language, or "common sense knowledge", is, in this respect, no different from the case of physics. (1968, p. 61)

For Chomsky, hence, the indeterminacy of translation is a natural and familiar phenomenon that we face in the case of physics or other natural sciences too, that is,

that a certain body of hypotheses about a subject matter may go beyond all the evidence and data we can collect. We can always have competing theories, or different sets of hypotheses, which are compatible with all possible evidence. We do not, in these cases, state that there is no fact of the matter about what our theories deal with; we do not, for instance, claim that, since our theories about chemistry are underdetermined by all possible observation, there is thereby no fact of the matter about chemistry. If indeterminacy is understood in this way, then Chomsky's problem regarding the case of translation is "why should all of this occasion any surprise or concern?" (1968, p. 67). Chomsky's suspicion about the importance of the indeterminacy of translation thesis has its roots, according to Quine, in his failure to appreciate the vital distinction between the indeterminacy of translation and the underdetermination of theory. This point is the one that "Chomsky did not dismiss … He missed it", as Quine says (1968, p. 276). The indeterminacy of translation is not the underdetermination of theory. To see why Davidson's appeal to the measurement scales analogy blurs, in a similar way, this distinction, we have to come back to Quine again and consider why he thought that there is such a difference between the indeterminacy of translation and the underdetermination of theory.[74]

Consider the case of measuring temperature. In this case, generally speaking, fixing the relevant physical facts, i.e. the facts about the object's molecules' kinetic energy at a certain time (figured in the formula "$1/2mv^2$") fixes the facts about temperature so that, once we know the amount of kinetic energy that the molecules have at a given time, we would know what the temperature of the object is: the fact about the molecules' kinetic energy uniquely determines the temperature of the object. Suppose that the amount of the kinetic energy of the molecules at the time t is α. This α, for sure, can be represented in different ways, either in T(C°), or in T(F°), or in any other linear transformations of them. This, however, is not the *Quinean* indeterminacy. The crucial point is that this alleged indeterminacy has no ontological/sceptical consequence: the claim that there is no fact of the matter as to which scale is the right one to represent the unique amount of the kinetic energy of the molecules (i.e. the object's temperature) is consistent with the claim that there *are* facts about temperature. Davidson has a similar analogy in mind: as different scales of measuring temperature can capture the same fact about temperature, the different manuals of translation can capture the same fact about

[74] Interestingly, for both Davidson and Chomsky, the indeterminacy of translation is not a source of worry, concern, or surprise. See, e.g. (Chomsky, 1968, p. 67) and (Davidson, 1991a, p. 214).

what a speaker means. We should be led to scepticism about the facts of the matter in neither case. However, Quine's main claim was that the case of translation is essentially different. According to Quine, the indeterminacy happens when settling the totality of physical facts falls short of fixing the facts about meaning so that we still have freedom to choose between intuitively incompatible translations. In this regard, considering Davidson's measurement scales analogy, it is hard to see what is supposed to play the role of something like α in the case of translation: if Davidson tends to stay close to Quine's project, then α, in the case of translation, would need to be stimulus meaning, but Quine's arguments already prevent stimulus meaning from being available to play this role. Quine's point was not that the same facts about meaning can be represented differently; rather that there are no such facts to be represented, since the only available physical facts, i.e. the facts about stimulus meaning, fail to discriminate between the manuals of translation respecting all such facts but intuitively delivering different fine-grained meanings. This is the reason we do not question the factuality of temperature: because when we settle the physical facts, our facts about temperature are thereby fixed so that we are left with no choice among different temperatures (we are provided with all we could know about the temperature of the object). We do, however, question the existence of fine-grained meanings because, with all of our physical facts fixed, we are left with choices among incompatible manuals of translation. This is, however, in conflict with Davidson's claim that "different assignments of objects can capture all the relevant information about situation without compromising the truth or 'reality' of the situation" (1997a, p. 75).[75]

Therefore, it seems that, by appealing to this analogy, Davidson is committed to a misunderstanding, of which Quine warned us before. The apparent indeterminacy in the case of temperature is radically different from the metaphysical/ontological indeterminacy in the case of translation. As Quine emphasizes, the difference is due to the fact that "in the case of natural science … there is a fact of the matter, even if all

[75] I take Gary Kemp (2012) to be in agreement with us on this point. He puts the problem in terms of Davidson's reading of the inscrutability of reference. According to Kemp, "but the point about meaning is precisely that where there are disagreements as regards reference amongst otherwise sound interpretation schemes, the intuitive disagreement *remains*, even when all the facts are before us" (2012, p. 126). Kemp thinks that Davidson's appeal to the measurement scales analogy cannot help him to resist such indeterminacy: "To choose a temperature scale … is not itself to measure anything; one can say what units we plan to use before measuring something. Whereas to decide on a particular scheme of interpretation just is to decide on reference. And that is just to say that the concept of reference is indeterminate, unlike the concept of a unit of length" (2012, p. 127).

possible observations are insufficient to reveal it uniquely" (1987, p. 10). Natural science is underdetermined by all possible observation, but "suppose that we have settled for one of the many overall theories of nature that fit all possible observation. Translation remains indeterminate, even relative to the chosen theory of nature. Thus, the indeterminacy of translation is an indeterminacy additional to the underdeterminaction of nature" (1987, p. 10). This is a vital distinction underpinning the sceptical conclusion of Quine's arguments, but it seems, however, to be neglected by Davidson.[76]

3.3. Conclusion: A Dilemma for Davidson

Therefore, Davidson's attempt to read Quine's indeterminacy of translation thesis as a non-dangerous, non-threatening claim with essentially epistemological consequences for meaning facts seems to fail, especially when we take into consideration the analogy he provides as a support. We can put the puzzle in the form of a dilemma for Davidson: either, in Davidson's view, facts about stimulus meaning are the only facts relevant to fine-grained meaning or fine-grained meaning is fixed by facts in addition to facts about

[76] An important question that arises is: Given that Davidson and Chomsky appear to be committed to the same confusion, why does Quine disagree with Chomsky but (apparently) not with Davidson? Interestingly, Davidson, at one point in his paper "Reply to Richard Rorty" (1999b), confesses that "what I am expounding is my own view of indeterminacy and underdetermination … I don't know about Quine, but my way of taking the distinction has no ontological significance whatever" (1999b, p. 597). It is, however, difficult to see in what sense, according to Davidson, there can be a real distinction between indeterminacy and underdetermination, without any ontological significance. Indeed, there is no problem with Davidson's taking the indeterminacy of translation, or the underdetermination-indeterminacy distinction, to *mean* something different from what Quine took them to do. The problem, however, is that, Davidson does not explain *why* the indeterminacy of translation does not amount to scepticism about fine-grained meanings. Except the passage quoted above, Davidson has always taken himself to be committed to Quine's project, especially the way Quine puts the indeterminacy of translation. Perhaps, it is just at this latest stage of his work (i.e. 1999s onward) that Davidson started realizing his deep divergence from Quine. But, if Davidson was inclined to give up on the idea that Quine's arguments lead to scepticism about fine-grained meaning, then what would be the resulting view which Davidson would endorse as an alternative? It seems that such a view would be non-reductionism about meaning. (In Chapter Four and Five, we will also argue that the best option for Davidson to resist KW's sceptic's sceptical problem is to hold a form of non-reductionism about meaning.) An explicit endorsement of such a view, nonetheless, is hard to find in Davidson's writings. Even if Davidson wanted to accept non-reductionism about meaning, it is still difficult to see how his use of the measurement scales analogy could him to support this view. We will see, in the next section, that what such an analogy suggests is rather reductionism about fine-grained meaning. Besides, our claim about Davidson's problematic reading of Quine will be supported by our claim, which will be outlined in Chapter Four, that Davidson's reading of KW's sceptical argument suffers from the same sort of problem: he takes KW's sceptical argument to be proposing an essentially epistemological problem about meaning, rather than a metaphysical/sceptical one.

stimulus meaning. If Davidson believes that facts about stimulus meaning are the only facts relevant to fine-grained meaning, then he has to accept Quine's indeterminacy of translation arguments. As a consequence, Davidson has to endorse the sceptical conclusions of Quine's arguments about fine-grained meaning: there are no facts about such meanings over and above the facts about stimulus meaning. If that is the case, then the question is whether Davidson's use of the measurement scales analogy can help him to justify his reading of Quine's thesis of the indeterminacy of translation. What does the analogy suggest on this horn of the dilemma? The analogy seems to imply that facts about fine-grained meaning are to be reduced to physical facts (facts about stimulus meaning) and thereby *eliminated*.[77]

However, if Davidson believes that fine-grained meaning is fixed by facts in addition to facts about stimulus meaning, he thereby admits that there *are* facts about fine-grained meaning. If this is Davidson's position, how could he claim to be faithful to Quine? How could he accept *Quine*'s indeterminacy of translation thesis? In this case, he has to embrace his radical divergence from Quine's project, which he is highly reluctant to do, at least in his most important works on the topic. But, again, the question is whether the measurement scales analogy is of any help for Davidson here. The answer seems to be negative. In this scenario, the relationship between fine-grained meaning and stimulus meaning is *not* like the relationship between temperature and molecules' kinetic energy, since, in the case of temperature, there are no facts over and above the physical facts about temperature, i.e. facts about molecules' kinetic energy. The facts about stimulus meaning, on this horn of the dilemma, are *not* supposed to be the only facts about fine-grained meaning, while the facts about molecules' kinetic energy are the only relevant facts about temperature.

In either case, therefore, the measurement scales analogy does not help Davidson to support his reading of Quine's indeterminacy of translation and block the sceptical outcome of Quine's arguments. The only way Davidson can accept the existence of fine-grained meaning facts and, at the same time, Quine's indeterminacy of translation thesis is to view the latter as epistemological. By doing so, however, he blurs the

[77] So here the relationship between fine-grained meaning and stimulus meaning turns out to be like the relation between polywater and ordinary-water-with-impurities rather than the relation between temperature and mean kinetic energy: as in the polywater/water-with-impurities case we have an eliminative reduction rather than the sort of tolerable or vindicative reduction we have in the water/H20 and temperature/mean kinetic energy cases. For tolerable or "vindicative" vs. eliminative reductionism, see (Miller, 2013a, pp. 185-186).

distinction between the underdetermination of theory and the indeterminacy of translation – the same mistake Chomsky allegedly made in his interpretation of Quine's thesis of the indeterminacy of translation. For instance, Davidson says that what he is actually concerned with is "whether the empirical and formal constraints on a theory of truth sufficiently limit the range of acceptable theories" (1977, p. 224). He believed that putting enough formal and empirical constraints on theories of truth will lead to an acceptable number of theories all of which equally well capture the facts about what the speaker means and believes: "The empirically equivalent theories it [indeterminacy] accepts as equally good for understanding an agent are *not* incompatible, any more than the measurement of weight in pounds or kilos involves incompatible theories of weight" (Davidson, 1997a, p. 76, emphasis added).[78] As I have stressed, however, this claim blurs the distinction between the indeterminacy of translation and the underdetermination of theory.[79] The incompatibility of the manuals of translation and

[78] Again, in "Radical Interpretation" (1973a), he states: "It is not likely, given the flexible nature of the constraints, that all acceptable theories will be identical. When all the evidence is in, there will remain, as Quine has emphasized, the trade-offs between the beliefs we attribute to a speaker and the interpretations we give his words. But the resulting indeterminacy cannot be so great but that any theory that passes the tests will serve to yield interpretations" (1973a, p. 139).

[79] Ramberg (2000) reconstructs Davidson's reading of the Quinean distinction between underdetermination and indeterminacy as follows: "The ineliminable possibility of alternative ways of stating the facts does nothing to threaten the factuality of our statements. … facts, for Davidson, are just what true statements express, and the truths expressed by psychological and semantic ascriptions cannot, according to anomalous monism, be restated in terms of non-intentional theory … once we give up the idea, as Davidson has done, that one vocabulary is especially suited to express the facts, …, the independence of theory-decisions across vocabularies becomes an ontologically insignificant fact" (2000, p. 356). Ramberg continues: "We can make sense of Davidson's use of Quine's indeterminacy-underdetermination distinction … without seeing him as pursuing *ontology* at all, but simply as discussing the relations between different descriptive strategies. The indeterminacy-underdetermination contrast, on my reconstruction, need not be backed by anything more philosophically potent than a distinction between vocabularies" (2000, p. 357). First of all, as mentioned before (Section 3.2, p. 127), Quine agrees with Davidson's anomalous monism. If that is the case, then it is not clear how anomalous monism can result in such a divergent reading of Quine's indeterminacy of translation thesis. But, the more important point is that, again, Quine's problem is supposed to be metaphysical: as a consequence of Quine's arguments, there will be no ontological gap between the facts about meaning and the physical facts about meaning, since there is no fact about meaning beyond the facts about stimulus meaning. Davidson's anomalous monism is compatible with the above claim: there is no ontological gap between the physical and the mental/semantical, though using semantical/intentional vocabularies is indispensible and irreducible to behavioral or physical descriptions (see Davidson, 1997a, p. 72). If Davidson agrees with this view, that is, with the claim that there are no facts about meaning other than the physical facts about meaning, then how can anomalous monism lead to anything but scepticism about fine-grained meanings? If, for Ramberg, Davidson's point is only that facts about stimulus meaning can be *re-described* by using different vocabularies, or even that using different vocabularies is inevitable in our description of such facts, then it is hard to see how this belief can block the sceptical outcome of Quine's arguments, considering that Davidson and Quine agree as far as physicalism is concerned.

the resulting sceptical conclusion about fine-grained meaning was the whole point of Quine's indeterminacy of translation arguments.[80]

In the next chapter, we will use these considerations in our discussion of Davidson's grasp of KW's sceptic's sceptical argument.

[80] Nonetheless, the appeal to the measurement scales analogy can be taken to be signaling Davidson's attitude towards a sort of factualist account of meaning. According to this analogy, what the radical interpreter does is at best keep track of an independent and complex structure in the world, that is, to "keep track of the attitudes of others, and of what they mean" (Davidson, 1999b, p. 596). The question, hence, will be: what sort of view does this claim amount to? We will take up this issue in Chapter Five.

4. Chapter 4: Davidson on Kripke's Wittgenstein

Introduction

The purpose of this chapter is to provide an answer to the question: How does Davidson respond to KW's sceptic's sceptical argument? Davidson has made a variety of criticisms of both KW's sceptic's sceptical argument and KW's sceptical solution, although these criticisms are not, for the most part, clearly characterized and defended by Davidson. We will characterize these criticisms and investigate whether they are plausible. Davidson has also proposed an alternative to KW's sceptical solution, which is supposed to defuse the sceptical paradox. Evaluating such an alternative solution is the second purpose of this chapter, i.e. to explore whether or not Davidson's alternative view of meaning is indeed capable of resisting KW's sceptic's sceptical challenge.

We shall start with introducing Davidson's interpretation of KW's sceptic's sceptical argument and KW's sceptical solution. Generally speaking, I will argue that the criticisms Davidson makes of KW's sceptical argument are not acceptable because of his problematic reading of KW's sceptic's sceptical argument. In addition, I will argue that Davidson's alternative solution to the sceptical problem fails, since he fails to fully appreciate the sceptical part of KW's sceptic's sceptical argument and KW's sceptical solution. I will also argue that Davidson's alternative solution suffers from the same sort of problem that vitiated his reading of Quine's indeterminacy of translation thesis.

4.1. Davidson on KW's Sceptical Argument and Sceptical Solution

Davidson believes that his alternative account of meaning is indeed "a weaker and more plausible alternative to Kripke's proposed account of what is required in order to mean something by what one says" (1992, pp. 116-117). This alternative is supposed to defuse KW's sceptic's sceptical problem and to differ from KW's sceptical solution with respect to "the flavor it gives the social aspect of language" (1992, p. 121). Davidson has characterized KW's sceptical problem in different ways. He interprets KW's problem as the problem, which one faces when one tries to explain the linguistic practices of an individual speaker, e.g. the speaker's practice of meaning something by an utterance, or following one rule rather than another, or speaking one language rather than another, especially when the speaker is considered in isolation. If it is

indeterminate whether a speaker means one thing rather than another by a word, or whether she follows one rule or the other, it will be indeterminate, considering the holistic feature of meaning and the compositional character of natural languages, whether the speaker speaks one particular language or the other. We can call these interrelated problems the deep problem of meaning-determination.

Generally speaking, Davidson claims that, to solve the problems, there has to be a second person linguistically interacting with the speaker. To this extent, he agrees with KW: the existence of a social setting is necessary if there is to be any language, meaning, or thought at all. He criticizes KW, however, for demanding more than that: the appeal to the shared practice of a community of speakers, such as their following the same agreed-on conventions, shared rules, standards, norms, and the like, is neither necessary nor sufficient in order to provide a criterion for the speaker's success in the practice of meaning something by her words. Rather, linguistic interaction between an individual speaker and a second person, i.e. the speaker intending to be interpreted in a particular way and her being interpreted as such, would equip the individuals with all they need to make meaningful responses.

Generally speaking, having seen matters in this way, the general problems with Davidson's grasp of KW's sceptical problem and solution can be put as follows: (I) KW's sceptic's sceptical problem does not merely concern the linguistic practices of an individual speaker so that the linguistic interaction between the speaker and a second person can by itself solve it; rather, it is applicable to two or more individual speakers linguistically interacting with each other, whose successful communication is supposed to be explained in terms of their following one rule rather than another, or in terms of the assumption that there are some facts of the matter as to what they mean by their words. (II) KW's sceptic's sceptical problem, as well as KW's sceptical solution, do not necessarily depend on taking for granted the rule-following picture of meaning, which is taken by Davidson to be highly problematic; rather, it arises against any sort of explanation, which relies on the assumption that there are facts of the matter as to what a speaker means by her words. (III) Davidson's suggested "mutual interpretation", or "linguistic triangulation", as the alternative solution to KW's sceptic's sceptical problem, fails to resist the sceptical problem, since it still depends on the claim that there are facts about the speaker as to what she means by her words. Davidson's suggested alternative, hence, reveals the deep problem with his treatment of KW's sceptic's sceptical argument as an argument with *sceptical* outcomes.

I start with introducing the three places, in which Davidson mentions KW's sceptic's sceptical argument and KW's sceptical solution. In each case, I will evaluate his characterization of KW's argument and solution and the criticisms he makes of each.

4.1.1. The First Place: The Rule-Following Picture of Meaning

Davidson's important paper, "The Second Person" (1992), is the main paper, in which he introduces KW's sceptical problem and tries to provide a response to it. The following passage is one of the clearest passages in this paper, which offers a characterization of KW's sceptical problem and solution:

> Kripke concentrates on the idea of following a rule. According to this idea, to speak a language is to follow rules. The rules specify what it is to go on 'in the same way'; how, for example, to use a word. There is, however, no inner mental act or process of 'grasping' or of 'following' the rule, so no study or knowledge of what is inside the speaker will reveal whether she is following one set of rules or another. Interpreters simply judge that a speaker is following the same rule they (her interpreters) are if the speaker goes on as they would. Put in terms of meaning: we judge that a speaker means what we would if we were to utter the same words if she goes on as we would. (1992, p. 113)

On this reading, KW's sceptic's argument rests on the rule-following picture of meaning: a language is a set of rules determining the correct use of words for a speaker so that to speak a language is to follow such rules. In this sense, the speaker uses a word correctly and, hence, means something by a word, only if she follows one rule rather than another. Regardless of whether or not there is any inner mental process of following, or grasping, a rule, Davidson aims to directly criticize the appropriateness of the rule-following picture of language: "Rules can be a help in learning a language, but their aid is available only in the acquisition of a second language. Most learning of how to use words is accomplished without explicitly learning any rules at all" (1992, p. 113). In the same way, he criticizes KW's sceptic's use of mathematical examples, since, according to Davidson, such examples rely on an analogy between speaking a language and doing mathematics: "We normally follow no procedure in speaking; nothing in normal speaking corresponds to adding a column of numbers" (1992, p. 114). For Davidson, it is in doing mathematics, and not in speaking a language, that we can legitimately talk about following rules. This line of criticism is directly related to the criticisms he made of Conventionalism or the Common View, which we discussed in Chapter Two when we introduced his negative argument against the Common View or

Conventionalism: "Linguistic communication does not require … rule-governed repetition" at all (1984b, pp. 279-280); "convention does not help explain what is basic to linguistic communication" (1984b, p. 280). For Davidson, the practice of winning a game, in which rules play an essential role, has a combination of features that the practice of speaking a language lacks. Davidson, hence, employs the material from his criticisms against Conventionalism in order to criticize KW's sceptic's sceptical argument, which shows that, in his view, KW's sceptical argument presupposes the Common View. His main objection to KW's sceptic's sceptical argument, thus, is that this argument relies on presupposing a highly problematic view.

Davidson also challenges KW's sceptical solution. If success in the practice of meaning something by words can be explained without any appeal to the existence of some shared rules or conventions, then any view which takes the existence of such rules, conventions, or standards as essential to the existence of our linguistic practices is wrong, including KW's sceptical solution. As we saw in our discussion of Davidson's alternative view in Chapter Two (see Section 2.2.1), Davidson believed that success in the practice of meaning something by an utterance can be explained in terms of success in mutual interpretation, which does not take for granted the existence of shared rules or conventions. Doing what others do, or going on as others do, does not give an adequate picture of the practice of meaning something by words. As Davidson insists,

> the theoretical possibility of communication without shared practices … shows that such sharing cannot be an essential constituent in meaning and communication. If I am right, then important claims by Tyler Burge, Saul Kripke, and perhaps Wittgenstein and Dummett must be false, for certainly the first two have insisted that speaking in the 'socially accepted' way is essential to verbal communication. (1994, p. 119)

Hence, he explicitly rejects any view implying that "linguistic communication requires that a speaker go on in the same way as others do – that to mean something in speaking, one must mean the same thing by the same words as others do" (1992, p. 114). He even ruled out the requirement that the speaker herself needs to continue speaking as she did in the past, as discussed in Chapter Two. Rather, for him, it is success in mutual interpretation that is both necessary and sufficient for explaining success in the practice of meaning something by word.[81] As a result, Davidson concludes that KW's sceptical solution is wrong:

[81] As Davidson emphasizes, "if you and I were the only speakers in the world, and you spoke Sherpa while I spoke English, we could understand one another, though each of us followed different 'rules'

> So, while it may be true that speaking a language requires that there be an interpreter, it doesn't follow that more than one person must speak the same language. … I conclude that Kripke's criterion for speaking a language cannot be right; speaking a language cannot depend on speaking as someone else does (or as many others do). (1992, pp. 114-115)

(a) Problems

Davidson took KW's sceptic's sceptical argument to be resting on the rule-following picture of meaning. The initial problem with this claim is that KW's sceptical argument does not necessarily rely on such a conception of the practice of meaning something by words. As discussed in Chapter One (Sections 1.1.1 and 1.3), KW's sceptic's question was: What is it that makes it the case that a speaker means one thing rather than another by the words she utters? In Chapter One (Section 1.4), we characterized two strands of KW's sceptic's sceptical argument: the first one took for granted the rule-following picture of meaning, while the other assumed the existence of meaning facts. According to the first strand of the sceptical argument, the Common View was taken as a view implying that if a speaker, S, means something by a word, then there is a rule, which determines the correct application of the word for S. However, according to the second strand of the sceptical argument, the Common View stated that, if S means something by a word, then there is a possible meaning fact, or a state of affairs, that determines the correct application of the word for S. We took KW's sceptic to be arguing that both characterizations of the Common View lead to the Radical Sceptical Conclusion, i.e. that there is no such thing as meaning anything by any word, since nothing about S can determine whether S follows one rule rather than another, *or* nothing about S (including the facts about her intention) can constitute the fact that S means one thing rather than another. The second strand of the sceptical argument does *not* bring in any talk of rules or conventions. This means that KW's sceptical argument will not be ruled out by criticizing the rule-following picture of meaning. In other words, one cannot rule out KW's sceptic's sceptical argument in general simply by arguing that following rules is neither necessary nor sufficient for explaining the success in the practice of meaning something by words, i.e. that the Common View, as characterized in the first strand of the sceptical argument in terms of the rule-following picture of meaning, is wrong.

(regularities)" (1992, p. 114). Hence, "what would matter, of course, is that we should each provide the other with something understandable as a language. This is an intention speakers must have; but carrying out this intention … does not involve following shared rules or conventions" (1992, p. 114).

Moreover, regarding the characterization of the Common View in terms of the rule-following picture of meaning, KW's sceptic can even claim that the rule-following picture of the practice of meaning is nothing but an *analogy* used to facilitate the way, in which the problem about meaning-determination can be put forward. In this sense, KW's sceptic, to run the first strand of his argument, does not need to claim that following rules and speaking a language are the same things. Hence, criticizing the rule-following picture of meaning does not necessarily amount to a rejection of even the first strand of KW's sceptic's sceptical argument.

Davidson, however, might respond to the above claim from the sceptic by arguing that even drawing such an analogy is implausible, since it was argued by Davidson that the practices, in which following rules is essential, such as playing and winning a game, are not even analogous to the practice of speaking a language. Hence, employing such an analogy by KW's sceptic is not legitimate. Nonetheless, KW's sceptic can again reply: Let's forget about the rule-following picture of meaning and the analogy altogether; I can instead challenge any proposal, which aims at explaining the success in the practice of meaning something by a word in terms of assuming the existence of any fact about the speaker's mental, behavioural, or social life, which is supposed to make it the case that the speaker means one thing rather than another. The second strand of KW's sceptic's sceptical argument led to the Basic Sceptical Conclusion that there is no fact of the matter as to what a speaker means by her words. We can conclude that, although Davidson's objections to KW's sceptic's sceptical argument may raise some serious problems for the first strand of the argument, it cannot result in a decisive rejection of the sceptical argument, since the second strand of this argument can still successfully establish the sceptic's desired Radical Sceptical Conclusion.

The second problem with Davidson's reading of KW is that his alternative solution fails to deal with the sceptical problem and, hence, to be counted as an *alternative* to KW's sceptical solution; it rather blurs the distinction between the metaphysical and the epistemological problems about meaning-determination. The sceptical problem questions any factually based explanation of the success in the practice of meaning something by words and is, thereby, applicable to two speakers mutually interpreting each other *if* one tries to explain their success in interpretation in terms of some fact about what is inside or outside their heads. As we will see in the following sections, KW's sceptic can raise his sceptical problem in the case of Davidson's alternative

solution, as put in terms of "mutual interpretation", "being interpreted as intended", or "linguistic triangulation".

4.1.2. The Second Place: "The"-Language-Determination

Davidson again considers the linguistic practices of an individual speaker and introduces the problem of determining 'the' language she speaks. He suggests that it is the following problem that troubles KW:

> The fact that all the publicly available evidence with regard to a speaker or group of speakers, even if imagined to exhaust all possible such evidence, might be consistent with many different languages … ought not in itself to worry us, for we can agree that it is enough to know that a speaker speaks any one of a set of empirically equivalent languages, as long as the empirical constraints clearly define the set.[82] (1992, p. 110)

Davidson, however, mentions that our available evidence is very limited so that "there will be endless languages consistent with all the actual utterances of a speaker none of which is 'the' language the speaker is speaking" (1992, p. 110). From the side of the interpreter, the problem is put as follows: "If you (the interpreter) do not know how a speaker is going to go on, you do not know what language she speaks, no matter how much she has said up until now" (1992, p. 110). Davidson, however, takes this to be the problem KW's sceptic proposes: "This difficulty, though it may have troubled Wittgenstein, and certainly troubled Kripke, seems to me to have a relatively simple answer" (1992, p. 111). As put by Davidson, the problem he attributes to KW's sceptic is basically similar to the problem of the indeterminacy of interpretation. Interestingly, at this stage, Davidson speaks in a way that reminds us of his measurement scales analogy:

> I regard the existence of empirically equivalent languages (that is, languages equally consistent with all possible empirical evidence) as no more threatening to the reality or objectivity of the correct interpretation of utterances and their accompanying mental states than the existence of various scales for recording temperatures or lengths is to the reality or objectivity of temperature or length. (1992, p. 110, fn. 3)

[82] As discussed in Chapter Three, Davidson considers the resulting indeterminacy of translation/interpretation to be unimportant if enough empirical and formal constraints are imposed on the process of translation/interpretation. Here again, Davidson appeals to the same sort of considerations, which, it will be argued, fail to provide an answer to the sceptical problem KW's sceptic has proposed.

Davidson has argued that a speaker, to mean something by her words, does not need to go on as others do; rather, all that is required is that how she does go on be understandable or interpretable by another speaker.[83] Davidson, for the problem he attributed to KW in the above passage, offers the following solution:

> The longer we interpret a speaker with apparent success as speaking a particular language, the greater our legitimate confidence that the speaker is speaking that language—that is, that she will continue to be interpretable as speaking that language. Our strengthening expectations are as well founded as our evidence and ordinary induction make them. … To the extent that we are right about what is in someone's head, and therefore are right about what he would mean by endless things he does not say, we are right about 'the' language he speaks. (1992, pp. 110-111)

Hence, for Davidson, in order to know 'the' language the speaker speaks, we need to know what she means and believes and, to know the latter, the speaker needs to be interpreted, that is, to be in linguistic communication with others. As discussed in Chapter Two (Section 2.2.1), successful communication, according to Davidson, essentially involves the speaker's intention to be interpreted in a particular way, since, among the variety of interpretations, the *correct* ones are those the speaker intended the interpreter to reach: "An interpreter (correctly) interprets an utterance of a speaker only if he knows that the speaker intends the interpreter to assign certain truth conditions to his (the speaker's) utterance" (1992, pp. 111-112).[84] Thus, what governs the correct application of a word for the speaker is the intention of the speaker to be interpreted in a particular way and his being interpreted in that way: "Intentions depend on the belief that one can do what one intends, and this requires that one believe nothing will prevent the intended action. Thus intention would seem to have just the properties needed to make sense of the idea that a speaker has failed to go on as before" (1992, p. 112). This means that "being interpreted as intended" is proposed by Davidson as the alternative solution which supposedly defuses the sceptical paradox: "Essentially these points about

[83] As we saw in Chapter Two, Davidson believes that the concepts like the concept of a language, rule, convention, and the like, are nothing but theoretical concepts employed by philosophers to theorize about speakers' successful linguistic practices. Hence, we do not really need to attribute to a speaker an actual "language", as the Common View defines it, that is, a language as a set of fixed rules; rather, it is enough to explain when speakers can understand each other. And, now, the problem is: how do they do that? How can successful interpretation happen?

[84] As he clarifies this point again, "meaning … gets its life from those situations in which someone intends (or assumes or expects) that his words will be understood in a certain way, and they are" (1994, p. 120). How a speaker "intended to be understood, and was understood, is what he, and his words, literally meant on that occasion" (1994, p. 120).

intention are made by Crispin Wright in attempting, like me, to defuse Kripke's view that he has extracted an essentially insoluble 'skeptical paradox' from Wittgenstein's treatment of meaning" (1992, p. 112, fn. 6). We will investigate, in the next chapter, Wright's response to KW's sceptical argument and the claim that Davidson's alternative solution is somehow inspired by the way Wright deals with the sceptical problem.

(a) Problems

What are the problems with Davidson's response to KW's sceptical argument and his treatment of KW's sceptical solution as introduced above? Briefly speaking, there are two. (I) Davidson assimilates KW's sceptical problem to the problem of the indeterminacy of interpretation. However, since, as discussed in Chapter Three (Section 3.2), he takes the indeterminacy problem to have epistemological outcomes only, he proposes the process of interpretation and induction as the solution to the problem. If that is true, then it is clear that his alternative solution to KW's sceptical problem faces the same sort of problem that his reading of Quine's indeterminacy of translation arguments faced. (II) Because of the same reason, that is, missing the sceptical part of KW's argument and solution, Davidson's alternative solution is still susceptible to KW's sceptical problem: KW's sceptic challenges explaining the practice of meaning something by an utterance in terms of any sort of fact about speakers, including, of course, the facts about how they intend their utterances to be interpreted.

First of all, how is interpretation, or more particularly, "being interpreted as intended", supposed to solve KW's sceptic's sceptical problem? Davidson's claim seems to be that if a speaker, S, means something by a word, say *table* by "table", then there is a fact, i.e. S's intention to be interpreted as speaking about tables via her use of "table", which determines the correct application of the word for S. Suppose that S, in the presence of a table, utters "table". The question the sceptic asks is whether S means *table* by "table"? The sceptic's answer was negative because there is nothing about S constituting the fact that S means *table* rather than something else by her utterance of "table". Davidson's answer, however, is affirmative, because, supposedly, the speaker intends the interpreter to interpret her utterance in a particular way, say, as meaning *table*. The interpreter's job is to successfully keep track of this fact about the speaker. Davidson thinks that the interpreter will be successful in doing so by engaging in the process of interpretation, induction, and evidence-collecting: the interpreter will successfully interpret the speaker

in the way the speaker intended to be interpreted if the speaker, intentionally or unintentionally, provides enough evidence and clues for such an intended interpretation. However, we can take KW's sceptic to be asking again: How can such a fact about S fix the (meaning) fact that S means *table* by "table" rather than *tabair*? What is it that makes it the case that S intends to be interpreted in one way rather than another? Metaphorically put, the intention of the speaker to be interpreted in a particular way, according to KW's sceptic, is nothing but a sign, which is in need of interpretation. The speaker can be taken as intending to mean *tabair*, and, for all that Davidson has so far said, there is nothing about S ruling out such sceptical hypotheses. In other words, suppose that the speaker utters, or assents to, the sentence "That's a table" where there is a table in view, and suppose that she intends her utterance of this sentence to be interpreted as meaning that *that's a table*. The sceptic, however, asks: What is it that makes it the case that the speaker really intended to be interpreted as meaning *that's a table* rather than *that's a tabair*, when something is a tabair if it is a chair anywhere else or a table in this place, say, in the speaker's office? The speaker's assent to "That's a table" is compatible with taking her to be intending to be interpreted as meaning *that's a tabair*. Similarly, her application of the world "table" to a table outside her office is correct *if* she intended her utterance to be interpreted as meaning *table*, and it is incorrect *if* she intended it to be interpreted as meaning *tabair*. What makes it the case that it is the speaker's intention to be interpreted as meaning *table* rather than *tabair* that determines the correct application of "table" for him?

Davidson's response seems to be that: it is a fact about the speaker that she has an intention to be interpreted in a particular way (it is a condition on being a speaker), and insofar as the formal and empirical constraints on the process of interpretation narrow down the range of interpretations to an acceptable level, we can take the alternative interpretations as equally good interpretations of the speaker's behaviour helping us to understand the speaker's behaviour as intelligible. However, as we saw in the case of Davidson's reading of Quine's indeterminacy of translation thesis, this response, as it stands, cannot solve the problem of meaning scepticism. The problem is that the interpretation that the speaker intends to mean *table* by "table" and the interpretation that the speaker intends to mean *tabair* by "table" are incompatible interpretations, which are compatible with all available evidence and clues the interpreter may be able to appeal to. What if the speaker really intends to be interpreted in a different way, that is, as meaning *tabair* by "table"? Any of the sceptic's sceptical hypotheses can be taken

to be the way the speaker intended her utterance to be interpreted. Such facts about the speaker cannot fix any meaning fact as to what she means by her words. Any use the speaker makes of her words is compatible with her intending to be interpreted in one way rather than another. And, thereby, KW's sceptic leads us to his sceptical paradox again.

While this response from Davidson fails to resist KW's sceptic's sceptical problem, Davidson might alternatively claim that: the speaker intending to be interpreted in a particular way is itself an irreducible fact. In other words, in each case of communication, and utterance by utterance, the speaker uniquely intends to be interpreted in a particular way, that is, to be interpreted as meaning something specific by her word, and, if she is interpreted as such, she actually means that thing by her word. These facts are to be taken as primitive, *sui generis*, facts about the speaker. We do not need to introduce a more basic fact about the speaker constituting the fact that she intends to mean one thing rather than another by her words. In this case, what was apparently taken to be a threatening, sceptical problem about meaning can now be treated as a non-threatening, epistemological consequence of the fact that there are always different ways of capturing such irreducible facts about the speaker.[85] As we saw in Chapter One (Section 1.2.9), this line of response is briefly discussed by KW's sceptic, where he considers a non-reductionist response to the sceptical problem. The main argument from the sceptic against this sort of response was the argument from queerness: taking the fact that the speaker means *table* by "table" to be an irreducible, primitive fact about the speaker will lead to the problem of explaining what such a fact is and how we know about it. What sort of state is this state, which is supposedly present in a potentially infinite number of new cases of use? If this state is not supposed to be introspectible, then how do we know it?

Indeed, defending such a non-reductionist response to the sceptical problem seems to be the best option for Davidson if he wishes to preserve his alternative solution based on mutual interpretation. For his alternative solution brings in the facts about the speaker's intention to be interpreted in a particular way and the speaker's being interpreted as such as what provide the criterion for assessing the speaker's success in the practice of meaning something by an utterance. Now, if such facts were to be constituted by some

[85] As a consequence of such indeterminacy with epistemological outcomes, there always remain different, but acceptable, interpretations of what the speaker intends to mean by her words, all of which correctly interpret the speaker, just as there are always different scales of measuring the temperature of an object.

other facts about the speaker, then it would be hard to find out what facts about the speaker, in a way immune to KW's sceptical problem, could constitute the fact that the speaker intends to be interpreted in a particular way. The sceptical problem would arise in the case of any fact cited to constitute the fact that the speaker intends to mean one thing rather than another. Because of this reason, the best way for Davidson to resist the sceptical problem is to support a non-reductionist response to KW's sceptical problem: whenever the speaker means something by an utterance, she intends her utterance to be understood in a particular way, and this is something that she noninferentially knows about herself without there being anything mysterious about such a state and the speaker's knowledge of this state. However, if this is the way Davidson would intend to respond to KW's sceptic's argument from queerness, then he would have to provide us with a plausible explanation of the speaker's first-person authority over her beliefs and intentions. Davidson has attempted to provide such an explanation which we will discuss in Chapter Five.

4.1.3. The Third Place: The "Cause"-Determination

In the last section above, we reached Davidson's conclusion that what makes it the case that the speaker means *table* by "table" is the fact that she intends to be understood in that way and she is understood as such. However, Davidson also says: "The view I have just sketched deals only with interpretation, and so presupposes a social environment rather than providing an argument for it" (1992, p. 112). Having taken for granted that "being interpreted as intended" is the solution to KW's sceptic's sceptical problem, Davidson now wants to argue that success in actual mutual interpretation is both necessary and sufficient for explaining the speaker's success in meaning something specific by her words. Here Davidson relies again on the notion of triangulation. It is possible to relate Davidson's alternative solution given in terms of mutual interpretation to his discussion of the notion of triangulation and, thereby, his remarks on the cause-determination process, i.e. the process of determining the actual cause of a speaker's utterances. As discussed in Chapter Two (Section 2.2.5), Davidson is an externalist about meaning, according to which "what a person's words mean depends in the most basic cases on the kinds of objects and events that have caused the person to hold the

words to be applicable" (1987, p. 37).[86] Also, recall that, according to Davidson, the interpreter understands what the speaker means by "table" if he interprets the speaker in the way the speaker intended her utterance to be understood. If that is true, then we can take Davidson to be claiming that the interpreter understands what the speaker means by her words if he, at least, successfully determines to what object the speaker intends to apply her words, i.e. if he finds out the intended object to which the speaker applies her words. Generally speaking, hence, the "actual external cause" of the speaker's utterance must be determined. However, how can the interpreter, or even the speaker, determine the actual cause of the speaker's utterance?

Davidson thinks that employing the notion of triangulation provides the answer to the above question. In his discussion of this notion, he again starts with advancing a problem for an individual speaker considered in isolation. Davidson can be seen to be in agreement with KW about the fundamental problem with the solitary speaker: a solitary speaker cannot be said to mean anything at all, especially if "it is, in the simplest cases, what causes a belief that gives it its content" (Davidson, 1999c, p. 129). Though we have discussed this issue in detail in Chapter Two (Section 2.2.5), we can briefly again put Davidson's reason for his agreement with KW as follows: the actual cause of the speaker's responses should be determinate if her beliefs and utterances are to have any content, or meaning, at all, since fixing the content of thoughts and the meaning of utterances depend, based on Davidson's externalism, on the external causes of the thoughts and utterances. The problem, as Davidson concedes, is that the actual cause of the solitary speaker is highly ambiguous: "Cause is doubly indeterminate: with respect to width, and with respect to distance" (1999c, p. 129). For a solitary speaker, it is not determinate (I) whether it is a proximal cause, e.g. some stimulation on the solitary speaker's skin, or an external cause, e.g. an object a certain distance from her in the external world, that actually causes the speaker's thoughts and utterances (this is the indeterminacy with respect to distance), and (II) whether it is one aspect of the cause rather than another that causes the speaker's thoughts and utterances (this is the indeterminacy with respect to width). However, any such cause-determination requires

[86] As Davidson puts it, "concepts, and the sentences and thoughts that employ them, are in part individuated by their causal relations to the world and in part by their relations to each other." (2001d, p. 138); "in the simplest and most basic cases, words and sentences derive their meaning from the objects and circumstances in whose presence they were learned. A sentence which one has been conditioned by the learning process to be caused to hold true by the presence of fires will (usually) be true when there is a fire present" (1988, pp. 44-45); "the stimuli that cause our most basic verbal responses also determine what those verbal responses mean, and the content of the beliefs that accompany them" (1991a, p. 213).

the speaker to possess the concept of objectivity, since Davidson, by appealing to Wittgenstein's considerations on private language, argued that, for a solitary person, what seems right to her would be right: there would be no way to draw the distinction "between thinking one means something and actually meaning it" (1992, p. 117). In this sense, whatever the solitary speaker takes to be the actual cause of her responses will always be the actual cause of her responses, no matter what it is. Thus, as Davidson indicates, his "argument shows that one's first language cannot be a private language, that is, a language understood by only one creature, and to this extent it is in agreement with Kripke's Wittgenstein" (1992, p. 121). Davidson concludes that the existence of a social setting is thereby necessary if there is to be any meaning for utterances and any content for thoughts and, to this extent, he is in agreement with KW. His disagreement with KW's sceptical solution emerges when, according to Davidson, KW brings in the requirement of speaking as others do. However, we will see again that the problem with Davidson's grasp of KW's sceptical argument and solution is more than this: the problem is his failure to take into account the vital *sceptical* part of KW's sceptical argument and solution. Surprisingly, Davidson, in a footnote in his paper "The Second Person" (1992), confesses: "I have ignored a very important aspect of Kripke's discussion, his claim that Wittgenstein's 'solution' to the problem of meaning is 'skeptical'" (1992, p. 113, fn. 7). KW's argument and solution's being "sceptical", however, is the most important feature of the sceptical argument and solution.

For Davidson, the role of the second person, via his interaction with the first person, is to help the first person, (I), to locate and fix the actual causes of her responses and, (II), to grasp the concept of objectivity, which is required for cause-, or content-, determination. As discussed in Chapter Two (Section 2.2.5), "primitive" or "prelinguistic" triangulation is put forward by Davidson to show how such an interaction, i.e. the correlation between two creatures' similar responses to a certain stimulus in the world, can help them to locate the actual cause of their responses and to make sense of the concept of objectivity, especially when the correlation between their similar responses is broken. Briefly speaking, if creatures which are similar in their innate responses (similar in grouping things together in the world) can correlate their similar responses to certain external causes in the world, then: (I) we can take the "actual" cause of their responses to be the "common" cause of their responses, highlighted by triangulation, and (II) the situations, in which such correlations break, provide them with the opportunity to perceive the dissimilarity between their responses.

These situations, hence, allow them to make sense of the distinction between what merely seems right to them and what is actually right independently of what each of them individually thinks, thereby enabling them to grasp the concept of objectivity or objective truth.[87]

However, Davidson thought that this prelinguistic triangulation is possible between animals with no language and, to this extent, although it is necessary, it is not sufficient for creatures to have thoughts with determinate content and utterances with determinate meaning.[88] According to Davidson, it is only engaging in linguistic triangulation that is both necessary and sufficient for creatures to fix the content of their thoughts and to grasp the concept of objectivity. As we concluded in Chapter Two (Section 2.2.5), what Davidson means by "adding language to primitive triangulation" is indeed nothing but bringing us back to the process of mutual interpretation, that is, the requirement that the triangulators must interpret, and be interpreted by, each other. Thereby, what Davidson eventually comes up with is again interpretation and induction as the solution to the

[87] Again, the point of this sort of primitive triangulation is that, through such triangulating "each creature learns to correlate the reactions of other creatures with changes or objects in the world to which it also reacts" (1999c, p. 128). As we saw in our discussion of Davidson's argument from triangulation in Chapter Two, prelinguistic triangulation could be established between two or more non-linguistic animals or between a teacher, with a language, and a child, with no language. For instance, in the latter case, the child has no meaningful responses; rather, the meaning is bestowed to the child's words and thoughts; as Davidson's remarks on ostensive learning clarify: "At the start, there could be no point in the learner questioning the correctness of the teacher's ostensions. The learner may or may not be learning how others in some linguistic community speak, but the learner can discover this only later. In the private lesson, a meaning is being bestowed on words quite apart from any use those words may have at other times and with other people. If we think of ostension only as the teaching of a socially viable meaning we miss the essential lesson, which is that for the learner ostension is not learning something already there. The learner is in at a meaning baptism" (1997b, p. 140). On the side of the learner, there is no such thing as meaning, concepts, causes, and the like. Hence, whether in the case of two interacting animals or the case of two interacting human beings, generally speaking, the prelinguistic, or primitive triangulation, is taken to be a very basic situation, in which two creatures responding to the world in a similar way respond to the same object, or stimulus, in their environment, as well as to each other's responses to that stimulus. They are observing each other's responses, or say, correlating their responses, to that specific object they are both responding to. As mentioned before, this triangulation is not a special situation only human beings can be in: "One sees this in its simplest in a school of fish, where each fish reacts almost instantaneously to the motion of the others" (1999c, p. 128). Because of that, engaging in primitive triangulation cannot be taken to be a sufficient condition for having thoughts and utterances with determinate content.

[88] Whether or not Davidson's remarks on triangulation successfully establish the claim that primitive triangulation is *necessary* for thoughts and a language to emerge is a matter of controversy. For discussion of this issue, see (Verheggen, 1997, 2006, 2007, 2016), (Glüer, 2006, 2011), (Talmage, 1997), and (Ludwig & Lepore, 2005, especially section 22).

problem of meaning-determination. The creatures in linguistic triangulation "must each be an interpreter of the other" (1992, pp. 121).

In the previous section, we concluded that interpretation, or more particularly, "being interpreted as intended", is Davidson's solution to KW's sceptical problem: the speaker means something specific by an utterance if she intends her utterance to be interpreted in a particular way and the interpreter successfully interprets the speaker's utterance as such through the process of interpretation and induction. Davidson, by means of his discussion of the notion of triangulation, has indeed added a causal explanation of such a process: the interpreter understands what the speaker means by her word if he determines to what object the speaker intends to apply her word, i.e. if he can fix the actual cause of the speaker's utterance, since, according to Davidson's externalism, it is the actual distal causes that partly determine the content of the speaker's utterances and thoughts. Davidson tried to show that two creatures can fix the content of their responses, i.e. the meaning of their utterances and the content of their thoughts, only via actual linguistic triangulation. In this sense, it is success in actual mutual interpretation that is both necessary and sufficient for the creatures to have meaningful responses, not the requirement of meaning the same thing by the same words.

(a) Problems

However, does Davidson's use of the notion of triangulation add anything new and constructive to his original response to KW's sceptical problem? Can Davidson's causal explanation, as presented in his discussion of the notion of triangulation, deal with KW's sceptical problem? The difficulty with Davidson's explanation is that KW's sceptic is indeed happy to take for granted any sort of fact one might cite to justify the claim that the speaker means one thing rather than another by her word, including Davidson's causal story about reference-, or cause-, determination. In other words, the sceptic can run his sceptical problem again even by granting Davidson's causal explanation of meaning-determination. For instance, suppose that a particular thing in the world, e.g. a specific aspect of the table in view, causes the speaker to utter "table". Having agreed on that, KW's sceptic would ask: What concept does the speaker express by her utterance of "table"? Suppose that it is that table in front of the speaker that caused her utterance of "table". What does the speaker's utterance of "table" mean, *table* or *tabair*? The sceptic would claim that the same object, or aspect of the object,

has caused the speaker to utter "table" not to mean *table*, but to mean *tabair*, where a tabair is a table up until time *t*, and a chair after *t*. In this sense, the sceptic would claim that the object caused the speaker to express the concept of a tabair, rather than that of a table, by her utterance of "table". The totality of the causal facts adduced by Davidson, hence, are consistent with both the *table* and *tabair* hypotheses. No matter what caused the speaker to utter "table", according to the sceptic, the speaker's application of the word can be considered as correct as well as incorrect dependent on what she means by the word: *if* the speaker means *tabair* by "table", applying "table" to a chair after *t* is correct, while *if* she means *table* by "table", such an application is incorrect. The sceptic concerns: What is it that makes it the case that the speaker means *table*, and not *tabair*, by her word?

Davidson's alternative, as this time put in the metaphor of triangulation, hence, cannot solve the sceptical problem, although Davidson apparently thinks it can:

> Kripke depends on the second person, or a community, to embody a routine which the speaker can share. In contrast, the argument I have outlined does not require (though of course it allows) a shared routine, but it does depend on the interaction of at least two speaker-interpreters, for if I am right, there would be no saying what a speaker was talking or thinking about, no basis for claiming he could locate objects in an objective space and time, without interaction with a second person. (1992, p. 121)

Davidson's suggestion, however, could not resist KW's sceptic's sceptical problem and nothing appears in Davidson's discussion of the notion of triangulation that helps to solve this problem. Both linguistic triangulation and his earlier remarks on mutual interpretation, i.e. "being interpreted as intended", lead us to the conclusion that he takes KW's sceptical problem as advancing an epistemological problem, rather than a metaphysical/sceptical one: he thinks that the problem will be solved through the process of interpretation. Through linguistic triangulation, each creature can grasp the concept of objectivity, having which is necessary for a creature to come up with other interrelated concepts, such as that of a table: "It is this triangular nexus of causal relations … that supplies the conditions necessary for the concept of truth to have application. Without a second person there is, as Wittgenstein powerfully suggests, no basis for a judgement that a reaction is wrong or, therefore, right" (Davidson, 1997a, p. 83). Without the concept of objectivity, the speaker's thoughts and utterances would not have a determinate, objective content, since whatever the speaker takes to be the cause of her responses, which supposedly gives content to her responses, will be the actual cause of her responses. To acquire the concept of objectivity, and to fix the actual cause

of her responses, the speaker, according to Davidson, must have linguistic interaction with at least another person. Once the actual causes of the speaker's responses are fixed, the concepts the speaker expresses by such linguistic responses are also fixed, that is, the content of her thoughts and the meaning of her utterances.[89] Nonetheless, such linguistic triangulation would not help to solve KW's sceptical problem, that is, whether, in virtue of being caused by a certain stimulus in the world to utter a word, the speaker expresses one concept rather than another via her use of that word.

Moreover, we can see that the difference between KW's sceptical solution and Davidson's suggested solution is more than just "the flavor [Davidson's solution] gives the social aspect of language" (1992, p. 121). Rather, it is the difference between a sceptical solution and a straight solution. KW's *sceptical* solution, as discussed in Chapter One (Section 1.5.2), does not allow for any previously grasped concepts, rules, and the like, to play a role in the explanation of the practice of meaning something by words. Doing so, according to KW, would bring us back to the Radical Sceptical Conclusion of the sceptical argument. However, this is exactly what Davidson is suggesting: to explain why two speakers speaking different languages can understand each other, Davidson appeals to the previously fixed and interrelated concepts that they have come to possess through their linguistic triangulation. Similarly, Davidson, in his criticisms of KW's sceptical solution, points out that his "triangulation depends not only on a plurality of creatures, but equally on shared external promptings" (2001d, p, 143), and such shared external promptings are absent in KW's solution: "Kripke's examples ... concern mathematical examples, and so lack the shared stimulus to provide the possibility of a shared content" (2001d, p. 144). And, thereby, "Kripke's account of what he takes to be Wittgenstein's 'skeptical solution' to the puzzle of rule-following is

[89] Davidson emphasizes that "our first guess as to what is meant by a perception sentence will be a shot in the dark, but given how much alike people are, getting it right is often like hitting a barn door; the most casual guess is often correct. The simplicity of this mode of entry into an alien language should not leave us thinking that a concept so identified is defined by its external causes without the aid of theory or a supporting nexus of further concepts. A concept is defined by its typical cause only within the framework of a system of concepts that allows us to respond to certain stimuli as tables, friends, horses, and flies. A concept is defined for those who speak languages like ours by its typical causes, given that we are already in the world of language and conceptualization (2001d, p. 138). He again states: "To have a concept, in the sense I am giving this word, is, then, to be able to entertain propositional contents: a creature has a concept only if it is able to employ that concept in the context of a judgment. It may seem that one could have the concept of, say, a tree, without being able to think that, or wonder whether, something is a tree, or desire that there be a tree. Such conceptualization would, however, amount to no more than being able to discriminate trees—to act in some specific way in the presence of trees—and this, as I said, is not what I would call having a concept (1995b, p. 9).

inadequate to serve as the whole story about conceptualization" (2001d, p. 143). Nonetheless, as discussed above, sharing such stimuli and fixing the actual causes of the speaker's responses by the triangulators fails to resist KW's sceptic's sceptical problem. Indeed, KW could make a similar criticism of Davidson's use of the notion of triangulation: How is triangulation supposed to help us fix the content of our utterances and thoughts in the case of mathematical examples? In such examples, nothing like a shared stimulus is available, and it is not clear how triangulation can answer the sceptical question: How can triangulation help fixing which mathematical function is denoted by a speaker's use of the "+" sign?

To make the problem with Davidson's reading of KW clearer, we can consider Davidson's claim that the problem we needed to solve was "how to account for failure to apply a concept correctly, given that what one person might count as an error may just be another person applying a different concept" (2001d, p. 143). For Davidson, KW's solution is that "if a learner fails to apply a concept (or word) as his teacher would, the learner has made a mistake" (2001d, pp. 143-144). Davidson argued that this suggestion (i.e. doing what others do) is neither necessary nor sufficient for explaining success in the practice of meaning something by an utterance, or that of expressing a concept correctly. Davidson's alternative solution was linguistic triangulation, through which the triangulators could fix the actual cause of their responses and, thereby, the concepts they express by their utterances. It can be done, according to Davidson, via the triangulators' grasp of the concept of truth or objectivity. Triangulation, in other words, presupposes the triangulators' grasp of the concept of truth or objectivity if they are to be capable of having responses with determinate content. As Davidson said, "to understand the speech of another, I must be able to think of the same things she does; … we must entertain the same propositions, with the same subject matter, and the *same concept of truth*" (1982, p. 105, emphasis added). However, KW's sceptic is asking: What fact about the speakers can make it the case that they have grasped one concept rather than another, that they express one concept rather than another by their utterances, or entertain one proposition rather than another, given that their utterances are indeed caused by specific objects or events in the world? According to KW's sceptic, no fact about a speaker, or a group of speakers, can make it the case that they have grasped one concept, such as the concept of truth, rather than another, that they express one concept, such as the concept of a table, rather than another by their utterances. Presupposing the fact that the triangulators can grasp the concept of truth

begs the question against the sceptic arguing that there is no fact of the matter about which concept a speaker grasps, that there is no fact of the matter about correctness or incorrectness at all. This claim is exactly what the sceptic has challenged. As Kripke emphasizes, "for Wittgenstein, an 'explanation' of this kind ignores his treatment of the sceptical paradox and its solution" (1982, p. 97).[90] There is no fact as to whether a speaker's use of a word is correct or incorrect, right or wrong, or true or false, according to KW's sceptic.

4.2. Mark Joseph on Davidson's Alternative Solution

Before summing up this chapter, it is worth reviewing the difficulties with Davidson's grasp of KW's sceptical argument and solution by considering Mark Joseph's characterization of KW's sceptical problem and Davidson's response to it. Joseph characterizes KW's sceptical problem as follows:

Suppose that before 1st of January 2001, Bert had used the word "arthritis" only to describe disease of the joints, but after that date he speaks of arthritic pains in his joints *and* muscles. Joseph asks: "Should we say (a) that Bert associates *one* concept with the word "arthritis" in the twentieth century, and later he associates a *different* concept with it in the twenty-first? … Or should we say (b) that his concept has *not* changed, where that concept is (and was) defined as follows:

- If *t* is before 1 January 2001, then *o* satisfies "*x* suffers from arthritis" at *t* if and only if *o* suffers from a disease of the joints at *t*;

[90] Indeed, KW's sceptic rejects Davidson's claim that "not only can others often learn what we think by noting the causal dependencies that give our thoughts their content, but the very possibility of thought demands shared standards of truth and objectivity" (Davidson, 1988, p. 52). As KW says, "on Wittgenstein's conception, a certain type of traditional – and overwhelmingly natural – explanation of our shared form of life is excluded. We cannot say that we all respond as we do to '68+ 57' *because* we all grasp the concept of addition in the same way, that we share common responses to particular addition problems *because* we share a common concept of addition. … For Wittgenstein, an 'explanation' of this kind ignores his treatment of the sceptical paradox and its solution. There is no objective fact – that we all mean addition by '+', or even that a given individual does – that explains our agreement in particular cases. Rather our license to say of each other that we mean addition by '+' is part of a 'language game' that sustains itself only because of the brute fact that we generally agree" (1982, p. 97). Kripke emphasizes that the success in our linguistic practices "cannot be explained by 'the fact that we all grasp the same concepts'" (1982, p. 109).

- If *t* is on or after 1 January 2001, then *o* satisfies "*x* suffers from arthritis" at *t* if and only if *o* suffers from a disease of the joints or ligaments at *t*" (2004, p. 94).

To use a different example, we can consider the case of a speaker, S, who now applies "triangle" to the things with the shape of □. The question is whether her application of the word is correct. The sceptic's answer is positive, since, according to him, what the speaker means by "triangle" is not *triangle*, but *quadangle*, where a quandangle is defined as follows:

- If $t < t^*$, then *x* satisfies "*x* is a triangle" at *t* if and only if *x* has the shape of Δ;
- If $t \geq t^*$, then *x* satisfies "*x* is a triangle" at *t* if and only if *x* has the shape of □.

KW's sceptic argued that no fact about S can determine what she means by "triangle", that is, what concept she is expressing by uttering "triangle". Any change in S's use of the word is merely apparent: S continues to use the same word with the same (odd) meaning. No matter how S uses "triangle", "we could always devise a further, odder concept to accommodate *that* use" (Joseph, 2004, p. 94). Therefore, as Joseph concludes, "a correct use of the word "arthritis" cannot be determined by a concept Bert intends it to express, since *any* use he makes of it can be understood to accord with some concept we can interpret him to have intended" (2004, p. 95).

According to Joseph, however, "Davidson resists this sceptical conclusion by rejecting the assumption that drives it; namely, that languages are systems of norms codified by a community's conventions" (2004, p. 95). As we saw in our discussion of Davidson's negative argument against Conventionalism in Chapter Two (Section 2.1.2), the assumption that Davidson rejects is actually the Common View's conception of a language as a rigid set of rules or conventions. According to Davidson's alternative view, a word means something only if it is uttered by a speaker intending her utterance to be interpreted in a certain way and it is interpreted in that way. As a result, before interpretation takes place, "triangle" has no meaning: we cannot say whether it means *triangle*, *quadangle*, or anything else. According to Joseph, Davidson claims that the nature of meaning *per se* is not related to what others do, rather it is identified with the invariant pattern a speaker's situated utterances and other actions make. We saw, in Chapter Three (Section 3.2), that such an invariant pattern, for Davidson, is the empirical relation between the speaker (his beliefs, intentions, and the like), the world (the external causes in the world), and the sentence the speaker chooses to utter (say,

what the cause prompts the speaker to utter). If the speaker is supposed to exemplify such an invariant pattern, "we can, everything else being equal, attribute to him either concept [*triangle* or *quadangle*] in pursuit of making the best overall sense of his linguistic and non-linguistic behaviour" (Joseph, 2004, p. 95). For instance, suppose that S applies "triangle" to triangles at time *t*. The interpreter forms the hypothesis that the speaker means *triangle* by "triangle". But, there are other candidates too, such as the one KW's sceptic suggested, i.e. that S means *quadangle* by "triangle". The interpreter's job is to confirm one of these hypotheses by using all available evidence. Although, according to Joseph, it is always possible to attribute a conceptual mistake to the speaker, which concept we attribute to the speaker as being expressed by his utterance "will entail making compensatory adjustments elsewhere to our theory of interpretation" (2004, p. 95).

This, however, leads us back to the problems with Davidson's reading of Quine's indeterminacy of translation thesis and his treatment of KW's sceptical argument and solution. For Davidson, if we come up with two theories of interpretation, one of which attributes the concept of *triangle* to the speaker and the other attributes the concept of *quadangle*, then we can take both as correctly describing the speaker's behaviour as intelligible, insofar as they are both equally compatible with all available evidence and respect the constraints Davidson has put on the process of interpretation. However, in our discussion of Davidson's reading of Quine's sceptical arguments in Chapter Three (Section 3.3), we concluded that Davidson's suggestion neglects the distinction between the epistemological problem of the underdetermination of theory and the metaphysical problem of the indeterminacy of translation. It is from such a perspective that Davidson sees KW's sceptical problem to have a relatively easy solution: interpretation and induction. What we tried to show was that Davidson's approach to the problem, both in the case of Quine and in the case of KW, is implausible. There is no success in interpretation at all without first providing an answer to KW's sceptic's question: what makes it the case that the speaker intends to be interpreted in one way rather than another? Without an answer to this question, mutual interpretation and induction can provide no response to KW's sceptic.[91]

[91] Stroud (1998) believes that Davidson and KW are both arguing that presupposing a certain conception of meaning leads to unacceptable conclusions about the practice of meaning something by a word; once we give up on such a misunderstood view, the implausible conclusions will disappear. As he says, "it is the very idea of something in the mind that instructs or guides a person's meaning or understanding what he does that Wittgenstein exposes as a widespread philosophical illusion. But without that idea, the

4.3. Conclusion

In this chapter, we first tried to clarify the way Davidson interprets KW's sceptical argument and solution. We saw that the same general treatment that Davidson offered in the case of Quine's indeterminacy of translation arguments is offered again in the case of KW's sceptic's sceptical argument. In both cases, Davidson attempts to consider the arguments as arguments with essentially epistemological consequences so that his alternative solution, i.e. that which proceeds via interpretation and induction, can provide a solution to the problems they propose. However, as in the case of Quine, his solution fails in the case of KW: unless it is argued that there are meaning facts, which empirically equivalent translation manuals, or interpretations, supposedly capture, the indeterminacy of interpretation will lead to nothing less than radical scepticism about meaning. KW's sceptic challenges any sort of view which attempts to explain success in the practice of meaning something by an utterance in terms of some fact about the speaker, such as her intention to be interpreted in a particular way. No such facts about the speaker can determine what she means by her utterance, according to KW's sceptic.

Moreover, we saw that Davidson's objection that KW's sceptic's sceptical argument is essentially dependent on the rule-following picture of meaning was misplaced, since KW's sceptic could run the second strand of his sceptical argument without appealing to the rule-following picture of meaning, rather on the basis of the presupposition that there are meaning facts determining the correct application of words for the speaker. Similarly, we saw that KW's sceptic's sceptical problem does not necessarily concern the linguistic practices of an individual speaker; rather, it could well function in the case of two or more individual speakers interacting with each other *if* the success in their linguistic practices is supposed to be explained in terms of some facts about them, such

---

sceptical or paradoxical conclusion cannot be reached in the way Kripke reaches it. So what Kripke actually establishes in the name of Wittgenstein is *equivalent* to what can be put in a way that brings it closer to the text from Davidson with which I began: "There is no such thing as a person's meaning something, not if meaning is anything like what many philosophers have supposed"" (1998, pp. 27-28, emphasis added). I am not, however, clear about the details of Stroud's claim. Indeed, the way we characterized Davidson's and KW's arguments in Chapter One and Two might point to a similar idea: Davidson argues against the "Common View" as what results in the implausible Radical Sceptical Conclusion about meaning, as the *first* strand of KW's sceptical argument argues against a version of Traditional Realism which may be treated as compatible with the "Common View". Nonetheless, as we argued throughout this chapter, this similarity is just superficial: Davidson seems to misunderstand KW's sceptic's *sceptical* argument and solution.

as their shared grasp of some concepts. Finally, KW's sceptical solution, contrary to Davidson's grasp of it, does not appeal to, and does not try to accommodate, the presumption that the members of a speech community are already equipped with, or grasp, certain concepts, such as that of truth or objectivity, *because* such an attempt has nothing but the Radical Sceptical Conclusion as its consequence. If the speakers use a word correctly *because* they express one concept rather than another by their utterance, then what is it that makes it the case that they express one concept rather than another by their utterance?

However, one important question remains: Are there still resources in Davidson's writings supporting a plausible response to KW's sceptic's sceptical problem? One way to resist the sceptical problem is to give up on any attempt to reduce meaning facts to any other facts about the speaker and, instead, to insist that meaning facts are primitive, non-reducible facts about the speaker. In this case, Davidson could claim that the fact that the speaker intends her utterance to be interpreted in a particular way, i.e. as meaning *table*, is itself an irreducible, primitive fact about her and the interpreter's job is to keep track of this fact about the speaker. In the next chapter, we will look for the potential resources in Davidson's works supporting such a non-reductionist response to KW's sceptic.

5. Chapter 5: Davidson's Non-Reductionism

Introduction

In this chapter, we will investigate whether Davidson can provide us with a plausible response to KW's sceptic's sceptical problem. Davidson responded to KW's sceptical problem by appealing to the process of interpretation and intention. According to Davidson, a speaker, in each case of communication, intends her utterance to be interpreted in a particular way and if she is interpreted as such, she can be said to successfully mean something specific by her utterance. For Davidson, having such intentions is a necessary condition on being a speaker and is essential to the success in the practice of meaning something by an utterance: "If communication succeeds, there must be these intentions on the part of the speaker, and therefore *if* successful communication is essential to meaning, these intentions are essential to meaning" (Davidson, 1992, p. 112). However, KW's sceptic argued that, if the speaker's intention to be interpreted in a particular way is a fact about the speaker, then this fact is compatible with different sceptical hypotheses about what she means by her word. Hence, there is no fact of the matter as to how the speaker intends to be interpreted and, consequently, there is no successful, or correct, interpretation at all. In response, Davidson could take a non-reductionist approach to the problem: the speaker's intending her utterance to be interpreted in a particular way is itself a primitive irreducible fact, not explainable in terms of any other facts about the speaker.

We first consider the most recent claim made by Claudine Verheggen (Verheggen & Myers, 2016) about whether Davidson can be taken to be supporting a non-reductionist response to KW's sceptic's sceptical problem. Verheggen believes that Davidson's remarks on the notion of triangulation can potentially sustain such a non-reductionist response, though she agrees that Davidson has misunderstood KW's sceptic's sceptical problem. I will argue that Davidson's use of the notion of triangulation presupposes non-reductionism, rather than provides an argument for it. The main concern in advancing such a view is to properly resist KW's sceptic's argument from queerness by providing an adequate account of first-person authority. We will see that Davidson has actually attempted to provide such an account. Wright, however, has criticized Davidson's explanation as deeply problematic. I will argue that Wright's reading of, and his objection to, Davidson's account is inadequate. Davidson's account will be argued

to be problematic for other reasons. Nonetheless, since Davidson seems to be inspired by Wright's remarks on intention, we will consider Wright's own alternative account of first-person authority and his response to KW's sceptic in terms of his judgement-dependent account of intention. This will help us to see how an alternative response to KW's sceptic, different from Davidson's actual one, can be provided. Our final inquiry will concern whether there are potential resources in Davidson's writings supporting such a judgement-dependent account of meaning and intention. We will see that Davidson, in his account of intention, has proposed such a general account, which can be extended to the case of meaning. This means that, putting his actual problematic response to KW's sceptic's sceptical argument to one side, Davidson could resist the sceptical problem by proposing his own judgement-dependent account of meaning and intention.

5.1. Verheggen on Davidson's Non-Reductionism

Verheggen claims that Davidson's considerations on the notion of triangulation can provide us with enough resources to establish a non-reductionist response to KW's sceptic. First of all, she sees the triangulation argument as one single, but complex, argument with the conclusion that "the meanings of one's thoughts and utterances can be fixed only if one has the concept of objectivity, possession of which requires triangulation" (2016, pp. 14-15).[92] Briefly speaking, according to Davidson's perceptual externalism, what partly determines the content of our utterance is the external cause of the utterance in the world. To have any meaningful utterance, hence, the cause of such utterances must be determined. As we saw in Chapter Two and Chapter Four, Davidson believed that, for a solitaire, the causes of his responses are doubly indeterminate, with respect to the location of the causes in the causal chains from the stimulation on the surface of the skin to the Big Bang, and with respect to the specific aspect of the cause, which bestows content on the responses. Even if we can determine that the cause of one's utterance is some external, or distal, object, rather than some proximal cause on

[92] To compare this reading with a different reading of Davidson's argument from triangulation, see (Glüer, 2006, 2011, especially Chapter 5 (Section 5.3)) and (Ludwig & Lepore, 2005, especially Section 22).

the surface of one's skin, we will still have the problem of determining which aspect of the distal cause actually determines the content of one's utterance.[93]

According to Verheggen, primitive triangulation, i.e. the situations in which two creatures similarly respond to a specific stimulus in the world and at the same time respond to each other's similar responses to that stimulus, cannot solve the aspect problem, and not necessarily even the problem of determining the location of the cause.[94] The first person's observation that the second person is similarly responding to a specific stimulus in the world cannot by itself determine to which specific aspect of the cause the second person is responding. Two creatures might produce similar sounds, say "tree", in response to a tree in view; but, these similar responses might be prompted by different aspects of the cause. For Verheggen, primitive triangulation must at best be taken to be showing how deep the problem of content-determination is for externalists like Davidson. As Wittgenstein's considerations on ostensive learning show, "no matter how similar shared responses may seem, they may be responses to different aspects of what cause them, and so they remain ambiguous" (Verheggen & Myers, 2016, pp. 27-28). Therefore, more than just mere interaction with another person, i.e. mere correlation between the individuals' similar responses to an external object in the world, is required. We can add that, hence, more than just mere associations between our responses and the causes in the world are needed in order for meanings to be determined, since, no matter how many times the individuals produce the similar response of uttering "tree", it by itself does not show whether they are both responding to the tree as a whole, to the tree plus the backyard, or to any other aspect of the external cause. According to Verheggen, for the triangulators to solve this problem, they must be able to use triangulation to fix the meaning of their utterances, that is, to fix the specific relation between their responses and the causes in the world. This primarily requires them to categorize, or sort, things in the world in a similar way – for instance, to put trees as a whole in the same class of things, and the backyard in a different class, and so on. Hence, if they can fix the causes of their responses, then they can understand, and

[93] For instance, suppose that we have determined the cause of my utterance of "tree" to be the tree in front of me in the backyard (and not some stimulation on my eyes' receptors). It is still to be determined whether it is the tree itself as a whole, its color, its shape, the tree with the yard, the tree with the soil, etc. that is causing and, thereby, partly determining the content of my utterance.

[94] For Verheggen, we can imagine situations in which the solitary person would be capable of determining whether the cause of his response is distal or proximal. In this sense, primitive triangulation would not be necessary for the solitary person to locate the actual cause of his responses. See (Verheggen & Myers, 2016, pp. 24-26).

successfully communicate with, each other. If you know that when I utter "tree" I apply it to the same thing in the world that you sort as trees, then you understand what I mean by the word even if you are inclined to apply a different word to that thing. The fact that the creatures similarly respond to a cause shows that the cause is taken by them to be the same cause.[95] Hence, if triangulation is supposed to be of any help, it must help the triangulators to fix the relation between their responses and the specific causes in the world.

According to Verheggen, the triangulators can use the triangular situations to solve the aspect problem and fix the content of their responses only when they can be said to possess the concept of objectivity, since such fixation necessarily requires that they both master the distinction between thinking that something is the same and that thing actually being the same, or thinking that something belongs to the extension of one concept and that thing actually belonging to the extension of that concept. The reason, again, is that to have a concept or to conceptualize is to classify objects in certain ways. And, to classify things as such, we, first of all, need to have a conception of misclassification or error: the fact that some things are tables means that some other things are *not* tables, and, hence, taking the latter to be tables is to make a mistake. Indeed, only in this situation can disagreement appear because, without such a conception of error, whatever the speaker takes to be similar would be similar. The triangulators must be in a position to recognize the prompting causes "as the specific aspects of one's environment that cause one's thoughts and utterances" (Verheggen & Myers, 2016, p. 29). In Verheggen's view, the second step of Davidson's argument from triangulation states that, to master this distinction, linguistic triangulation is required. It is so because acquiring and applying the concept of objectivity needs *genuine* disagreement between the triangulators' correlated responses to emerge. The

[95] As an example, suppose that I see a table in front of me and utter "table". My hearer observes my response, and she is seeing, or perceiving, the table in front of me too. Tomorrow, as I stand in front of the same table, I utter "table". My hearer, the second person, observing my response will understand that I am talking about the same thing, which caused my similar response yesterday, *only if* she knows to which specific cause I was and I am responding. The reason is that, if it is not the case, then as it was not clear which aspect of the table caused me to respond as I did yesterday, it would still be ambiguous which aspect of the cause is causing the *similar* response of mine today. This means that it is possible that, yesterday, one aspect of the cause (say, its surface) and today another aspect (say, its leg) causes my similar responses. Therefore, merely observing the similarity between responses is not enough for the content-determination problem to be solved. Similarly, as a result, mere associations between my responses and some objects in the world cannot be taken to be the whole story about meaning-determination.

emergence of such disagreement, however, is no help if the triangulators cannot take advantage of the triangular situations to solve or settle the disagreement. In this situation, it would not be just up to one of them to settle the disputes regarding the external facts. Rather, what this settling requires is discussing, speaking, and talking, that is to say, interpretation.[96]

Nonetheless, the disagreements could not happen and settling the disputes could not be achieved if there was not a background of agreement between the triangulators on many other things. According to Verheggen, to agree or disagree, the triangulators need to agree at least on what their utterances mean to the other. For instance, they should agree on what "cow" means for them, or at least what it means for the speaker, in order to settle the dispute about whether this cow is the same cow they saw yesterday. And, for Verheggen, this is what Davidson means by "linguistic triangulation". Therefore, (1) since acquiring the concept of objectivity essentially requires linguistic triangulation, and (2) since cause-determination requires the possession of the concept of objectivity, and (3) since, without a determinate cause which bestows content on our thoughts and utterances, there would be no language and thought at all, we can conclude that, without linguistic triangulation, there will be no language and thought at all. Therefore, Verheggen believes that Davidson's answer to the problem of meaning-determination can be put as follows: "Since no non-intensional magical trick will do to fix the causes, and hence the meanings, of one's thoughts and utterances, only those producing the thoughts and utterances could achieve this feat. And, they could achieve it only by having the concept of objectivity and triangulating *linguistically* with others" (2016, p. 29). Only those with a language can disambiguate the cause of their responses and thereby fix the meaning of their utterances and the content of their thoughts. Only linguistic creatures, in other words, can fix a certain relation between their words and the items (causes) in the world, which gives content to their responses. This conclusion seems to capture what Verheggen takes to be Davidson's non-reductionist response to KW's sceptic. However, what exactly is this non-reductionist response?

[96] For instance, if at time t_1 the triangulators both responded by "tree" to an object in front of them and, then, at time t_2, one responded by "table" and the other by "tree" to a seemingly similar stimulus, then they are in a position to realize that the similarity that their responses had at the time t_1 is broken at the time t_2. One of them is responding differently this time. How can this situation help them to see what has went wrong? According to Verheggen, they should settle the disagreement between their responses by speaking to each other.

For Verheggen, the non-reductionist response, which Davidson's triangulation argument sustains, is what Wittgenstein himself indicated before:

> Wittgenstein suggests that unless one already has a language one cannot establish a connection, e.g., by pointing, between the utterance of the word and the color, for the sheer pointing and utterance leave it indeterminate whether it is the color or, say, the material or the shape, of an object that one is talking about. (2016, p.72)

Indeed, primitive triangulation showed this problem: responding in a similar way to an object cannot by itself solve the aspect problem, and thereby "the problem Wittgenstein uncovers is similar to that developed by Davidson in the first step of the triangulation argument" (2016, p. 74). Davidson, according to Verheggen, solves this problem by the second step of his argument, i.e. by arguing that "only those who have distinguished, at least on some occasions, between what is the case and what seems to them to be the case, and thus, only those who have the concept of objectivity, are in a position to have fixed the causes, and hence the meanings, of their utterances" (2016, p. 75). And, as we saw above, acquiring the concept of objectivity essentially requires linguistic interactions. Hence, there is no reductionist solution to the meaning-determination problem, and we must take the speakers as already having a language and their utterances as already having determinate meanings or contents. Therefore, what Davidson did, according to Verheggen, was to take "meaning facts to be primitive" (2016, p. 84). Now, how is this related to KW's sceptic's sceptical argument?

First of all, Verheggen agrees that "Davidson misunderstands the problem Kripke is grappling with" (2016, p.142). The reason, for Verheggen, is that Davidson's suggested solution in terms of interpretation and induction assumes "that what the speaker means is determined by some fact, and that the only problem is to figure out what the fact is" (2016, p. 142). But, this is what was in dispute at the outset: "In the absence of a fact to point to, apparent success in interpreting the speaker will not cut it – we would have been equally successful if we had interpreted her as meaning quaddition rather than addition" (2016, p. 142). As we discussed in Chapter Four, without an argument in favor of the existence of the meaning facts, the apparent success in interpretation is just superficial. Similarly, regarding our discussions in Chapter Three, Davidson must either confirm the existence of meaning facts (and, thereby, diverge from Quine) or reject the existence of such facts (and embrace the sceptical outcome of Quine's arguments). However, what are such meaning facts, according to Verheggen?

In her view, the facts that Davidson introduces are indeed nothing but the relations between the speaker, the external objects, and the sentences the speaker chooses to utter. We saw, in Chapter Three (Section 3.2), that Davidson introduced the facts as "the empirical relations between a speaker, her sentences, and her environment" (1999b, p. 596). This, according to Verheggen, means that Davidson's solution is given by concentrating "on external things as the determinants of meaning" (2016, p. 143). We argued, in Chapter Three (Section 3.2.2), that Davidson's acceptance of Quine's physicalism and the indeterminacy of translation thesis makes it puzzling whether he counts such empirical relations between the speaker and the world as facts about stimulus meaning or facts about fine-grained meanings. Verheggen, however, thinks that such relations are taken by Davidson to be the facts about fine-grained meanings. Speakers fix a certain relation between their words and specific items (causes) in the world through linguistic triangulation. How can this deal with KW's sceptic's sceptical problem? Verheggen, to turn Davidson's remarks on the notion of triangulation into a non-reductionist response to the sceptical problem, first of all, offers the following reading of KW's sceptic's sceptical problem: mere associations between our responses (our uttered words) and external (as well as internal) things, according to KW's sceptic, cannot determine the meaning of our words, since this relation can be interpreted in many different ways. KW's sceptic's problem, as Verheggen emphasizes, was about "a certain conception of how meanings should be determined, through sheer association between expression and extra-linguistic items" (2016, p. 143). We can always take one aspect of the cause rather than another to be what determines the content of the utterance.

At this stage, however, the question that arises regarding Verheggen's reading of KW's sceptic's sceptical problem and Davidson's argument from triangulation, is whether the aspect problem, which emerges in Davidson's discussion of cause-determination, can be identified with KW's sceptic's sceptical problem? It is true that Davidson's perceptual externalism implies that external items, at least partly, determine the meaning of our utterances and the content of our thoughts; as a result, hence, solving the aspect problem is crucial to fixing meanings. However, as argued in Chapter Four (Section 4.1.3.a), one worry is that fixing the external cause of the speaker's responses does not by itself solve the sceptical problem. KW's sceptic could grant the entire causal story Davidson has suggested in his triangulation argument and still raise his sceptical argument: In what way does the fixed external cause determine the meaning of

the utterance? What concept does a speaker express by being prompted by a certain cause to make such-and-such an utterance? The aspect problem, at least as it is presented in Davidson's discussion of the notion of triangulation, is not KW's sceptic's sceptical problem. We still need an account of what it is that the speaker means by her utterance, given that her utterance is caused by some specific thing in the world. One may provide a non-reductionist response to deal with this latter problem. Verheggen, however, seems to take both problems to be the same. According to her, the first step of the triangulation argument states that "invoking features of the world around us, when considered by themselves, yields the interpretation problem just as much as invoking any other entities. The skeptic may insist that external features will therefore not do as meaning-determining facts, any more than any of the facts considered by Kripke" (2016, p. 143). Verheggen, however, continues:

> The reply should be that external features will not do only insofar as we think of the relation between them and our utterances in a certain way. There is indeed no way, without a language, to isolate the aspect or part of the world any particular expression used by a speaker is about. (2016, p. 143)

Even if one claims that both the aspect problem and the sceptical problem may be solved by the same solution, it does not mean that they are the same problem.

Nonetheless, the last passage from Verheggen quoted above indicates her reading of Davidson's alleged non-reductionism: external features can be taken as meaning-determining facts only if the relation between them and the utterances is established by those who already have a language. For her, Davidson's triangulation argument shows that it is not the case that any association between words and extra-linguistic reality leads to the sceptical paradox, and hence it shows that KW's sceptic's "insisting on this conception is incoherent, that it is hopeless to try to ground meaning on sheer associations between expressions and extra-linguistic items, because the associations could only proceed by means of language" (2016, p. 143). In this sense, we can take the speaker as meaning something specific by a word via being caused by an external object, or event, to utter that word *because* the speaker has already fixed the relation between what he utters and what his utterance is about, i.e. the object or the event. Doing so requires the speaker to actually engage in linguistic triangulation, since, otherwise, there would be no answer to the question to which cause the speaker is responding. This means that, the speaker's utterance of "table", as prompted by the table in the world, is taken by the speaker as already meaning *table*, and not anything else,

because the speaker has fixed a certain relation between the table in the world and "table".

So far, we have tried to find out what facts, according to Verheggen, are counted as meaning facts by Davidson. If the speaker is prompted by tables in his environment to utter "table", and if he has associated those items with his word via linguistic triangulation, then we can say that he means *table* by "table". More specifically, Verheggen introduces the meaning facts as follows:

> The meanings of speakers' words are determined, in the first instance, by the features of their environment on which they have triangulated linguistically and which they have agreed to take as the relevant aspect determining the meanings of their words. Thus the external facts, … that serve as determinants of meaning are those that triangulating speakers agree to take to be so. They do not have to agree to take the same facts to determine the meaning of both interlocutors' words. But, they have to agree as to what facts determine the meanings of the speaker's words. (2016, p. 144)

As this passage suggests, hence, the external facts determining meanings of the speaker's utterances are nothing but the external items, features, events, or objects in the world, which the triangulators agree to be what the speaker's utterances are about. The interpreter must understand what external fact the speaker takes to be the meaning fact. As there are endless facts available to be treated as meaning facts, say, endless aspects of a cause, only that fact is counted as the meaning fact that the speaker picks up *and* the interpreter understands to be the fact constituting the meaning of the speaker's utterance. As a result, Verheggen concludes that "focusing on the external facts is what provides Davidson with a solution to the problem of meaning determination" (2016, p. 141). The association between my words and extra-linguistic reality gives meaning to my words, though, among an endless number of such associations, only one such association is meaning-giving, i.e. the one on which the hearer and I agree to be what determines the meaning of my word. And, according to Verheggen, "we need a language in order to develop" one such association between the world and the words (2016, p. 143). We need a language to do so because we need to possess the concept of objectivity in order to establish such associations, which in turn requires genuine disagreements between our responses to emerge and to be settled via linguistic interpretation. For instance, suppose that I utter "black" when I am pointing to the snow in view. What prompted me to utter this word has been some specific feature of some item in the world: I associated "black" with the snow's having a specific colour. It is the fact that snow has that specific property (the white colour) that gives meaning to my

utterance, though the same feature prompts the interpreter to utter “white”, not “black”. But it does not matter, since, insofar as the interpreter understands what in the world is taken by *me* to be the meaning-giving fact, she would understand what I mean by my utterance of “black” on that occasion. The difference is that *I* chose to apply “black” to the same cause, or feature, in the world, to which the interpreter applies “white”. In this way, the interpreter understands that I mean *white* by “black”. As a result, I am not required to mean the same thing by the same words (for the interpreter, “black” means *black*). Rather, the interpreter should be able to find out which fact is taken by me as the determinant of the meaning of my word. The latter view is what Verheggen has called an “interpersonal view”, as opposed to a “communitarian view”. As Verheggen characterizes it, an interpersonal view implies that “having a (first) language essentially depends on having used (at least some of) one’s words, whatever one means by them, to communicate with others”, while a communitarian view demands that “having a (first) language essentially depends on meaning by one’s words what members of some community mean by them” (2006, p. 203). Davidson’s view implies the first sort of view, rather than the second.

5.1.1. On Verheggen’s View of Davidson’s Non-Reductionism

Therefore, Verheggen believes that Davidson’s considerations on the notion of triangulation give rise to non-reductionism about meaning: only those who have a language can fix certain facts to be what govern the correct application of their words, and such facts are the external facts, which the speaker and the interpreter agree to be what give meaning to the speaker’s utterances. Nonetheless, the question is whether Davidson’s argument from triangulation really establishes this view or just presupposes it. Does this argument provide us with any explanation of or justification for non-reductionism about meaning? I believe that Verheggen’s claim that Davidson’s triangulation argument could establish non-reductionism fails, since the argument from triangulation as presented by Verheggen has already presupposed non-reductionism, rather than establishing it. The argument from triangulation could establish non-reductionism if it could lead us to such a view *without* presupposing it in advance, that is, if Verheggen could show how the argument can resist KW’s sceptic by providing a non-reductionist response to the sceptical problem. But, what Verheggen’s reading of Davidson’s argument from triangulation shows is that this argument leads to KW’s

sceptic's sceptical problem, *unless* we assume that non-reductionism is true in advance. Below, I try to clarify this point.

Verheggen agrees that Davidson's actual response to the sceptical argument in terms of the notions of mutual interpretation and induction is mistaken. Davidson's response failed because the sceptic's problem was metaphysical, not epistemological: it was not about the problem of explaining how we can know the facts about what a speaker means by his word; rather, it was about whether there is any fact of the matter as to what a speaker means by his words. Davidson's alternative solution might help us to deal with the problem *if* it was epistemological: by respecting the formal and empirical constraints on the procedure of interpretation, we could get closer and closer to the *facts* about what the speaker means by his words through engaging in the process of interpretation, induction, and evidence-collecting. This solution, however, could not deal with the sceptic's claim that there is no fact about meaning at all. And if there is no fact about what a speaker means by his words, there would be no fact about any success in interpretation either. In Chapter Four (Section 4.1.3), we also considered Davidson's remarks on the notion of triangulation and argued that the causal facts that Davidson adduces fail to resist the sceptical problem, since the sceptic can take such facts for granted and still run his sceptical argument. In other words, the fact that a speaker is prompted by a certain object in the world to utter "table" cannot by itself answer the sceptic's question: What is it that the speaker means by his utterance, given that he is caused by that specific object in the world to utter "table"? Having been prompted by a certain object in the world to utter "table", the speaker could have intended to mean different things by his word, such as *table*, *tabair*, etc.: What is it that makes it the case that one of such sceptical hypotheses is true, i.e. that the speaker means *table* rather than *tabair* by his utterance, when a tabair is a table up until time *t*, and a chair after *t*?

Verheggen's idea was that this problem could be solved by Davidson if we take him to be claiming that the causal facts are counted as meaning facts only if the speaker already fixes them as such: the fact that that table in the world caused the speaker to utter "table" gives meaning to the speaker's utterance only if the speaker already determined what he means by "table" and the interpreter understands it. As Verheggen said, "The meanings of speakers' words are determined, in the first instance, by the features of their environment on which they have triangulated linguistically and which they have agreed to take as the relevant aspect determining the meanings of their words" (2016, p. 144). This is obviously what we took to be a question-begging way of

establishing non-reductionism: the argument from triangulation as it stands leads to KW's sceptic's sceptical problem, *unless* we take non-reductionism to be true in advance. In other words, *because* the causal facts as the determinants of meaning failed to deal with KW's sceptic's sceptical problem, we have thereby no choice but to treat them as irreducible, primitive meaning facts. But, why couldn't one endorse non-reductionism in the beginning by claiming that the fact that a speaker means such-and-such by his word is itself a primitive *sui generis* fact, without bringing in the talk of causal facts and the notion of triangulation? Verheggen herself conceded that the association between words and objects will by itself lead to scepticism about meaning, since an endless number of such associations can be proposed by the sceptic to be meaning-giving, and nothing about the speaker can pick up one of them as correct. These associations would not lead to meaning scepticism, according to Verheggen, if we take it to be an irreducible primitive fact about the speaker that he takes one such association to be meaning giving in advance. However, the question is: What is our justification for such an assumption? I think Verheggen can offer no better reason but the question-begging claim that, if we do not make such an assumption, the sceptical problem would arise. But, in this sense, I think KW's sceptic would be right to treat this sort of response as "desperate" (1982, p. 51), because it really "leaves the nature of this postulated primitive state … completely mysterious" (1982, p. 51) and says nothing about how we know the content of such a state. To resist KW's sceptic's sceptical problem, it is not plausible to claim that, since the causal facts as what constitute meaning facts lead to meaning-scepticism, we have no other choice but to treat them as irreducible, primitive meaning facts. Rather, to establish non-reductionism, the argument from triangulation must be capable of leading us to non-reductionism about meaning, without presupposing it in advance. It is, however, hard to see how Davidson's triangulation argument can accomplish this job. Hence, we can conclude that Verheggen fails to convince us that Davidson's use of the notion of triangulation can establish, or even justify, a non-reductionist response to KW's sceptic, since it begs the question against the sceptic by presupposing a view which has been in dispute in the beginning, namely, non-reductionism about meaning.

To make this point clearer, we can come back to our discussion of Davidson's argument from triangulation in Chapter Two and Chapter Four and see to what conclusion this argument drove us. We faced two problems with the process of cause-determination: the problem of determining whether the cause of the creature's responses

is distal or proximal and the problem of determining which aspect of the cause bestows meaning on the creature's responses. Primitive triangulation, as Verheggen agreed, could not be much of a help to solve any of the mentioned problems. In other words, even if primitive triangulation can help to deal with the first problem, it obviously fails to solve the aspect problem, since the mere correlation between similar responses to some distal cause cannot by itself determine to which aspect of the cause each creature is responding. There are always different aspects of an object that can prompt similar responses. From these considerations, we can at least conclude that engaging in primitive triangulation is not sufficient for the creatures to possess thoughts and utterances with determinate content, since non-linguistic animals can triangulate in this sense.

What about linguistic triangulation? In Chapter Four, we argued that if the fact that two creatures have their similar responses to a certain stimulus in the world cannot solve the aforementioned problems with cause-determination, then it is hard to see how their responses' being linguistic can by itself make any difference. We can in fact detect an implicit assumption in the suggestion that linguistic responses are to be treated as essentially different from non-linguistic ones and that linguistic triangulation can solve the problems with meaning-determination. What is this implicit assumption? The assumption seems to be that the linguistic responses have certain characteristics: they are already meaningful. And, this was the reason why we claimed that Davidson's argument from triangulation presupposes a solution to KW's sceptic's sceptical problem, rather than provides one. In other words, linguistic responses would help to solve the problems with meaning-determination only if we presuppose that the triangulators have already fixed the meaning of their expressions. Davidson put this point in terms of the requirement that the triangulators must engage in the process of *mutual interpretation*, as he says: if the triangulators are to know of each other to what thing the other is responding, they must "be in communication. Each of them must speak to the other and be understood by the other. … They must each be an interpreter of the other" (1992, pp. 121). He, however, did not explain why and how this suggestion can solve the sceptical problem. As argued in Chapter Four (Section 4.1.2), engaging in the process of interpretation and induction would not by itself solve KW's sceptic's sceptical problem: unless Davidson can first answer the question as to what the meaning facts are, there would be no success in interpretation at all. There would be, in other words, no meaning for the interpreter to grasp. In this sense, we concluded that

Davidson's response to KW's sceptic, whether it is put in terms of mutual interpretation or in terms of linguistic triangulation, is a failure.

Verheggen, however, attempted to use Davidson's remarks on triangulation to respond to KW's sceptic: the causal facts are meaning constituting facts only if they are already taken by the speaker to be what bestow meaning on his words. Nonetheless, throughout this section we argued that this response is just question-begging. Without taking non-reductionism to be true in advance, the sceptic would be able to run his sceptical argument in the case of Davidson's claim that success in "interpretation" can determine meaning, just as the sceptic could run his sceptical argument in the case of Davidson's claim that success in "linguistic triangulation" can determine meaning.[97] Therefore, the argument from triangulation as presented by Verheggen does not establish non-reductionism about meaning. As we saw above, Davidson's triangulation argument leads us again to his earlier suggestion that engaging in the procedure of interpretation and induction can solve the sceptical problem, which this time is combined with a causal explanation of this procedure. In this sense, the important point regarding Verheggen's claim is: if one insists that it is "triangulation" that can support non-reductionism about meaning, then one could instead make the same claim about Davidson's earlier suggested criterion for success in interpretation, i.e. the speaker's utterance "being interpreted as intended", as offering such non-reductionism about meaning too. In other words, one can alternatively claim that it is an irreducible, primitive fact about the speaker that, in each case of communication, he intends his utterance to be interpreted in a certain way, and the interpreter's job is to keep track of this fact. What does this claim amount to? It amounts to the claim that, to take Davidson to hold non-reductionism about meaning, we do not really need to rely on his considerations concerning the notion of triangulation. I think, the notion of triangulation is better taken as an analogy, or a metaphor, that at best shows how Davidson's externalism and the process of interpretation are reconcilable, without doing anything constructive to deal with KW's sceptical problem.

[97] Moreover, Verheggen's, and Davidson's, emphasis on the fact that, without possessing the concept of truth or objectivity, meaning-determination would not be possible is to the same degree question-begging. As argued in Chapter Four (Section 4.1.3.a), the sceptic has indeed challenged the claim that a speaker can be said to grasp one concept rather than another, since there is no fact of the matter as to what concept a speaker grasps including that of truth or objectivity. Presupposing a grasp of this concept begs the question against the sceptic.

The important question at this point, however, is: What does it take for one to establish a non-reductionist response to KW's sceptic? Of course, KW's sceptic's sceptical problem would be dealt with if Davidson could provide us with a non-reductionist account of meaning. The sceptic could not raise his sceptical problem again by putting forward alternative sceptical interpretations of the way the speaker associates his words with specific items in the world *if* the fact that the speaker meant something specific by his words could be claimed to be an irreducible primitive fact. Interestingly, KW's sceptic himself did not say that one cannot simply take meaning facts to be primitive. As he says, "perhaps it is simply a primitive state, not to be assimilated to sensations or headaches or any 'qualitative' states, nor to be assimilated to dispositions, but a state of a unique kind of its own" (1982, p. 51). As the sceptic confesses, "such a move may in a sense be irrefutable" (1982, p. 51). Non-reductionism is an option, and Davidson could choose it. The sceptic's problem with this option, however, was the following: How can we justifiably establish this response? In order to hold such a non-reductionist response to KW's sceptic, what Davidson is required to do is to clarify what such a primitive fact is and especially explain how we know about it "with some fair degree of certainty whenever it occurs" (Kripke, 1982, p. 51). This brings us back to KW's sceptic's "Argument from Queerness", which was put forward by the sceptic to rule out such a response to the sceptical problem (see Section 1.2.9). This argument, in this case, leads to the problem of accounting for self-knowledge: How do we have noninferential knowledge of what we mean by our word in each case of using it, considering that there can be a potentially infinite number of such cases?

Indeed, Davidson could claim that such a primitive state of meaning something by an utterance is the state of the speaker's intending his utterance to be interpreted in a particular way. Nonetheless, the main problem for Davidson is to resist the sceptic's argument from queerness by providing a plausible account of first-person authority. If it is a primitive fact that what the speaker intends to mean by his words is what the words mean for him, then it has to be explained how the speaker directly knows what he means, believes, and intends, in a way different from the way the interpreter knows what the speaker means, believes, and intends. But, again, this means that what we need do is to look for the potential resources in Davidson's works supporting a proper account of self-knowledge. This is the point that Verheggen neglects. Even if we could take Davidson's argument from triangulation to be supporting non-reductionism about

meaning, unless Davidson could offer a plausible response to KW's sceptic's argument from queerness, we would not yet be able regard to his non-reductionism as plausible.

In the following sections, we will look at two different ways of accounting for self-knowledge and, hence, defending a form of non-reductionist response to KW's sceptic: Davidson's account of first-person authority and Wright's judgement-dependent account of intention. Generally speaking, Davidson believes that intentions have the following characteristics: "Intention, like belief and expectation, does not require attention or reflection, and intentions are not usually arrived at by conscious reasoning. Intentions are not normally attended by any special feelings, nor is our knowledge of our own intentions arrived at (usually) by inference or resort to observation" (1992, p. 112). We will discuss in detail his sophisticated account of intention later in this chapter (Section 5.7.2), but, for now, we can note the following point: to intend to do something (such as to intend to be interpreted in a particular way) is not, for Davidson, a qualitative state, and our knowledge of such a state is not introspectible, observational, or inferential. Rather, we noninferentially know what we intend. However, putting it in terms of the problem of self-knowledge, how can Davidson plausibly explain speakers' noninferential knowledge of what they intend and mean?

5.2. Davidson on First-Person Authority

In his paper, "First Person Authority" (1984a), Davidson attempts to provide an explanation of the phenomenon of self-knowledge. Generally speaking, Davidson takes the requirement that a speaker has noninferential knowledge of his own intentions and meanings as what is implied by the possibility of interpretation. As he says, "a new explanation of first person authority is offered which traces the source of the authority to a necessary feature of the interpretation of speech" (2001b, p. xiii).

Davidson introduces the problem of self-knowledge as follows:

> When a speaker avers that he has a belief, hope, desire or intention, there is a presumption that he is not mistaken, a presumption that does not attach to his ascriptions of similar mental states to others. … What accounts for the authority accorded first person present tense claims of this sort, and denied second or third person claims? (1984a, p. 3)

Davidson is after an explanation of such an asymmetry between the attributions of certain attitudes to ourselves and the attributions of similar attitudes to others, the

existence of which we intuitively concede. For Davidson, the special authority a speaker has over his attitudes applies to all sort of propositional attitudes, but, in the case of belief, on which he is going to focus, it is more direct, clear, and simple.[98]

First of all, according to Davidson, although speakers have special authority over their beliefs, error is always possible. First-person knowledge is not infallible: "We do not always have indubitable or certain knowledge of our own attitudes. Nor are our claims about our own attitudes incorrigible. It is possible for the evidence available to others to overthrow self-judgements" (1984a, p. 4). This claim, as we will see later in this section, has its own manifestation in Davidson's discussion of the process of interpretation too. A speaker can be said to be successfully meaning something if he is interpreted as he intends; otherwise, the speaker fails to mean anything at all, and, thereby, we can say that he does not know what he means and believes. In this case, the speaker has just made a mistake in thinking that he meant something specific by his word.[99]

For Davidson, what we need to do is to explain why there is an asymmetry in the ascriptions of attitudes to ourselves and to others. Davidson puts forward his own proposal, which starts with a distinction between two sorts of asymmetry:

> On the one hand, there is the familiar difference between self- and other-attributions of the same attitude to the same person: my claim that I believe Wagner died happy and your claim that I believe Wagner died happy. If these claims are put into words, we have the difficulty of deciding what pairs of utterances are suitably related in order to guarantee that the claims have the 'same content'. On the other hand, we may consider my utterance of the sentence 'I believe Wagner died happy', and then contrast my warrant for thinking I have said something true, and your warrant for thinking I have said something true. (1984a, p. 11)

Davidson is inclined to deal with the second sort of asymmetry, that is, to find an answer to the question: "What explains the difference in the sort of assurance you have

[98] As Davidson clarifies, "the question whether someone intends to lock the door by turning the key depends in part on whether he wants to lock the door and believes that turning the key will lock the door; and whether this belief and desire have caused, in the right way, a desire to turn the key. Special authority attaches directly to claims about the desire and belief, less directly to claims about the necessary causal connection" (1984a, p. 4).

[99] However, there is another aspect of making mistakes, which we will discuss further: it is not always the case that the failure is on the speaker's part; it may be the interpreter who makes a mistake in correctly interpreting the speaker, while the speaker has provided, or believed that he has provided, enough clues and evidence for the intended interpretation to be achieved. This is an important aspect of error, which Davidson's account of error has to accommodate.

that I am right when I say 'I believe Wagner died happy' and the sort of assurance I have?" (1984a, p. 11). He prefers to deal with the second sort of asymmetry because, not only are these two asymmetries related, but Davidson also has a linguistic explanation of the asymmetry, which is matched with the second characterization of the asymmetry. As Davidson says, "if one can speak with special authority, the status of one's knowledge must somehow accord" (1984a, p. 3). His account concerns the process of interpretation, the meaning-belief relation, and holding-true attitudes.[100] According to Davidson, "if you or I or anyone knows that I hold this sentence true on this occasion of utterance, and she knows what I meant by this sentence on this occasion of utterance, then she knows what I believe – what belief I expressed" (1984a, p. 11). For Davidson, however, "we can assume without prejudice that we both know, whatever the source or nature of our knowledge, that on this occasion I do hold the sentence I uttered to be true" (1984a, p. 12). In other words, Davidson allows that the interpreter, as well as the speaker, knows that when the speaker utters a sentence, he holds it to be true on that occasion. The speaker, according to Davidson, holds a sentence to be true because of what he means by the sentence and what he believes to be the case.[101] Thereby, if the interpreter knows what the speaker means by the sentence, she knows what the speaker believes, and if she knows what the speaker believes, she knows what he means by his word. So far, no asymmetry between the situation occupied by the first-person and the situation occupied by the second-person regarding the first person's attitudes emerges. According to Davidson, the difference emerges in the following way:

> You and I both know that I held the sentence 'Wagner died happy' to be a true sentence when I uttered it; and that I knew what that sentence meant on the occasion of its utterance. And now there is this difference between us, which is what was to be explained: on these assumptions, I know what I believe, while you may not. (1984a, p. 12)

The difference has its roots in the interdependence of meaning and belief: "The assumption that I know what I mean necessarily gives me, but not you, knowledge of what belief I expressed by my utterance" (1984a, p. 12). As outlined in the Appendix to Chapter Two (see p. 256), Davidson believes that, at least in the most basic cases, when a speaker utters a sentence, he expresses what he believes. In this way, the process of

[100] These notions have been introduced in the Appendix to Chapter Two.

[101] As Davidson says, "a speaker's assent to a sentence depends both on what he means by the sentence and on what he believes about the world" (1983, p. 147).

interpretation and its implications explain the asymmetry we intuitively think there is in the case of attributions of attitudes: meaning and beliefs (as well as intentions) are interrelated, and if I know what I mean, then I know what I believe. But, the interpreter reaches my belief only after engaging in the process of interpreting my utterance; she must first interpret my utterance correctly, by utilizing the available evidence and clues, in order to realize what I believe (by presupposing that I hold my uttered sentence to be true on this occasion).

What remains to be explained in order to explain the asymmetry is the claim that the interpreter does not know what the speaker means in the same direct way: "Why there must be a presumption that speakers, but not their interpreters, are not wrong about what their words mean" (1984a, p. 12). If speakers are not wrong about what they mean by their words, then they are not wrong about their beliefs, while interpreters are always apt to make mistakes in interpreting what the speakers mean and believe. But, why must there be such an assumption? Davidson's answer is that "the presumption is essential to the nature of interpretation – the process by which we understand the utterances of a speaker" (1984a, p. 12). In other words, if speakers were not mostly right about what they mean (and, thereby, about what they believe), then there would be no interpretation at all. If that is true, then the asymmetry is explained: "A hearer interprets (normally without thought or pause) on the basis of many clues; … The speaker, though he must bear many of these things in mind when he speaks, since it is up to him to try to be understood, cannot wonder whether he generally means what he says" (1984a, p. 12). The interpreter, who is inevitably dependent on evidence and clues, "is liable to serious error", that is, "there can be no general guarantee that a hearer is correctly interpreting a speaker" (1984a, p. 12). The speaker does not know himself via interpretation, induction, and evidence-collecting: the interpreter or the hearer "may always be regarded as interpreting a speaker. The speaker cannot, in the same way, interpret his own words" (1984a, p. 12). Such a contrast is emphasized by Davidson by saying that

> the speaker, after bending whatever knowledge and craft he can to the task of saying what his words mean, cannot improve on the following sort of statement: 'My utterance of "Wagner died happy" is true if and only if Wagner died happy'. An interpreter has no reason to assume this will be *his* best way of stating the truth conditions of the speaker's utterance. (1984a, p. 13)

Concerning the available evidence and clues, the interpreter is always trying to arrive at the best interpretation of the speaker's utterances, i.e. that which makes the behaviour of the speaker intelligible; from the interpreter's point of view, however, disquoting the

sentence uttered by the speaker may not always be the best way to understand the speaker's utterance. Nonetheless, the knowledge that we can credit the speaker with is at best the speaker's knowing that "'Wagner died happy' is true if and only if Wagner died happy". We cannot find evidence changing the interpretation of the right-hand side sentence of the above theorem but leaving the interpretation of the left-hand side sentence untouched; rather, whatever evidence there is for or against the right-hand side sentence will be counted as evidence for or against the left-hand side sentence. The speaker directly knows what he means, while the interpreter need not necessarily take the disquoted sentence on the right-hand side to be a specification of the uttered sentence's truth-condition.[102]

What we have so far considered does not imply that the speaker is *always* right about what he means and believes, or that he is immune to error, since he may fail to speak in such a way to make his utterance understandable or interpretable to the interpreter. In other words, we do not "assume that no question can arise concerning a speaker's interpretation of his own words"; rather, what a speaker's words mean "depends in part on the clues to interpretation he has given the interpreter, or other evidence that he justifiably believes the interpreter has. The speaker can be wrong about what his own words mean" (1984a, p. 13). If that is true, then "first person authority is not infallible" (1984a, p. 13). Davidson emphasizes that this fallibility does not affect the asymmetry: "The asymmetry rests on the fact that the interpreter must, while the speaker doesn't, rely on what, if it were made explicit, would be a difficult inference in interpreting the speaker" (1984a, p. 13). Hence, as Davidson concludes, there is no special, or mysterious kind of knowledge that the first-person, but not the interpreter, is supposed to have: "Neither speaker nor hearer knows in a special or mysterious way what the speaker's words mean; and both can be wrong" (1984a, p. 13). As a reply to KW's sceptic's argument from queerness, therefore, Davidson can be taken to be claiming that the speaker's knowledge of what he means, or of the way he intends his utterance to be interpreted, is by no means mysterious. There is no mystery in knowing what I believe, since I directly know what I believe simply because I know what I mean; and I directly know what I mean because if I don't, there would be no interpretation, and, thereby, no meaning at all. As Davidson says,

[102] We will discuss this characterization of first-person authority further in Section 5.4. .

> the speaker may fail in this project [of being interpretable] from time to time; in that case we can say if we please that he does not know what his words mean. … But if [speaker's behaviour] is [interpretable], then what his words mean is (generally) what he intends them to mean. … There is a presumption – an unavoidable presumption built into the nature of interpretation – that the speaker usually knows what he means. So there is a presumption that if he knows that he holds a sentence true, he knows what he believes. (1984a, pp. 13-14)

If the speaker fails to provide enough evidence and clues for the interpreter to understand what he meant by his words, then there is simply no meaning to be known. As discussed in Chapter Two, in Davidson's view, meaning emerges as a consequence of successful communication, without which there would be no meaning. If interpretation is to be possible at all, the speaker must have direct, noninferential knowledge of what he means and believes: the speaker is mostly right about what he means and believes. No massive error in beliefs and meanings is possible; otherwise, there is no ground for interpretation to proceed.[103]

5.3. Wright's Objection to Davidson's Account of Self-Knowledge

Crispin Wright, in "The Problem of Self Knowledge (II)" (2001c), has raised an objection to Davidson's account of first person authority. In Wright's view, there is an "essential self/other asymmetry in the means of knowledge": we know ourselves "*differently* from the way in which we know others and they know us" (2001b, p. 320). This asymmetry is the phenomenon we need to explain, which, according to Wright, is the phenomenon of *avowal* – the phenomenon of authoritative, non-inferential self-ascription" (2001b, p. 320), such as (the attitudinal avowal of) "I believe that it is raining". According to Wright, Davidson's proposal is based on the features and implications of the process of interpretation, in which the interpreter's knowledge of the speaker's beliefs and meanings is achieved on the basis of evidence and induction, contrary to the case of the speaker's knowledge of himself. In the case of the second person, "any claim made about what a (putative) fellow English-speaker means by a particular sentence, even one that uses that very sentence to specify the content in question, is hostage to the evidence of interpretation" (Wright, 2001c, p. 348). In this way, what one means by uttering a sentence depends on his utterance's being

[103] For more on Davidson's account of self-knowledge, see e.g. (Ludwig & Lepore, 2005, Section 20), (Hacker, 1997), (Thöle, 1993), (William, 2007), and (Holly, 1986).

successfully interpreted by an interpreter so that different, non-standard uses someone makes of his words may defeat, or change, the interpreter's opinion about what the speaker means, or was expected to mean, by his words. By contrast, in the case of the first person, "when I use a sentence to specify disquotationally what *I* mean by it, there is no such hostage" (2001c, p. 348). For example, suppose that Mrs. Malaprop utters "That's a nice derangement of epitaphs". Her interpreter will try, by utilizing the available evidence and clues, to specify what she means by her utterance, for example, by giving the following disquotationally specified truth-conditions of her uttered sentence: (I1) "'That's a nice derangement of epitaphs' as uttered by the speaker is true if and only if that's a nice derangement of epitaphs". However, although this interpretation may seemingly be giving the right sort of truth-conditions for Mrs. Malaprop's utterance, it might not be the correct interpretation of the utterance, since she might intend to mean something different by uttering that sentence, e.g. that *that's a nice arrangement of epithets*. Suppose that the interpreter finds out that "derangement", as Mrs. Malaprop is using it, means *arrangement* and "epitaphs" means *epithets*. For instance, when Mrs. Malaprop applies "epitaphs" to the things that are not epitaphs, but epithets, and uses "*x* is a nice derangement" in the cases in which *x* is nicely arranged, this evidence will motivate the interpreter to change his putatively given truth-conditions of the speaker's utterance as follows: (I2) "'That's a nice derangement of epitaphs' as uttered by Mrs. Malaprop is true if and only if that's a nice arrangement of epithets". Therefore, the evidence that the interpreter has for the sentence on the right-hand side of the suggested interpretation is not necessarily the evidence for the left-hand side sentence, so that the evidence which supports the inclusion of the right-hand side sentence in (I2) does not support the inclusion of the right-hand side sentence in (I1). Hence, the interpreter's claims about what Mrs. Malaprop means by her sentences are hostage to the available evidence and clues.

Nonetheless, according to Wright's reading of Davidson's account, the same does not hold in the case of Mrs. Malaprop's considering her own utterances, that is, when she ascribes meanings to her own utterances. What she means by a sentence is directly known to her: she does not interpret her utterance on the basis of evidence and clues; rather, what she takes the sentence to mean is just what it means for her. In the cases in which she attributes meaning to her own utterance of "That's a nice derangement of epitaphs", (I1) would be the best she can get to specify the meaning of her utterance. For instance, suppose she says to herself "Snow is white". For her, as the first person,

“Snow is white” is true if and only if snow is white. There is no evidence changing the truth-conditions of the sentence which does not apply to the sentence itself. What is counted as evidence for the right-hand side sentence – the used sentence – would be the same evidence for the left-hand side sentence – the mentioned sentence. Davidson’s view is summarized by Wright:

> Suppose, then, that I now assent to S. Both you and I can, so to speak, observe my assent. But I am proof against error about *what* I have assented to in a way that you are not. We are both in position to conclude that I believe something. But I am uniquely assured of no error in the specification of *what* I believe. Therein, suggests Davidson, lies the source of my special authority. (2001c, p. 348)

In Wright’s view, Davidson tries to solve the problem of self-knowledge via defending the idea that the speaker does not, while the interpreter must, go through the process of interpretation, which essentially relies on evidence and induction.

According to Wright, however, Davidson’s proposal fails, since it faces a dilemma regarding the interpretation of ‘assent’. My assent to a sentence, S, indicates my holding the sentence true on that occasion, and if such evidence is supposed to be publicly available to the interpreter too, then it must have a behavioural form. As Wright says, “Davidson takes my assent to S to be available as a datum to both me and you. So ‘assent’ ought to mean: overt assenting behaviour (however that is characterised)” (2001c, p. 349). Such an idea was endorsed by Davidson himself in his discussion of radical interpretation, that the ultimate source for interpretation is the behaviour of the speaker. My holding a sentence, S, true (or, equivalently, my assent to S) is the publicly available evidence which is the source of the interpreter’s extracting what I mean by S and what I believe to be the case. The asymmetry between the speaker and the interpreter emerges regarding the fact that the speaker, but not the interpreter, directly knows what he means and believes, while the interpreter indirectly, via interpretation, knows what the speaker means and thereby what the speaker believes. Wright’s problem is about how we have to interpret such publically available evidence and how we can explain the asymmetry on the basis of that interpretation.

For Wright, there is a difference between what I really believe and what is extracted from my behaviour: “Overt behaviour as of assent is one thing; genuine belief another” (2001c, p. 349). The issue in question is how the speaker, but not the interpreter, can directly and noninferentially know what she genuinely believes. According to Wright,

Davidson's account faces a dilemma, since my assent to S must be taken to be either mere behavioural evidence or behaviour expressive of what I genuinely believe. The first horn of the dilemma is about taking my assent to S as purely behavioural, i.e. as not already containing, or expressing, what I genuinely believe. If this is the way Davidson treated my assent to S, then it is not in advance presupposed that the speaker has any specific knowledge of such evidence as an expression of her genuine belief. Wright's objection is that, in this case, we will lose the asymmetry altogether: if my assent to S is just mere behavioural evidence, then it does not show that I have noninferential knowledge of what I genuinely believe *before* the interpretation takes place, that is, I cannot be said to have expressed any of *my* beliefs until my behaviour is successfully interpreted. I have a belief only *after* interpretation takes place. Hence, if the explanation goes in this way, although we intuitively want to credit the speaker with direct and noninferential knowledge of what she believes, we cannot achieve this, and what the speaker believes will thereby depend on how she is interpreted. As a consequence, what is credited to me as what I believe by my interpreter (based on the behavioural evidence) must be taken to be what I genuinely believe – since it has not been presupposed in advance that my assent to S is an expression of my genuine belief as I directly know it. But, what we intuitively expected was an explanation of the asymmetry between the interpreter's and my knowledge of my beliefs, while what we have come up with is nothing but the view that the interpreter and I have a similar position regarding my beliefs. My first-person authority over my beliefs is thereby left unexplained. As Wright says, "part of the phenomenon to be explained is that I can know, effortlessly and non-inferentially, whether I hold it or not. When 'assent' is interpreted behaviourally, Davidson's reconstruction leaves me no better placed on that matter than you are" (Wright, 2001c, p. 349).

However, Davidson could alternatively claim that my assent to S is an expression of what I genuinely believe: when I assent to S, I know what belief I am expressing. However, Wright's problem with this reading of "assent" is that it is now presupposed *in advance* that I have direct access to, or knowledge of, the belief that is expressed in my assent to S, but the interpreter does not. For the interpreter, my assent to S is just mere behavioural evidence, which should be first interpreted if it is to be treated as an expression of what I believe; for me, however, my assent to S is primarily taken to be containing what (I know) I genuinely believe. According to Wright, Davidson begs the question if he takes the behavioural evidence as expressive of a genuine belief of mine,

since the asymmetry between my knowledge of my belief and the interpreter's knowledge of my belief is thereby presupposed, rather than explained. If "'assent' is interpreted as behaviour expressive of a genuine belief, then there is an asymmetry in our knowledge of the supposed datum – my assent to S – about which his proposal says nothing, and which is effectively the very thing it was supposed to account for" (Wright, 2001c, p. 349). The problem was to explain how the speaker, but not the interpreter, can noninferentially know what belief she is expressing by her assent to S. But now it is already granted that the speaker, independently of the interpreter or interpretation, knows that her behavioural assent to S expresses what she believes. The problem, hence, is to explain how I can know such a thing *before* interpretation takes place. Why is this behaviour, for me, an expression of what I genuinely believe, but not for the interpreter (unless he successfully interprets me)?

As a result, Wright concludes that "Davidson has simply missed the matter for explanation" (2001c, p. 349). For Wright, Davidson's appeal to interpretation to solve the problem of self-knowledge cannot provide the plausible explanation we were after.[104]

5.4. Evaluating Wright's Objection

In this section, I will examine Wright's objection to Davidson by searching for an answer to the following three questions. (I) What is Davidson's *explanation* of the asymmetry? (II) Does Wright successfully characterize Davidson's explanation? (III) In either case, is Wright's objection plausible?

[104] As Wright states, in Davidson's proposal, "the authority over my beliefs with which you will normally credit me pertains to what I say about them *after* you take it that you are interpreting those remarks correctly – when my distinctive advantage, construed as by Davidson, has already been neutralised" (2001c, pp. 349-350). Hence, "*after* you assume the correctness of your interpretation of my ungrounded, unsupported claims about my beliefs, you are required to treat them as a criterion of what my beliefs are" (2001c, p. 350). However, if Davidson wants to claim that my assent to S is expressive of my genuine belief, which I know without interpreting myself first, then he is presupposing the asymmetry rather than explaining it: "it kicks in after that asymmetry is allowed for" (2001c, p. 350).

5.4.1. Davidson's Explanation of the Asymmetry

First of all, it seems to be clear that the source of Davidson's account of the asymmetry comes from his remarks on interpretation. As he emphasizes, his explanation is supposed to trace "the source of the authority to a necessary feature of the interpretation of speech" (2001b, p. xiii). It is what the possibility of interpretation implies that is supposed to explain the asymmetry. In "First Person Authority" (1984a), although the solution to the problem of self-knowledge is given at the end of the paper, Davidson, in the beginning of the paper, points to a *presumption* that we have in our interaction with others, which needs to be properly accounted for: "When a speaker avers that he has a belief, hope, desire or intention, there is a presumption that he is not mistaken, a presumption that does not attach to his ascriptions of similar mental states to others" (1984a, p. 1). This means that, for Davidson, there *is* an asymmetry whose existence is taken for granted, and the problem is to provide an explanation of it, that is, of why there exists such an asymmetry. It is a fact that "all propositional attitudes exhibit first person authority" (1984a, p. 4). As Davidson emphasizes in his criticism of Ryle, the problem with Ryle's account of self-knowledge is that he "neither accepts nor explains the asymmetry; he simply denies that it exists. … I think it is obvious that the asymmetry exists" (1984a, p. 6).[105] As we will argue, presupposing the asymmetry in order to explain why it exists is one thing, and presupposing the asymmetry without explaining it (and thereby begging the question, or taking for granted what ought to be explained) is another. Moreover, we should note that Davidson's approach to the problem of self-knowledge is linguistic, which means that he is going to explain the asymmetry regarding the attributions of beliefs (or propositional attitudes, in general) through explaining another asymmetry regarding the attributions of meaning, since, according to Davidson, "if one can speak with special authority, the status of one's knowledge must somehow accord" (1984a, p. 1). Davidson intends to explain the first-person authority in speech in order to allow himself to explain the asymmetry which emerges in belief-attributions.

[105] Briefly speaking, Ryle denies the existence of the asymmetry. According to Ryle, there is no privileged access to our beliefs and mental states; rather, we, as the first persons, just occupy a better position, a better place, than others, to observe ourselves. There is, hence, no essential difference between my way of knowing myself and your way of knowing me. The differences, as Davidson mentions, "are difference of degree, not of kind" (1984a, p. 5). For Ryle's view on this matter, see (Ryle, 1949).

The important point on which Davidson insists throughout his paper is that appealing to the notion of evidence in order to explain the asymmetry is *not* a plausible way to deal with the problem of self-knowledge. He does not think that the availability or unavailability of evidence can be taken to be what explains the asymmetry, although it is a tempting way to deal with the problem of self-knowledge[106]: "This feature of first person authority, suggestive as it may be, does not help explain the authority" (1984a, p. 5). According to Davidson, this evidence-based account of self-knowledge has its roots in Wittgenstein's remarks on the matter. As Davidson puts it, for Wittgenstein, "What is the criterion for the redness of an image? For me, when it is someone else's image: what he says or does. For myself, when it is my image: nothing" (1984a, p. 4). Davidson tends to sympathize with this Wittgensteinian claim, since for Davidson too the self-attributer does not base his claims on evidence or observation. But he does not think that this by itself *explains* the asymmetry. Rather, what explains this feature is something else: "First person authority isn't explained by the fact that self-attributions are not based on evidence" (1984a, p. 5). What is Davidson's reason for such a claim? According to Davidson, "the chief reason is simply that claims that are not based on evidence do not in general carry more authority than claims that are based on evidence, nor are they more apt to be correct" (1984a, p. 5).[107] Davidson emphasizes this point elsewhere by saying that

> it is a strange idea that claims made without evidential or observational support should be favored over claims with such support. Of course, if evidence is not cited in support of a claim, the claim cannot be impugned by questioning the truth or relevance of the evidence. But these points hardly suffice to suggest that in general claims without evidential support are more trustworthy than those with. (1987, p. 16)

Hence, we can say that, for Davidson, there is a difference between *describing* the asymmetry by appealing to the notion of evidence and *explaining* the asymmetry.[108] For

[106] As Davidson characterizes this tempting, but misleading, view, "it comes closer to characterizing first person authority to note that the self-attributer does not normally base his claims on evidence or observation, nor does it normally make sense to ask the self-attributer why he believes he has the beliefs, desires, or intentions he claims to have" (1984a, p. 4).

[107] In Davidson's view, "first person attributions are not based on better evidence but often on no evidence at all" (1984a, p. 6). See also (Davidson, 1991a, pp. 207-208)

[108] Davidson emphasizes this point in his criticism of Strawson's account of self-knowledge. As he characterizes Strawson's view, "according to Strawson, if the skeptic understands his own question ('How does anyone know what is going on in someone else's mind?'), he knows the answer. For if the skeptic knows what a mind is, he knows it must be in a body, and that it has thoughts. He also knows that we attribute thoughts to others on the basis of observed behavior, but to ourselves without such a basis"

instance, Davidson criticizes Rorty for not giving an explanation (but just a description) of the asymmetry by appealing to evidence and observation:

> But the question remains: what reason has Rorty given to show that self-ascriptions not based on evidence concern the same states and events as ascriptions of the same mental predicates based on observation or evidence? There is a difference in kind in the ways the two sorts of ascription are made, and how they explain behavior is different. (1984a, p. 8)[109]

As we discussed before (Section 5.2), Davidson offers his own explanation of the asymmetry as follows. First of all, he thinks that "it is no help to say I have access to a way of knowing about my own beliefs that you do not have" (1984a, p. 11). For Davidson, there is no essential difference in what I, as the first person, and you, as the second person, can know about my beliefs, or about whatever evidence there is about what I believe or mean. What you and I know about me, hence, is the same thing.

Davidson is also aware of the danger of begging the question by presupposing the asymmetry in order to explain it: "There must *be* such an asymmetry, of course, but it cannot be allowed to contribute to the desired explanation" (1984a, p.12). Because of that, he thinks that "it would … make the account circular to explain the basic asymmetry by assuming an asymmetry in the assurance you and I have that I hold the sentence I have just uttered to be a true sentence" (1984a, p. 12). If that is true, then the "basic" asymmetry regarding the attributions of beliefs needs to be explained in terms of something else, not just by presupposing my privileged, unique access to evidence, or to a different sort of resource. These remarks from Davidson support our claim that Davidson's talk of evidence, as well as his disquotational description of the asymmetry, must not be taken to be his *explanation* of the asymmetry, since he thinks that explaining the asymmetry by adverting to the difference in the assurance that I have for my knowledge of my beliefs and you have for your knowledge of my beliefs would commit us to a question-begging way of explaining the asymmetry. In this regard, the

---

(1984a, p. 7). Davidson, however, believes that "the skeptic will reply that though Strawson may have correctly described the asymmetry between first and other person ascriptions of mental predicates, he has done nothing to explain it" (1984a, p. 8). For more on Strawson's account, see (Strawson, 1959).

[109] The sort of explanation Rorty suggests, according to Davidson, is the observation that self-ascriptions of mental properties can be made without using any behavioural evidence or making observations and, in this way, they can provide a better explanation of our behaviour than the ascriptions made by others. Hence, treating self-ascriptions in this way will lead us to a better explanation of the agents' behaviour and, in this regard, as Davidson puts it, "it became a linguistic convention to treat self-ascriptions as privileged" (1984a, p. 8). For more on Rorty's view, see (Rorty, 1970).

interpreter and I know the same thing, that is, I know that I hold this particular sentence to be true on this occasion and the interpreter too knows that I hold this particular sentence to be true on this occasion. My assent to a sentence as expressive of what I genuinely believe is a fact that both I and the interpreter are aware of.

I think it is also very important to note that Davidson aims to explain what he called above the "basic" asymmetry between my knowledge of my beliefs and your knowledge of my beliefs, that is, the asymmetry between the attributions of propositional attitudes. For explaining this asymmetry, we cannot just presuppose another asymmetry: "It would beg the question to explain the basic asymmetry by appeal to some asymmetry in our knowledge of the fact that I know what my sentence, as uttered on this occasion, meant" (Davidson, 1984a, p. 12). To presuppose another asymmetry without explaining it – in order to explain the basic asymmetry – cannot provide us with any plausible explanation of the basic asymmetry. Nonetheless, Davidson says that we can take for granted the following two assumptions: "The assumptions are just these: you and I both know that I held the sentence 'Wagner died happy' to be a true sentence when I uttered it; and that I knew what that sentence meant on the occasion of its utterance" (1984a, p. 12). Now, Davidson claims that, by having these two assumptions at hand, the basic asymmetry between the attributions of beliefs emerges: "Now there is this difference between us, which is what was to be explained: on these assumptions, I know what I believe, while you may not" (1984a, p. 12). If I know what sentence I hold true and what I mean by that sentence, I will know what I believe, while it is only after interpreting what I meant by my utterance that the interpreter can reach what I believed.

Davidson's suggested explanation looks question-begging, since it appeals to another asymmetry, that is, the one between my knowledge of what I mean and the interpreter's knowledge of what I mean. However, we should note that this explanation would beg the question if his explanation stopped at this point. But, Davidson himself already mentioned that we cannot explain the basic asymmetry by simply presupposing another. Davidson is aware of his appeal to a second asymmetry: "The difference follows, of course, from the fact that the assumption that I know what I mean necessarily gives me, but not you, knowledge of what belief I expressed by my utterance" (1984a, p. 12). Hence, although the basic asymmetry between my knowledge of my beliefs and your knowledge of my beliefs is explained by citing a second asymmetry, the second asymmetry needs to be explained. Why is it the case that I noninferentially know what I mean, while you can know what I mean only through interpretation? The explanation of

the basic asymmetry is successful *only if* the second asymmetry is explained. As Davidson himself states, "it remains to show why there must be a presumption that speakers, but not their interpreters, are not wrong about what their words mean" (1984a, p. 12). Davidson's explanation of this presumption goes as follows: "The presumption is essential to the nature of interpretation – the process by which we understand the utterances of a speaker. This process cannot be the same for the utterer and for his hearers" (1984a, p. 12). If understanding, in general, is to be possible, then the speaker cannot be taken to be the interpreter of herself: "The asymmetry rests on the fact that the interpreter must, while the speaker doesn't, rely on what, if it were made explicit, would be a difficult inference in interpreting the speaker" (1984a, p. 13). If the speaker were to be engaging in such difficult inferences, it would make interpretation impossible. In this way, the basic and the second asymmetries are explained. If meanings are noninferentially knowable to us, so are our beliefs. And, meanings are noninferentially knowable, since, otherwise, no understanding, that is, no interpretation, is possible: the interpreter "may always be regarded as interpreting a speaker. The speaker cannot, in the same way, interpret his own words" (1984a, p. 12).[110]

As we saw before (see Section 5.2), Davidson then *re-describes* the second asymmetry between attributions of meaning by talking about disquotational specifications of self-ascriptions of meaning: "The speaker … cannot improve on the following sort of statement: 'My utterance of "Wagner died happy" is true if and only if Wagner died happy'. An interpreter has no reason to assume this will be *his* best way of stating the truth conditions of the speaker's utterance" (1984a, p. 13). The question is *why* the speaker can't improve on the disquotational specification of what he means. The reason seems to have its roots in Davidson's *explanation* of the asymmetry, that such an improvement is impossible if interpretation and understanding is to be possible at all: "There is a presumption – *an unavoidable presumption built into the nature of interpretation* – that the speaker usually knows what he means. So there is a presumption that if he knows that he holds a sentence true, he knows what he believes" (Davidson, 1984a, p. 14, emphasis added). Such a description of the asymmetry, that is,

[110] Davidson also clarifies what he means by the *same way* of interpreting, or the difficult inference in interpreting a speaker: "A hearer interprets (normally without thought or pause) on the basis of many clues: the actions and other words of the speaker, what he assumes about the education, birthplace, wit, and profession of the speaker, the relation of the speaker to objects near and far, and so forth. The speaker, though he must bear many of these things in mind when he speaks, since it is up to him to try to be understood, cannot wonder whether he generally means what he says" (1984a, p. 12).

the claim that the speaker cannot improve on a disquotational specification of what he means by a sentence, seems to be just another re-description, and not an explanation, of the asymmetry: it is just a characterization of the asymmetry, according to which the speaker knows what he means in a direct way, without any appeal to evidence and interpretation, while the same does not hold for the interpreter. Recall that Davidson already rejected the plausibility of any appeal to the existence of some special resources, assurances, or evidence, that the first-person has access to, but the interpreter does not.

We now have enough material for our evaluation of Wright's objection to Davidson's account. I have argued for two claims. (I) The way Davidson describes the asymmetry is different from the way he explains it. (II) Davidson does not explain the basic asymmetry regarding belief-attributions by an ungrounded appeal to another asymmetry regarding meaning-attributions; rather, while he explains the former in terms of the latter, he gives an explanation of the latter by adverting to the necessary features of interpretation. I now use these points to evaluate Wright's objection to Davidson's account.

5.4.2. Wright's Objection Revisited

If the way we have characterized Davidson's account of first-person authority is correct, then the immediate question that arises is whether Wright's reading of Davidson's account acceptable?

Wright's discussion of Davidson's account is very brief. According to Wright's reading of Davidson's account, "Davidson's well-known proposal about attitudinal avowals takes off from an analogue of the last point – that there cannot be one kind of mismatch between contemporaneous first- and second-order thought – and from a point about disquotation" (2001c, p. 348). In this way, he takes Davidson's proposal to be based on a difference between the role evidence plays in the speaker's self-attributions of meaning and others' attributions of meaning to the speaker. Wright continues:

> Any claim made about what a (putative) fellow English-speaker means by a particular sentence, even one that uses that very sentence to specify the content in question, is hostage to the evidence of interpretation. ... By contrast, when I use a sentence to specify disquotationally what I mean by it, there is no such hostage. For whatever interpretation may teach you about the content I attach to the sentence in question will apply to both the mentioned and the used

> occurrences in my specification: whatever I mean by a sentence, S, I am guaranteed to be able to say with perfect accuracy what that sentence means merely by using it. ... I am uniquely assured of no error in the specification of *what* I believe. Therein, suggests Davidson, lies the source of my special authority. (2001c, p. 348)

To make Wright's reading of Davidson's account clearer, we can consider the following example. Suppose that a speaker, X, utters a sentence, S, such as "two plus two equals four". In the case of an interpreter's attributing meaning to X's utterance, we will have the following specification:

(I) X means that *two plus two equals four* by her utterance of "two plus two equals four".

The interpreter interprets X's utterance as meaning that *two plus two equals four*. Even this obvious disquotational specification of what X means by her utterance, according to the process of interpretation, is hostage to evidence. For example, by recalling KW's sceptic's example, it is possible that the interpreter later finds out that X answers by "5" to the question "57 + 68 =?". In this case, the evidence now overturns the previous claim about what the speaker means by her utterance and supports the following:

(II) X means that *two quus two equals four* by her utterance of "two plus two equals four".

X seems to mean *quus* by "plus" when she uses the word. Any specification of what X means by her words, from the interpreter's point of view, is hostage to evidence. Davidson agreed that "an interpreter has no reason to assume this [disquotationally specified interpretation, i.e. (I) above] will be *his* best way of stating the truth conditions of the speaker's utterance" (1984a, p. 13). The evidence suggests that the interpreter's previous interpretation of X's utterance should be withdrawn. However, the same is not the case in the case of self-attributions, for instance, when we have the following specification:

(III) I mean that *two plus two equals four* by my utterance of "two plus two equals four".

There is no possibility of mismatch, according to Wright, between the right-hand sentence and the left-hand sentence of the above judgement so that any evidence which appears to be suggesting a re-interpretation of the right-hand side sentence will suggest

the same re-interpretation of the left-hand side sentence. Any evidence suggesting that "plus" means *quus* applies *both* to the mentioned sentence (on the right) and the used sentence (on the left). As a result, my self-attributions of meaning are not hostage to evidence, while the interpreter's attributions of meaning are, since further evidence can overturn the disquotationally specified meaning of my utterance.

The first worry about Wright's construal of Davidson's account is that it characterizes it as if it is based on the claim that the basic asymmetry is *explained* in terms of the fact that self-attributions of meaning cannot be based on evidence and interpretation, while others' attributions of belief and meaning to the subject are based on evidence and interpretation, that is to say, that evidence cannot overturn disquotational self-ascriptions of meaning, while evidence can overturn others' ascriptions of meaning to the speaker. Obviously, if that was the case, then it would have to be presupposed either that I have special access to my assent to a sentence as expressive of what I genuinely believe, in which case the asymmetry is presupposed rather than explained, or that I do not have such access, in which case the asymmetry disappears. But, as discussed in the previous section, this is not the way Davidson's explanation of the asymmetry proceeds. Let's call what Wright attributes to Davidson a "Disquotational Argument" for self-knowledge. It seems that what Davidson proposes is, by contrast, a "Transcendental Argument", according to which the asymmetry between my knowledge of what I mean and the second person's knowledge of what I mean necessarily exists because, without such an asymmetry, interpretation and understanding would not be possible. The disquotational way in which Wright presents Davidson's account is just the way Davidson *re-describes* the asymmetry, rather than *explaining* it. Wright's reading of Davidson's explanation of the asymmetry, therefore, does not seem to be acceptable.

Nonetheless, one can raise the following question: the Disquotational Argument is indeed an argument, that is, it appears to be an explanation of the asymmetry, and Davidson has indeed used it in his description of the asymmetry. Why shouldn't we take this argument as what explains the asymmetry even if Davidson himself has not used it as his own explanation? What is wrong with this explanation in general? We cannot use this Disquotational Argument as what explains the asymmetry because the asymmetry exists even in cases where the specification of what the subject means is not disquotational. Isn't it possible to take the speaker as knowing, of herself, that she means *rabbit* by "gavagai"? Why can't the speaker ascribe to herself the following statement: "I mean *rabbit* by 'gavagai'"? It seems to be possible that the speaker

ascribes to herself the above specification of meaning, even if the meaning is not disquotationally specified. Recall Davidson's example of Mrs. Malaprop: she meant *a nice arrangement of epithets* by her utterance of "A nice derangement of epitaphs". According to Davidson, Mrs. Malaprop's hearer has no trouble understanding her utterance in the way she intended it to be understood. Can't we suppose that Mrs. Malaprop knows herself to mean *epithet* by "epitaph" and *arrangement* by "derangement"? In Davidson's view, not only is it conceivable, but it is common: we use the same words to mean different things and our hearers have no problem understanding what we say (see Section 2.1.3.c). Mrs. Malaprop has noninferential knowledge of what she means by "epitaph", which is *epithet*. Hence, we can say that Mrs. Malaprop cannot improve on the following sort of statement:

I mean *epithet* by "epitaph".

Whatever she means by uttering the word at the time of the utterance, she has noninferential knowledge of it. However, the interpreter has no immediate reason to take either of the following statements as his *best* interpretation of Mrs. Malaprop's utterance of "epitaph":

Mrs. Malaprop means *epitaph* by "epitaph".

Mrs. Malaprop means *epithet* by "epitaph".

The interpreter might choose one of the above, or any other, interpretations of Mrs. Malaprop's utterance and check its plausibility against further evidence. However, this is not what happens in the case of Mrs. Malaprop's self-ascriptions of meaning: she does not know herself by means of interpretation and induction. This by itself can show that the asymmetry cannot be plausibly explained in terms of the disquotational argument. In other words, some self-ascriptions of meaning are not disquotational so that the "Disquotational Argument" cannot be invoked to explain first-person authority with respect to them. We can conclude that the disquotational explanation of the asymmetry is not, and I think cannot be, Davidson's actual explanation of the asymmetry in general. Hence, Wright's reading of Davidson's account of the asymmetry is not acceptable.

5.4.3. Is Wright's Objection Plausible?

The next question is: if Davidson's explanation of the asymmetry is *not* based on the difference between the availability of evidence for the speaker and the interpreter, then is Wright's objection to Davidson's account plausible? The dilemma Wright introduced for Davidson's account was: either the evidence – the speaker's assent to S – is expressive of what the speaker genuinely believes or not. If it is, then it is presupposed that the speaker has a special access to the evidence, of which Davidson does not provide an explanation. And, if it is not, then there is no essential difference between my knowledge of what I believe and your knowledge of what I believe, and we will thereby lose the asymmetry. Davidson, I think, accepts the first horn of the dilemma, and it can be argued that this is consistent with the availability of the relevant sort of explanation.

As we saw (in Sections 5.2 and 5.4.1), Davidson did not believe that there is an essential difference in the sort of things I can know about myself and the sort of things you can know about me. What I utter is an expression of what I genuinely believe, and what the interpreter will achieve, after his successful interpretation of my utterance, is a grasp of that belief of mine. The important point that we need to note is that, by making such a presupposition, Davidson does not beg the question, i.e. does not unjustifiably presuppose the asymmetry, contrary to what Wright claims. It is true that Davidson presupposes the asymmetry in meaning-ascriptions: I directly know what I mean, while the interpreter knows what I mean through interpretation and induction. And it is true that Davidson explains the basic asymmetry in belief-attributions in terms of the above asymmetry in meaning-attributions: since I know what sentence I hold true and I directly know what I mean, I directly know what I believe. But, it is not by itself a question-begging way of explaining the basic asymmetry *if* there is an explanation of the asymmetry in meaning-ascriptions. We cannot rule out a putative explanation of a phenomenon by arguing that it presupposes the existence of the phenomenon in order to explain it. How could an explanation of something be given without presupposing its existence? Davidson's account would fail if all he did to explain the basic asymmetry in belief-attributions was to presuppose the second asymmetry in meaning-ascriptions.

Davidson's explanation of the asymmetry regarding the knowledge of meaning (and thereby the knowledge of beliefs), as we saw, was based on the claim that the speaker knows what he genuinely believes because he knows what he means, and he knows

what he means because assuming that he does is necessary for interpretation and understanding to be possible. The essential requirements for the possibility of interpretation by themselves explain, and justify the use of, the assumption that we are mostly right about what we mean by our words. And, if we are right about what we mean by our words, in Davidson's view, we are right about what we believe. Moreover, since my assent to a sentence expresses what I genuinely believe, the interpreter will achieve knowledge of what I genuinely believe too through her successful interpretation of what my utterance means.

Therefore, we can conclude that Wright's reading of, and his objection to, Davidson's account is not plausible. (1) His reading misses the main explanation Davidson proposed of the asymmetry; it fails to distinguish between the descriptions Davidson offers of the asymmetry and the explanation he suggests of it. (2) For a similar reason, Wright's objection (that Davidson's account is question-begging) is implausible, since Davidson's account would be so only if it took the asymmetry for granted without explaining it.

5.5. Objections to Davidson's Account of Self-Knowledge

Although Wright's objection to Davidson's account was argued to be implausible, I think Davidson's account fails due to some other problems.

One problem with Davidson's account of self-knowledge can be put as follows: the requirement that without taking the speaker to be right about his beliefs there would be no interpretation implies that, in understanding others, our default view must be that they are right about what they mean and believe. Nonetheless, one may make a weaker claim, according to which for interpretation to be possible, we must take speakers *as if* they are always right about what they mean and believe. In this sense, however, there is no guarantee that they are *really* right. Although this weaker claim would allow interpretation to proceed, it does not explain the asymmetry in the way we expected: it does not say anything about whether the speakers genuinely know what they mean and believe. In other words, we take their utterances *as if* they are expressive of their genuine beliefs, but there is no claim as to whether their utterances *are* expressive of what they genuinely believe. The problem is that Davidson's account of first-person authority is compatible with this weaker claim, since one may urge that, in order to

accommodate the asymmetry, all we need to do is to take others *as if* they are right about what they mean and believe. Interpretation can proceed even by taking for granted this weaker claim. But, does it imply that the asymmetry really exists? It does not seem so. As an analogy, for example, consider a game, for playing which it is constitutive that I assume that you are a giant. I can start playing the game only if I imagine that you are a giant. Nevertheless, it does not mean that you really are one; rather, I take you *as if* you are a giant. The same can be held true in the case of Davidson's account of first-person authority: I can play the game of interpretation only by taking you *as if* you are mostly right about your beliefs, since, without this assumption, interpretation would not be possible. This, however, does not explain whether you are, as the first person, really right about your beliefs. Therefore, since Davidson's account of self-knowledge is compatible with this weaker claim, Davidson does not succeed in his account.

The other problem concerns Davidson's account of error. According to Davidson, first-person authority does not imply the infallibility of first-person attributions so that I may think I meant something specific and thereby believed something specific, while I actually failed to mean or believe anything (see Section 5.2.). The reason for this claim has its roots in Davidson's criterion for success in communication: I have to be successfully interpreted by my interpreter if I am to be said to mean anything at all, and I may fail to provide enough clues and evidence for such an intended interpretation to be achieved by the interpreter. If I fail to be interpreted as I intended, then there is simply nothing to be known or believed. However, as discussed in Chapter Two (Section 2.1.7), such a criterion for the speaker's success in her practice of meaning something by her utterance does not necessarily imply that the failure in interpretation is always due to the speaker's failure in providing enough clues and evidence. The speaker, we can imagine, provides enough clues and evidence for her intended interpretation, while it is the interpreter who fails to properly employ them to reach the speaker's intended interpretation. In this case, is it plausible to claim that the speaker fails to mean something by her words, that she does not know what she means and believes? It seems that, in such cases, although interpretation *fails*, the speaker has to be treated as knowing what she means and believes.[111] But, can't we treat any failure in communication in this way? Davidson does not develop, and indeed it is hard to see

[111] Davidson might respond to this objection by taking the interpreter to be immune to such failures. In Section 2.2.3.a (pp. 94-95), we argued that this is not the way Davidson has actually treated the way two people communicate with each other.

how he could develop, a criterion distinguishing between the cases in which the speaker fails to provide enough evidence and the cases in which the interpreter fails to properly utilize the available evidence. If that is true, then Davidson's account of self-knowledge seems to face a problem: this account fails to explain the fallibility of self-knowledge. In other words, if the fallibility of self-knowledge is supposed to be explained in terms of Davidson's criterion for success in interpretation, then this criterion falls short of explaining such fallibility, since it will lead to the cases in which, although interpretation has supposedly failed, we intuitively take the speaker to know what she means and believes.[112]

The next problem with Davidson's account can be put as follows: even by granting Davidson's explanation of the asymmetry, there is still another step that Davidson must take in order to complete his explanation, i.e. Davidson's explanation as it stands is incomplete. Davidson explains the basic asymmetry regarding belief-ascriptions in terms of the asymmetry regarding meaning-ascriptions. The latter is explained by Davidson in terms of the necessary conditions on the possibility of interpretation. However, one might ask: What is Davidson's reason for the claim that if speakers did not know what they mean and believe, then interpretation would not be possible? Indeed, we need an explanation of why interpretation would not be possible if it was not the case that speakers have noninferential knowledge of what they mean.

Davidson, I think, would respond to this objection by drawing our attention to the process of radical interpretation. As discussed in the Appendix to Chapter Two, Davidson believes that the possibility of radical interpretation depends on some assumptions. Although he takes the ultimate evidence on which the interpreter must base her interpretation to be behavioural evidence and the goings-on in the speaker's environment, he believes that the interpreter must start her interpretation by presupposing that, when the speaker utters a sentence, she holds the sentence to be true on that occasion. The speaker's holding true attitudes were something that Davidson

[112] The following example may roughly clarify the point: Suppose that there are two interpreters interpreting my utterance. I believe that I have provided enough evidence and clues for them to interpret my utterance in the way I intended it to be understood. One of the interpreters successfully understands what I intend to mean by my utterance (and hence his communication with me is successful, since the evidence and clues are sufficient for him to reach the intended interpretation of my utterance), while the second interpreter does not (and hence his communication with me fails, since, for him, the evidence and clues are not sufficient). In this case and based on Davidson's criterion, do I know what I mean and believe? Davidson's explanation of the fallibility of self-knowledge in terms of his criterion for success or failure in communication does not seem to be capable of explaining this feature of first-person authority.

allowed the interpreter to know about in order to be able to start his interpretation. However, for Davidson, a speaker holds a sentence true partly because of what she means by the sentence and partly because of what she believes to be the case. For Davidson, propositional attitudes, such as beliefs and intentions, and meanings are deeply interrelated. Now, the question is: Why does the speaker utter one string of words rather than another? We can take Davidson to be claiming that the speaker chooses the specific words she does because of what they mean to her and because of the belief she wants to express by uttering those words. For instance, the speaker utters "Snow is white". According to Davidson, the speaker utters this sentence because of what these words mean for her and because of what she believes to be the case. If the speaker did not know what she means by those words and what beliefs she intends to express by them, then what reason would she have for choosing the words she uttered? What reason would the interpreter have to take those uttered words to have any meaning at all? In Davidson's view, even when we utter some words in the absence of any audience, for example, when we say to ourselves "What a nice day!", "yet it matters what words are used, what they mean. ... there must be some *reason* for using those words, with their meaning, rather than others" (1984b, p. 272). If the speaker did not directly know what she means by the words and what she believes about the world, then it would not be clear why she made that utterance rather than another and how her utterance is supposed to be interpreted by the interpreter in a certain way. Interpreting the speaker depends on taking her to know what she means and believes and to have intelligible beliefs about the world.

The interpreter is able to interpret the speaker's utterances only if he can attribute intelligible beliefs to the speaker, and he can attribute intelligible beliefs to the speaker only if he can interpret the speaker's utterances correctly. The interpreter can interpret the speaker only if he can take the speaker to know what she means and believes and to have beliefs that are mostly right. Without these presuppositions, there would be no ground for the interpreter to start his interpretation, since there would be no answer to the question: Why did the speaker utter those words? What did she express by uttering those words? The possibility of radical interpretation, hence, relies on the application of the principle of charity (and other relevant principles).[113] Mere observation of the speaker's behaviour, e.g. that the native speaker responds by "Gavagai" to the presence

[113] I have introduced this principle, and the related principles, in the Appendix to Chapter Two.

of a rabbit in view, cannot by itself lead the interpreter to any intelligible interpretation of the speaker, unless he can take the speaker as saying something intelligible and having intelligible beliefs about the goings-on in her environment. Putting the holistic feature of interpretation and the formal constraints on the Davidsonian truth-theory to one side, what allowed the interpreter to conclude that the native means *rabbit* by "Gavagai" was the fact that he could take the speaker to have intelligible beliefs about the happenings in her environment and to know the beliefs she has. Therefore, the application of the principle of charity, which is necessary for interpretation to be possible, implies that speakers know what they mean and what beliefs they express by their utterances. Without these assumptions, there would be no interpretation possible at all, as Davidson states,

> what is impossible is that she [the agent] should be wrong most of the time. The reason is apparent: unless there is a presumption that the speaker knows what she means, i.e. is getting her own language right, there would be nothing for an interpreter to interpret. (1987, p. 38)

What we were after here was to find an explanation of Davidson's claim that if interpretation is to be possible, then the asymmetry between the attributions of meaning must exist. We saw that Davidson could argue that if the interpreter is to be able to start his interpretation, he has to apply the principle of charity, according to which speakers must be taken to be expressing what they believe by choosing the words they utter, i.e. the words that have specific meanings for them. Unless they know what they mean and hence what they believe, interpretation would not be possible.

An important worry about Davidson's explanation, however, can be put as follows: let's assume that Davidson's proposed explanation of the asymmetry is successful. But, this explanation is an explanation of *why* the asymmetry exists. Has Davidson provided us with an explanation of *how* the asymmetry exists? Davidson's suggested explanation states that the speaker has to know what he means, that is, that the existence of the asymmetry is *necessary* for the possibility of interpretation. But it does not tell us how it can be that the speaker has noninferential self-knowledge.[114]

[114] In this regard, contrast Davidson's approach with Wright's: Wright argues that we must credit speakers with authoritative first-person knowledge in order to avoid KW's sceptical paradox. But, in addition, he develops the judgement-dependent account of intention and meaning to show *how* authoritative and noninferential first-person knowledge is possible. He criticizes McGinn (1984) for failing to provide an account of self-knowledge in his response to KW's sceptical argument. McGinn in order to support his claim that we have some "conception of how the primitive, non-experiential state of

Considering the objections discussed above, it seems that Davidson's account of self-knowledge is not successful. If that is the case, then it does not seem that Davidson can resist KW's sceptic's argument from queerness by appealing to his solution to the problem of self-knowledge, since our hope was to take Davidson's account of first-person authority as capable of providing an answer to the sceptic's claim that our knowledge of the alleged primitive state of meaning something by a word is mysterious.

Wright, however, has proposed his own explanation of the asymmetry, which is different from Davidson's and provides him with a response to KW's sceptic's argument from queerness. It is now important to look at Wright's own explanation of self-knowledge and the way he responds to KW's sceptic because not only will it help us to see how a judgement-dependent account of meaning and intention can plausibly support a non-reductionist response to KW's sceptic's sceptical problem, but it will also assist our attempt to find a similar kind of account in Davidson's works. Indeed, as Davidson himself states, his alternative solution to KW's sceptic's sceptical problem, put in terms of the requirement of "being interpreted as intended", has been somehow inspired by Wright's remarks on intention: "Essentially these points about intention are made by Crispin Wright in attempting, like me, to defuse Kripke's view that he has extracted an essentially insoluble 'skeptical paradox' from Wittgenstein's treatment of meaning" (1992, p. 112, fn. 6).

5.6. Wright's Response to KW's Sceptic

Wright, in his paper "Kripke's Account of the Argument against Private Language" (1984), raises an objection to KW's sceptic's sceptical argument, according to which this argument relies on the assumption that knowledge of what we mean by our words is inferential. In other words, KW's sceptic presupposes a sort of reductionist conception

---

meaning something is an object of distinctively first-person knowledge" appeals to "the ordinary idea of first-person authority for psychological states like – believing, thinking, intending, etc." (2002, p. 119). McGinn believes that "how to give a philosophical *theory* of this kind of knowledge is … a difficult and substantive question", but lack of such a theory does not lead "to doubt the *existence* of the phenomenon'" (Wright, 2002, pp. 119-120). Wright, however, thinks that this response is "an instance of philosophical stone-kicking" (2002, p. 120). For him, it is not enough to claim that states like meaning, intending, and the like are primitive states, without explaining our noninferential knowledge of the content of such states; rather, we need to an answer to the question "how is it possible to be, for the most part, effortlessly and reliably authoritative about, say, one's intentions?" (Wright, 2002, p. 120). In other words, McGinn has to answer the "how" question regarding first-person authority.

of meaning.[115] For instance, putting it in terms of our perceptual claims, the sceptic challenges the claim about my former perception that "Yesterday, I saw it raining". In order to justify the correctness of my former perceptual claim, the sceptic allows us to appeal to "any relevant fact so long as I do not thereby presuppose that there is such a thing as knowledge of what I formerly perceived – since it is of belief in the very existence of the genre of knowledge that the skeptic is demanding justification" (1984, p. 774). Hence, the sceptic argues that all the relevant facts, (except the fact that I perceived that it was raining), fail to constitute the fact that I perceived that it was raining. According to Wright, "the trouble, evidently, is the assumption that knowledge of a former perception has to be *inferential,* that the ultimate grounds for such knowledge must reside in knowledge of a different sort" (1984, p. 774). Similarly, in the case of meaning, the sceptic demands citing "facts from which the character of his former understanding of [the expression] E may be *derived.* And that is fair play only if knowledge of a *present* meaning has to be inferential; otherwise the skeptic is satisfactorily answered simply by recalling what one formerly meant" (1984, p. 774).

Wright, however, thinks that KW's sceptic's presumption is not justified; rather, we can potentially resist the sceptical challenge by bringing into play a sort of non-reductionism about meaning. But the question is: "If it is to be possible simply to recall the character of former meanings, can the requisite presupposition, that knowledge of present meanings may be noninferential, really be made good?" (1984, p. 775). If one attempts to defuse the sceptic's sceptical paradox by defending non-reductionism about meaning, then the problem of self-knowledge needs to be properly accounted for: Why do the avowals we make about ourselves have the unique and special features they do, such as being authoritative? Why do we know ourselves in such a special way?

KW's sceptic's discussion of non-reductionism – the view that meaning something by a word is an irreducible, *sui generis* state, which is not qualitative or introspectible – is brief. KW's sceptic's objection was that this response is a desperate attempt, which leaves the nature of such a primitive state mysterious (see Section 1.2.9). As Wright puts it, the sceptic claims that "it is utterly mysterious how such a state could have the requisite properties, in particular how, although a finite state realized in a finite mind, it

[115] As Wright clarifies, "Kripke's skeptic discounted the attempt to bring your previous general thoughts against unwelcome interpretations of your previous use of E [an expression], on the grounds that you thereby presuppose knowledge of the proper interpretation of those thoughts – which is, in detail, knowledge of the very putative species currently under suspicion" (1984, p. 774).

could nevertheless have the potential infinity of content that the normativity of meaning requires" (1984, p. 775). Hence, the difficulty is to provide a plausible answer to the following question: "How can there be a state which each of us knows about, in his own case at least, noninferentially and yet which is infinitely fecund, possessing specific directive content for no end of distinct situations?" (1984, p. 775). Wright, first of all, criticizes KW's sceptic for putting the above questions in such a way as "to leave the impression that it is rhetorical, that we have not the slightest idea what such a state might be" (1984, p. 775). Wright, on the contrary, believes that "a little reflection shows that both these features – noninferentiality and indefinite "fecundity" – are simply characteristic of our standard intuitive notion of *intention*" (1984, p. 775). Hence, he wants to respond to KW's sceptic by showing that the fundamental features of intentions, i.e. their noninferentiality and indefinite fecundity, are not problematic or mysterious.

Of the criticisms KW's sceptic made of non-reductionism, one was that it is mysterious how a finite mind can contain intentions to employ words in a specific way in a potentially infinite number of cases. Wright calls this problem, "the mysterious fecundity" of intentions (1984, p. 776). For Wright, however, there is no mystery in this feature:

> Well, suppose I intend, for example, to prosecute at the earliest possible date anyone who trespasses on my land. Then there can indeed be no end of distinct responses, in distinct situations, which I must make if I remember this intention, continue to wish to fulfill it, and correctly apprehend the prevailing circumstances. (1984, p.776)

As another example, if I intend to tell the truth on any occasion, then it does not matter whether the potential situations are finite or infinite in number. Our intentions having the general content they do must not be taken to be a mysterious feature of them. [116]

The second problem KW's sceptic introduced was about how, in each case, we can have direct knowledge of our intentions, or meanings, without introspection, inference,

[116] According to Wright, to make better sense of such infinite fecundity, we can consider the case of universal quantifiers. If we agree that our intention has a general content, then we should not be worried about the consequence of having such general content more than we are to be worried about universally quantified conditionals: "There is no more a puzzle about the "infinity" of this content [of our intention] than there is a puzzle about the capacity of any universally quantified conditional, $(x)(Fx \rightarrow Gx)$, to yield indefinitely many consequences of the form, *Ga, Gb,* . . . . , when conjoined with corresponding premises of the form, *Fa, Fb,* . . . ." (1984, p. 776).

and the like. According to Wright, as a matter of our ordinary conception of first-person authority, we usually take ourselves to have a special privileged access to the features of our own intentions. When we are asked about our own intentions, what we answer, putting aside the possibility of lying, slips of the tongue, self-deception, and the like, should be taken to be right and reliable. As he clarifies, "each of us regularly carries out intentional acts without necessarily thinking about what we are doing at all" (1984, p. 775). Hence, "we can in general make no ready sense of the question, 'How do you know?' directed at an avowal of intention" (1984, pp. 775-776). In this way, knowledge of our own intentions is not mediated by reflection in general.[117]

To have a clearer picture of Wright's view, we can consider the way he introduces the main intuitive characteristics of avowals by distinguishing between two sorts of such avowals. (I) The first group is called "*phenomenal avowals*" (2001b, p. 321), involving claims such as "I have a headache", "I feel tired", and so on. (II) The second kind of avowals are called "*attitudinal avowals*" (2001b, p. 322). The characteristics, processes, and states which such avowals concern "are partially individuated by the propositional content, or intentional direction, which they contain" (2001b, p. 322). Examples of such avowals are "I believe that term ends on the 27th", "I am frightened of that dog", and so on. The first group of avowals has three main marks. (1) "They are *groundless*" (2001b, p. 321), that is, we do not demand the person to give a reason or cite evidence for his claims about himself. (2) "They are *strongly authoritative*", that is, "if somebody understands such a claim, and is disposed to make it sincerely about themselves, that is a guarantee of the truth of what they say" (2001b, p. 321). Hence, any doubt about these claims would have its roots in either the sincerity of the person or his understanding of it. Granted that the subject is credited with sincerity and understanding, the subject's making such claims about himself is "a criterion for the correctness of the corresponding third personal claim made by someone else" (2001b, p. 321). (3) Phenomenal avowals display a sort of "*transparency*" (2001b, p. 321). It is absurd for someone to say "I don't know whether I have a headache". If we make such avowals, we are clear about what we avow. The second sort of avowals, i.e. attitudinal avowals, display the three features of phenomenal avowals too, except that the authority they

[117] As Wright emphasizes, "to come to know that you have a certain intention is not to have it dawn on you that you have an intention of *some* sort and then to recover an account of what the intention is by reflecting upon recent or accompanying thoughts. It is the other way round: you recognize thoughts as specifying the content of an intention that you have *because* you know what the intention is an intention to do" (1984, p. 776).

carry is weaker because of the possibility of self-deception and confusion: "To the extent that there is space for relevant forms of self-deception or confusion, sincerity-cum-understanding is no longer a guarantee of the truth of even basic self-ascriptions of intentional states" (2001b, p. 324). We can imagine situations in which the subject, for instance, deceives himself or has other relevant forms of confusion about his attitudes; in these cases, we do not simply take the first-person avowals as a guarantee for their truth. Attitudinal avowals, hence, display a weaker authority: they provide "justification for the corresponding third-personal claims" (2001b, p. 324). If the subject makes such avowals, and there is no obvious evidence of self-deception or confusion, then we are justified in regarding his avowals as true.[118]

Therefore, Wright believes that our intuitive notion of intention implies that "a subject has, in general, authoritative and noninferential access to the content of his own intentions, and that this content may be open-ended and general, may relate to all situations of a certain kind" (1984, p. 776). For him, such a notion of intention, if it can be successfully defended, can deal with KW's sceptic's sceptical challenge: "The skeptical argument is powerless against it" (1984, p. 776). Wright puts forward his general proposal as follows:

> In order, then, to rebut the skeptical argument, it would have sufficed, at the point where the skeptic challenged you to adduce some recalled mental fact in order to discount the grue-interpretations, to recall precisely your former intention with respect to the use of 'green'. To be sure, any *specification* that you might give of the content of that intention would be open to unwelcome interpretation. But, if you are granted the intuitive notion of intention, you can reply that you do not in any case know of the content of an intention via a specification of it; rather, to repeat, you recognize the adequacy of the specification because you know of the content of the intention. (1984, pp. 776-777)

Wright thinks that this proposal can successfully defeat KW's sceptic's sceptical argument.[119] However, again, such a proposal needs to be properly supported, especially against KW's sceptic's argument from queerness. That is to say, we need "to validate our claim to noninferential authority for our present (and previous) intentions without succumbing to the mythology of infinite, explicit introspectible content"

[118] For Wright having such special authority over our attitudes is a condition on being an intentional subject at all: "Wholesale suspicion about my attitudinal avowals – where it is not a doubt about sincerity or understanding – jars with conceiving of me as an intentional subject at all" (2001b, p. 325).

[119] As Wright says, "the skeptical argument has absolutely no destructive force against that proposal" (1984, p. 777).

(Wright, 1984, p. 778). The problem is to explain "*why* avowals display the marks they do" or "how is it possible for subjects to know these matters non-inferentially" (2001b, p. 330).[120] Unless we can provide plausible answers to these questions, we have not yet defused KW's sceptic's sceptical problem: "One way or another we have to answer, or undermine, the question: how is first-person authority for intentional states possible? Until we do, we have not answered the Skeptical Argument" (Wright, 1989, p. 114).

Wright leads us to his own answer to the above questions by considering the natural, but highly problematic, response from Cartesianism, according to which the avowals "are the product of the subject's exploitation of what is generally recognized to be a position of (something like) *observational privilege*" (2001b, p. 331). Such Cartesianism explains the asymmetry between the way we know ourselves and the way others can know about us by claiming that "the truth values of such utterances [or avowals] are noninferentially known to the utterer via her immediate awareness of events and states in a special theatre, the theatre of her consciousness, of which others can have at best only indirect inferential knowledge" (2001b, p. 332). According to Wright, this response, however, has its own "unaffordable costs" (2001b, p. 333). In other words, it gives rise to the problem of other minds: the subject is the only one who enjoys such a special authority over his own mind, in the special realm of his consciousness, and such knowledge is not at all transferable to others. Others at best can indirectly and inferentially know about the subject's attitudes. But, how can we be sure that there are other minds like us if what we can know about them is always indirect, that is, if the way in which we know about our minds is radically different from the way we know about other minds?

According to Wright, the later Wittgenstein offered an even deeper problem for Cartesianism, which shows that it is completely hopeless: "Cartesianism is based on a *grammatical* misunderstanding, a misinterpretation of the language-game of self- and other-ascription of mental states" (Wright, 2002, p. 135). The grammatical role of such first-person avowals deceives us into thinking that there is a special realm, in which we ourselves can privately know what we think and believe. Wittgenstein thought that "the

[120] Wright puts this question in a more general form: "What is it about [the avowals'] subject matter, and the subject's relationship to it, which explains and justifies our accrediting her sincere pronouncements about it with each of Groundlessness, Strong Authority and Transparency in the case of phenomenal avowals, and with Groundlessness, Weak Authority and Transparency in the case of attitudinal avowals?" (2001b, p. 330). It is this "how" question which, at the end of the previous section, I accused Davidson of neglecting.

Cartesian picture attempts a lay-philosophical explanation of something which needs no explanation – aspects of the 'grammar' of avowals: the rules of the ordinary psychological language game" (Wright, 2001d, p. 442). Therefore, according to Wright, if, following the Cartesian picture, we treat the avowals as reports of some self-standing subject matter, we face a serious problem. On the one hand, what is it that determines the truth or falsity of such reports? The events, or states of affairs, which are supposed to make such avowals true or false, according to the Cartesian picture, are available privately only to the subject alone. Hence, the subject can take anything he wishes to be what makes his avowals true: there would be no criterion for assessing the truth or falsity of such avowals. On the other hand, if we want to avoid such privacy, then we should consider publicly available states of affairs: the state of affairs must be items which are fully manifestable and available to public observation. However, in this case, our view takes a behaviouristic shape, and consequently, as Wittgenstein says, "it looks as if we had denied mental processes" (Wright, 2001c, p. 359). How can we properly account for the subject's first-person authority over his mental states by appealing to the behaviour of the subject? The intuitive asymmetry, which we hoped to explain, seems to disappear, since "as soon as you go public, it becomes obscure what advantage selves can enjoy over others" (Wright, 2001c, p. 359). Therefore, as Wright concludes, we face a dilemma: "*Either* accept the Cartesian view, which cannot accommodate ordinary knowledge of others, *or* accept some form of externalization – perhaps behaviourist, more likely physicalist – which cannot sustain the special place of self-knowledge" (2001c, p. 359).

Wittgenstein's own view, as Wright sees it, is to undermine the validity of the "*entire explanatory project*" (2001c, p. 366). This is how Wittgenstein's quietism manifests itself. Instead of meeting the demand of theorizing about, and finding a philosophical explanation of, the features of our avowals, we only need, as philosophers, to reveal the misunderstandings leading us to such a demand. We faced such problems *because* we wanted to explain why our avowals have the features they do. For providing such an explanation, we appealed to problematic philosophical views, such as Cartesianism, behaviourism, physicalism, and the like. We think there should be something deeper, on which such features rely. Instead, in Wittgenstein's view, we can see the avowals' having such features as primitively constitutive of their being psychological claims, which we ordinarily make of ourselves. In this sense, we can count the asymmetries, which we have considered so far, as what "belong primitively to the 'grammar' of the

language-game of ordinary psychology, in Wittgenstein's special sense – 'grammar' which 'is not accountable to any reality' and whose rules 'cannot' be justified by showing that their application makes a representation agree with reality'" (Wright, 2001c, pp. 366-367). The resulting view, which emerges from Wittgenstein's remarks, is what Wright calls "the Default View":

> [T]he authority standardly granted to a subject's own beliefs, or expressed avowals, about his intentional states is a *constitutive principle*: something which is not a consequence of the nature of those states, and an associated epistemologically privileged relation in which the subject stands to them, but enters primitively into the conditions of identification of what a subject believes, hopes and intends. (2001c, p. 368)

In this way, it is a primitive and constitutive feature of our psychological claims that our opinion about what we believe, intend, and the like, are "default-authoritative and default limitative" (2001c, pp. 368-369).

Wright, however, does not accept Wittgenstein's quietist approach to the problem of self-knowledge; he does not agree that we should just abandon the search for an explanation of this phenomenon, since, if we accept that the Default View is an option, then

> the immediate instinct is to ask what might *justify* the idea that it is the whole truth? That is the instinct to attempt to understand when and why it is a good move to dismiss the attempt to understand. To succumb is to reenter the space of explanatory philosophy. To resist is to have no reason for the Default View. (2001c, p. 373)

Wright attempts to provide his own explanation of why our avowals have the characteristics they do, and how the best opinions, which the first-person, under ideal conditions, forms about his intentions, can be plausibly treated as constituting the facts about what he intends. In the following section, we will consider Wright's judgement-dependent account of meaning and intention.

5.7. Wright's Judgement-Dependent Account of Intention/Meaning

In "Meaning and Intention as Judgement Dependent" (2002), Wright is concerned with "the relation between *best* judgements – judgements made in what are, with respect to their particular subject-matter, *cognitively ideal* conditions of both judge and circumstance – and truth" (2002, p. 130). His concern, in other words, is to account for

the relationship between the best opinions we can make about a subject matter, under the ideal conditions concerning our abilities and the circumstances under which the opinions are formed, and the truth of our judgements about that subject matter: whether what makes them true is independent of our best opinions, or whether it is our best opinions that make them true. Now suppose that a subject, S, under the ideal conditions, forms some opinion about a subject matter, say, the colour of an object. Such conditions, [121] for the case of colour, can be put as follows: "Conditions like those which actually typically obtain out-of-doors and out-of-shadow at noon on a cloudy summer's day" (Wright, 2002, p. 132). The subject must also be taken to be equipped with normal perceptual apparatus or equipment, i.e. the "equipment which is actually statistically usual among human beings" (2002, p. 132). These conditions are the best conditions under which a subject can judge whether an object is, for instance, red. In this regard, Wright distinguishes between two ways of viewing judgements that are made under these optimal conditions: we can view them as "*extension-determining*" or we can view them as "*extension-reflecting*" (2002, p. 130). As Wright clarifies,

> it is presumably necessary, in order for our judgements, appropriately constrained (partially) to *determine* the extension of some concept, that it be *a priori* true that the concept applies when, so constrained, we judge that it does. And it is, I suggest, *a priori* true that [when certain conditions are met], the fact of the object's colour … and the subject's judgement of the fact will, as it were, covary. (2002, pp. 130-131).

In other words, suppose that we find out that a subject's best opinions about the colour of an object covary with the facts about the object's colour so that "it is true that if the C-conditions obtain, a suitable subject will judge that *x* is red if and only if *x* is red" (Miller, 2007a, p. 229). Now, this covariance between the subject's best opinions about the colour of the object and the facts about the colour of the object can be accounted for in two ways: either the best opinions play at most a tracking role or they play a constitutive role. If the best opinions play the first role, then they are at best good at tracking the facts about the colour of the object, i.e. its being red. Our best opinions, in this case, only reflect the independent facts about the object's colour: "The determinants of a judgement's being true and of its being best have to be somehow independent" (Wright, 2002, p. 130). Our best opinions, hence, are *extension-reflecting* in such cases, that is, they merely reflect, rather than constitute, the extension of a colour predicate, such as "… is red". However, our best opinions and the facts about redness may covary

[121] Such cognitively ideal conditions are also called "C-conditions" (see Wright, 2002, p. 132).

because our best opinions determine or constitute, rather than just reflect, the extension of the colour predicate, in which case they play an *extension-determining* role: it is our best opinions made under ideal conditions that determine which objects have which colours. The first way of explaining the covariance between our best opinions and the facts about the subject matter, namely, our best opinions as extension-reflecting, leads to the view that the predicate, in the discourse under study, is *judgement-independent*. The latter explanation, however, leads us to the conclusion that the predicate is *judgement-dependent*.

Now, a predicate, such as a colour predicate, is judgement-dependent if the following *provisional equation*, i.e. the equation that "under certain conditions, C, a subject will hold a certain belief if and only if it is true" (Wright, 2002, p. 131), satisfies four conditions. For example, the equation, in the case of redness, can be put as follows:

> For any *x*, if C-conditions hold, then a suitable subject, S, judges that *x* is red if and only if *x* is red.[122]

Now, the quality of being red can be said to be judgement-dependent, (I), if and only if the provisional equation is *a priori* true, and, (II), if the C-conditions can be specified *non-trivially*, (III), in such a way that the question about whether they are satisfied is *independent* of facts about redness, and, (IV), if there is no better explanation regarding why the covariance between best opinion and facts is *a priori* and substantial in this way (see Wright, 2002, 2001e; and Miller, 2007a, pp. 229-234, 2009, 2013b; Miller & Divers, 1999).

Wright employs an application of these ideas to the case of intention to defend a view, according to which "subjects' best opinions about their intentions, *both past and present*, are properly conceived as provisionally extension-determining and which explains how and why the opinions which they typically hold are indeed best" (2002, p. 140). To see whether our best opinions about our intentions are extension-determining, and hence whether the facts about our intentions are judgement-dependent, we should specify the provisional equation in the case of intention and see whether it satisfies the four mentioned conditions: if it satisfies them, then we have a judgement-dependent

[122] We can see that Wright's judgement-dependent account is not reductionist. The reason is that, in the above equation which is central to this account, the term "red", or the concept of red, appears on both sides of the biconditional. Hence, no reduction of this concept is proposed in this account.

account of self-ascriptions of intentions, on the basis of which we can provide a response to KW's sceptic's argument from queerness. First of all, for the case of intention, we can specify the C-conditions as follows: "The subject S is not lying, is prey to no material self-deception, is not making a simple slip of the tongue, has an adequate grasp of the concepts requisite for the expression of the intention, and is adequately attentive to the question of the content of his intention" (Miller, 2007a, p. 230). These are the best conditions under which the subject's avowals about his intentions can be made. Now, if one wishes to show that the subject's best opinions about his intentions are not extension-determining, one should show that at least one of the conditions on the provisional equation does not hold. But once we characterize the equation for a subject's intentions, we can see that it is *a priori* true and non-trivial, and satisfies the rest of the conditions Wright has introduced. The provisional equation can be specified in this way:

> For any subject, S, if the C-conditions hold, then S intends to φ if and only if S judges that he intends to φ.

This equation satisfies all the mentioned four conditions. (I) If the C-conditions are met, then it is *a priori* true that the facts about the subject's intention and the subject's judgements about the facts covary. In other words, if the C-conditions can be formulated in a way to prevent the cases of self-deception, confusions, misunderstandings, and the like – which make the subject unable to realize his intention – then it will be *a priori* true that whenever the subject, under ideal conditions, judges that he has a specific conditions, he has that intention, and *vice versa* [123] (see Wright 2002, p. 136-137). (II) The C-conditions are not trivially specified, since it does not appeal to "whatever-it-takes", that is to say, it is not presupposed in the specification of the C-conditions that the subject, under the C-conditions, can anyway make true

[123] As Wright says, "the truth, if it is true, that the extensions of colour concepts [and, in our case, the concept of intention] are constrained by idealised human response – best opinion – ought to be available purely by analytic reflection on those concepts, and hence available as knowledge *a priori*' (1992, pp. 116-117). Wright asks, "how might a subject who had the conceptual resources to form a belief appropriate to the presence, or absence, of a certain intention [i.e. possession of the concept of intention], nevertheless fail to do so?" (2002, p. 136). For Wright, one possibility is that the subject is self-deceived; another is that the subject is not "appropriately attentive to the question what his/her intentions are" (2002, p. 136). Now, if the subject can be said to have a "grasp of the appropriate concepts, lack of any material self-deception or anything relevantly similar, and appropriate attentiveness" (Wright, 2002, p. 136), then it will be *a prior* true that, under ideal conditions, the subject has an intention if and only if he judges that he has that intention.

judgments about his intentions (see Wright, 2002, p. 136). (III) The C-conditions can be satisfied independently of the facts about whether the subject has particular intentions. The facts about whether the C-conditions are satisfied are independent of the facts about S's intending to φ. (IV) And, finally, there is no better explanation of the fact that S's best opinions and the facts about S's intentions covary.[124] Therefore, Wright comes up with the claim that, for our judgements about our intentions, "there is no distance between being true and being best; truth, for such judgements, is constitutively what we judge to be true when we operate under cognitively ideal conditions" (2002, p. 130).

In this way, Wright could show that truths about intentions are judgement-dependent so that when we judge, under ideal conditions, that we have such-and-such intentions, we do have them, and our relevant avowals should be taken as authoritative and true. Wright believes that to employ this explanation to respond to KW's sceptic's sceptical problem regarding the concept of meaning we do not necessarily need to reduce the concept of meaning to that of intention. Rather, it would be enough to show that the concept of intention and the concept of meaning, as ordinarily conceived, are similar with respect to the characteristics they have, so that the provisional equation for the subject's meaning something specific by a word satisfies the aforementioned four conditions. Wright believes that "it is enough that the concepts are relevantly similar – that both sustain authoritative first-person avowals" (2002, p. 140). More specifically, in Wright's view, the concept of meaning and the concept of intention have similar features, i.e. they both exhibit first-person authority and "disposition-like theoreticity"[125], and viewing such concepts as judgment-dependent would help us to

[124] As we saw in the previous section, Cartesianism, as well as behaviourism, physicalism, and the like, as the alternative options for explaining this covariance, could not survive Wittgenstein's criticisms, according to Wright. And Wright rejected the Wittgensteinian therapeutic approach to this problem too.

[125] According to Miller (2007a), "intention … displays the same sort of "disposition-like theoreticity" as our intuitive conception of meaning: just as the ascription of an intention to a person may be withdrawn if he behaves in certain ways in the future, the ascription of understanding to him may be withdrawn if he applies the expression in certain bizarre ways in future situations" (2007a, p. 227). In other words, "just as an ascription of brittleness to a glass has to answer to how that glass would behave (crucially, whether it would shatter) in unactualised situations, so an ascription of understanding to a speaker has to answer to how that speaker would go on to apply the relevant expression (certain patterns of application would lead us to withdraw the original ascription of understanding, just as certain patterns of behaviour by the glass would lead us to withdraw the ascription of brittleness)" (2007a, p. 228). For more discussion, see (Wright, 2002, p. 138).

explain both such features.[126] As a result, Wright can use his judgment-dependent account to respond to KW's sceptic:

> It will be, similarly, a perfect answer to Kripke's Skeptic to explain how judgements concerning one's own meanings, both past and present, are likewise provisionally extension-determining in the most ordinary circumstances. Challenged to justify the claim that I formerly meant addition by 'plus', it will not be necessary to locate some meaning-constitutive fact in my former behaviour or mental life. A sufficient answer need only advert to my present opinion, that addition is what I formerly meant, and still mean, and to the *a priori* reasonableness of the supposition, failing evidence to the contrary, that this opinion is best. (2002, p. 140)

Therefore, Wright takes a completely different path, from Davidson's, to argue against KW's sceptic. Wright's account could establish a non-reductionist response to KW's sceptic's sceptical problem and account for the problem of self-knowledge. The explanation Davidson has actually provided, however, was not satisfactory. But the question is whether there is, in Davidson's writings, potential for providing a similar sort of judgement-dependent account of meaning. What form would Davidson's judgement-dependent account take if he wished to offer such an account?[127] In the next section, we will search for an answer to these questions by looking at the way Davidson accounts for the notion of intending. What is it to intend to do one thing rather than another? What is it that makes it the case that a subject intends to do one thing rather than another?

---

[126] According to Wright, "we may think of meaning so-and-so by [an expression] E as either consisting in, or relevantly similar to, possession of an intention or complex of intentions" (1987, p. 127). As Miller (2007a) explains, "If meaning such and such by an expression can be taken to consist in the possession of a certain sort of intention then the application of Wright's judgement-dependent account of intention to the case of meaning will be straightforward. But even if this is not the case, we can still construct something like the story we constructed about intention for the case of meaning" (2007a, p. 232). KW's sceptic himself thinks of meaning and intention as playing a relatively similar role: "Granted that I mean addition by '+', then of course if I were to act in accordance with my intentions, I would respond, given any pair of numbers to be combined by '+', with their sum; but equally, granted that I mean quaddition, if I were to act in accordance with my intentions, I would respond with the quum" (1982, p. 28). He emphasizes, "Wittgenstein questions the nexus between past 'intention' or 'meanings' and present practice" (1982, p. 62). See also (Kripke, 1982, pp. 54, 39, 37). For more discussion of the characteristics of intention and meaning, see (Wright, 1987, pp. 125-127). See also (Miller, 2007a, pp. 226-234).

[127] We should note that Wright's judgement-dependent account is controversial. There are many objections to this account, which I do not have space to consider. For some objections, see e.g. (Miller, 1989, 2007b), (Boghossian, 1989, sections 29 and 30), (Edwards, 1992), and (Holton, 1993). Moreover, this means that if Davidson can be taken to be capable of providing a judgement-dependent account of meaning and intention, the next issue would be to investigate which of these criticisms is applicable to Davidson's alleged judgement-dependent account.

5.8. Davidson's Judgement-Dependent Account of Intention/Meaning

Davidson's actual alternative solution to KW's sceptic's sceptical problem, that is, adopting "being interpreted as intended" as the criterion for success in the practice of meaning something by an utterance, was rejected by KW's sceptic. For instance, suppose that a speaker utters, or assents to, a sentence, and suppose that he intends his utterance to be interpreted by his interpreter as meaning that *that's green*. The sceptic, however, asks: What is it that makes it the case that the speaker intended his utterance to be interpreted in one way, i.e. as meaning that *that's green*, rather than another, i.e. that *that's grue*? The speaker's assent to that sentence is apparently compatible with both sceptical hypotheses. The sceptic left Davidson with two difficult options: either introduce a fact about the speaker constituting the fact that he intends to be interpreted in one way rather than another, or take a non-reductionist approach to the problem, i.e. that it is a primitive, *sui generis*, fact that the speaker, in each case of communication, intends to be interpreted in one particular way, and properly support the idea that the speaker directly and noninferentially knows what he intends. The problem with the first option was that, if Davidson tries to reduce the meaning facts to any other sort of facts about the speaker, the sceptical problem will be run again at that level. Hence, again, whether or not Davidson *actually* embraces a non-reductionist response to the sceptic, non-reductionism seems to be the best, perhaps the only, option for Davidson to respond to KW's sceptic. But, if Davidson hopes to defend such non-reductionism, then he has to provide us with a plausible story about (I) what it is for an agent to intend to φ, and, in particular, to intend to be interpreted in a particular way; and (II) how the agent noninferentially knows what he intends, in a way which does not leave the nature of such knowledge mysterious.

Insofar as KW's sceptic's sceptical problem is concerned, we are interested in a metaphysical account of what it is to intend to do something, that is, Davidson's potential account, if any, of what constitutes the fact that an agent intends to do one thing rather than another.[128] We will briefly look at Verheggen's characterization of

[128] Regarding Davidson's remarks on intention, what we are concerned with here is not to investigate whether Davidson believes in the reducibility or irreducibility of the concept of intention. Davidson has always emphasized the irreducibility of mental concepts to physical concepts, or any other concepts: "Psychological concepts, I have been arguing, cannot be reduced, even nomologically, to others. But they are essential to our understanding of the rest. We cannot conceive a language without psychological terms or expressions—there would be no way to translate it into our own language" (1974b, pp. 243-244); "as I

Davidson's non-reductionism again and will investigate whether it can be developed along the lines of Wright's judgement-dependent account of meaning and intention. We will see that Davidson's remarks on the notion of triangulation would not have any clear application in such a desired judgement-dependent account of meaning. Then we will consider Davidson's account of intending, which has characteristics that might be suggestive of a judgement-dependent account of meaning and intention. In this regard, we will end with the suggestion that Davidson could resist KW's sceptic's sceptical problem by providing a proper non-reductionist response to it.

5.8.1. Verheggen's Reading of Davidson's Non-Reductionism Revisited

We saw (Section 5.1) that, for Verheggen, Davidson's argument from triangulation can be seen as a manifestation of a non-reductionist view about meaning: meaning can be fixed only by those who already have a language, that is, by those who have the concept of objectivity, the possession of which requires linguistic triangulation. It is, according to Verheggen's reading, the association of words with extra-linguistic items that gives meaning to our words, though, among endless such associations, only those bestow meanings on the words that have already been established by the speaker and understood by the interpreter. Only those "external facts", as Verheggen calls them, are counted as meaning facts that are chosen by the speaker, and the interpreter agrees to take them as what determine the meaning of the speaker's utterances. Therefore, something is the actual cause and, hence, the meaning-giving cause, of the speaker's utterance if the speaker takes it as the actual cause of his utterance, and the speaker and the interpreter agree to take it as the actual cause of the speaker's utterance. How can we mix this view with a proper sort of judgement-dependent account? One can characterize such a view by claiming that, under proper characterization of the C-conditions, the following provisional equation is *a priori* and non-trivial:

> If ideal conditions obtain, then I mean *table* by "table" if and only if I judge that my utterance of "table" is caused by tables.

How can this be taken to be a judgement-dependent view, which accommodates Verheggen's reading of Davidson's causal explanation of meaning-determination? The

---

am urging, psychological characteristics cannot be reduced to the others, nevertheless they may be (and I think are) strongly dependent on them" (1973b, p. 253).

main concern here is to find a place in the equation to locate the requirement of linguistic triangulation. One possible suggestion is to locate linguistic triangulation somewhere among the ideal conditions. If that is the case, then

> if I am not self-deceived, not lying, not confused, and the like, *and* I am in actual linguistic triangulation, then I mean *table* by "table" if and only if I judge that my utterance of "table" is caused by tables.

However, it does not seem plausible to demand that an agent must linguistically triangulate with another person in each case of making a judgment about the cause of his utterance. I cannot optimally judge that something is the cause of my utterance if I am self-deceived, or if I lie, when I am making that particular judgement. But it seems that I can make my best judgements about whether or not something is the cause of my utterance even if I am not *now* in actual linguistic triangulation with others. At best, actual linguistic triangulation can be regarded as what is required if a creature is to have the ability to make any judgement at all, such as having the necessary concepts, such as that of objectivity. If a creature, in enough cases, has been engaged in actual linguistic triangulation, so that it can be said to have possessed the concept of objectivity together with many other relevant concepts, then it can be said to have been equipped with what is required to make proper judgements about the cause of its utterances and thoughts, since, according to Davidson, there is no meaning and intention at all without linguistic triangulation. However, being in linguistic triangulation is not a condition that must obtain in order to make a particular judgement. Hence, it is not really clear what role linguistic triangulation must be playing in such a judgment-dependent account of meaning. One can directly aim at supporting a judgement-dependent account without necessarily bringing in the discussion of the notion of linguistic triangulation.

However, it seems that Davidson, in his account of intention, has attempted to provide us with such an account, which has the general characteristics of a judgement-dependent account of intention.

5.9. Davidson's Judgement-Dependent Account of Intending

In his paper, "Intending" (1978), Davidson offers an account of what it is to intend to φ. First of all, he distinguishes between an agent's intending to do something and his acting with an intention: "Someone may intend to build a squirrel house without having

decided to do it, deliberated about it, formed an intention to do it, or reasoned about it" (1978, p. 83). As Davidson mentioned before (see Section 5.1.1), intending to do something is not a qualitative state, and our knowledge of such a state is not introspectible, observational, or inferential; it is not equivalent to, or even necessarily accompanied with, doing what is intended, thinking about it in advance, forming a belief about it, and the like. If that is true, then what makes it the case that I intend to φ rather than something else cannot be explained by reducing it to my actions, thoughts, beliefs, and the like. It is, however, important to note that Davidson's target is to account for such pure intending by using materials from his account of intentional action. Hence, we need to start with his account of action with an intention.

5.9.1. Intentional Actions

For intentional action, Davidson's explanation of the action is given by describing the action as caused by certain propositional attitudes, which rationalize the action. If someone acts with an intention, he has a reason to do so: "A man who nails boards together with the intention of building a squirrel house must want to build a squirrel house, or think that he ought to (no doubt for further reasons), and he must believe that by nailing the boards together he will advance his project" (1978, p. 83). To put it in more detail, according to Davidson's general account of intentional action, "whenever someone does something for a reason, therefore, he can be characterized as (*a*) having some sort of pro attitude toward actions of a certain kind, and (*b*) believing (or knowing, perceiving, noticing, remembering) that his action is of that kind" (1963, pp. 3-4). For Davidson, pro-attitudes are attitudes such as desires, wantings, wishes, hopes, and the like.[129] In this way, he defines primary reason as follows: "Giving the reason why an agent did something is often a matter of naming the pro attitude (*a*) or the related belief (*b*) or both; let me call this pair the *primary reason* why the agent performed the action" (1963, p. 4). In Davidson's view, we can simplify our talk of reasons: "To know a primary reason why someone acted as he did is to know an intention with which the action was done. If I turn left at the fork because I want to get to Katmandu, my

[129] As Davidson clarifies, such pro-attitudes are taken to be "desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values in so far as these can be interpreted as attitudes of an agent directed toward actions of a certain kind" (1963, p. 4).

intention in turning left is to get to Katmandu" (1963, p. 7). Hence, the primary reason, or the intention, one may have for his action is given by referring to the beliefs and the pro-attitudes that he seems to have for his action. In this way, although the beliefs and the pro-attitudes that the agent has, as the reason for his action, may be false, unreasonable, or wrong, they can nonetheless provide an explanatory reason for that action "only if those attitudes are appropriately related to the action as viewed by the actor" (1978, p. 84). In other words, the beliefs and desires may not themselves be reasonable from our point of view, but the action should be viewed as reasonable by the agent, under those desires and beliefs, if they are to be considered as the reason for that action: "If someone acts with an intention, he must have attitudes and beliefs from which, had he been aware of them and had the time, he *could* have reasoned that his action was desirable" (1978, p. 85). Therefore, "a reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action" (1963, p. 3).

However, for Davidson, we face a problem if we want to fully characterize acting with an intention in this way, since the attitudes the agent has must be "appropriately related to the action"; but, someone might have the beliefs and desires which might give a good reason to perform an action with an intention, while the act is done unintentionally. If that is true, then having specific beliefs and desires is not *sufficient* to make the claim that the action is done by an intention. For instance, "someone might want tasty stew and believe sage would do the trick and put in sage thinking it was parsley; or put in sage because his hand was joggled" (1978, p. 87). Hence, it is not yet enough to simply cite a belief-desire pair to explain the action as intentional. As another example, suppose that someone has the desire to drink vodka and the belief that there is vodka in that glass on the table. As he is thirsty too, he first goes for another glass on the table, which he thinks is a glass of water. This glass, however, is actually full of vodka so that, by drinking it, he unintentionally drinks vodka. If this person is to be said to do something with a specific intention, he must do it *because of* the specific beliefs and attitudes he had for that action: "Central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the reason" (Davidson, 1963, p.9). Nonetheless, adding "because" will bring a new element into this account of action: it involves the notion of cause and the casual relation between reasons (beliefs and pro-attitudes) and actions.

In Davidson's view, the beliefs and pro-attitudes that the agent has for his action must be considered as the cause of the action, as he says, "the primary reason for an action is

its cause" (1963, p. 4).[130] Therefore, Davidson's account has so far turned into the claim that an action is intentional if a belief-desire pair causes the action. However, the problem regarding this causal explanation of action now emerges. As Davidson puts it, "an agent might have attitudes and beliefs that would rationalize an action, and they might cause him to perform it, and yet because of some anomaly in the causal chain, the action would not be intentional in the expected sense, or perhaps in any sense" (1978, p. 87). It is not the case that any causal relation between the belief-pro-attitudes pair and the action makes the action intentional. Davidson does not provide a clear example of this problem, but we can borrow his previous example and change it in a way to fit this case. Someone might want tasty stew and believe adding sage would make the stew tastier; he grabs the metal box of sage to put in the stew; however, the kitchen table is wet and the nearby toaster's electric wire has been corroded and worn out at the point it touches the table. He grabs the box of sage and is then shocked by the electric current passing through his arm; the electric shock, at the same time, makes his arm move in a way to position the sage box on the top of the stew so that the sage in the box is put in the stew. His act of putting the sage in the stew was caused by his having that specific desire and belief, but he eventually put the sage in the stew unintentionally: some anomaly in the causal chain – the causal chain from the belief-desire pair to his action – made his action unintentional at the end.

Therefore, what Davidson has come up with is at best an (incomplete) account of acting with an intention, according to which "an action is performed with a certain intention if it is caused in the right way by attitudes and beliefs that rationalize it" (1978, p. 87). It is the belief-desire pair causing your action *in the right way* that actually rationalizes your action. Although this view is in need of more clarification, especially regarding the clause "in the right way", it shows that acting with intention is not something mysterious that requires introducing a special kind of attitude; it is nothing but describing an action by citing certain attitudes, with which we are already familiar. Hence, what intentional action requires, according to Davidson, are "only

[130] According to Davidson, our actions are considered as intentional under specific descriptions of the action, and not intentional under the others: "I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home. Here I need not have done four things, but only one, of which four descriptions have been given" (1963, p. 4); however, "reasons may rationalize what someone does when it is described in one way and not when it is described in another" (1963, p. 5). In this way, "when we explain an action, by giving the reason, we do redescribe the action; redescribing the action gives the action a place in a pattern, and in this way the action is explained" (1963, p. 10).

desires (or other pro attitudes), beliefs, and the actions themselves" (1978, p. 87). There is no ontological gap between acting with an intention and having certain beliefs, desires and action: "The reduction is not definitional but ontological" (1978, p. 88).

This investigation, at the same time, shows that acting with an intention is different from *mere intending* to do something, namely, intending without an accompanied action. My action can be rationalized by describing it as caused by certain beliefs and desires. But what about my intending to φ without actually doing φ? What about my intending to φ without believing that φ? What is it that makes it the case that I intended to φ rather than ψ? As we saw, Davidson believed that intending to φ is not a matter of having a reason for φing, or going through a process of deliberation or reasoning: pure intending is "intending without conscious deliberation or overt consequence" (1978, p. 89). By investigating intentional action and the characteristics of pure intending, we found out that it is not easy to relate intentional action to pure intending: "Intending is a state or event separate from the intended action or the reasons that prompted the action" (1978, p. 89). However, how can these remarks help Davidson in his account of pure intending?

According to Davidson, we can suppose that if I intend to do something, I take doing that thing to be desirable. In other words, suppose that I want to eat something sweet; eating something sweet is thereby desirable for me. Nonetheless, the problem with the case of explaining pure intending is that there is no action to be evaluated, or to be reasoned about, in terms of certain causal relations between the action and the beliefs and pro-attitudes: "There is no action to judge simply good or desirable" (1978, p. 97). In Davidson's view, hence, "all we can judge at the stage of pure intending is the desirability of actions of a sort" (1978, p. 97). If intending to φ is just to judge, or to make the judgement, that doing φ is desirable, then the problem is that such judgements "do not always lead to reasonable action" (1978, p. 97). In other words, such judgements may lead to contradictory outcomes. To see the problem, we should find out why Davidson distinguishes between two sorts of judgements about our desirable actions: "prima-facie judgements" and "all-out judgements". This distinction will play a significant role in Davidson's account of intending.

5.9.2. Prima-Facie Judgements vs. All-Out Judgements

The characteristic of the first sort of judgements is that they are conditional. The reason for having such a feature is that they are often law like, for instance, the judgments like "any action that is an eating of something sweet is desirable" (Davidson, 1978, p. 97). This kind of judgement, as characterized in this way, will lead to contradictions because, for instance, "while holding it desirable to eat something sweet, we may also hold that it is undesirable to eat something poisonous" (1978, p. 98). In other words, an object can be both poisonous and sweet at the same time and, hence, the same desired action of eating that thing can be considered as the act of eating something sweet and eating something poisonous. The same action, hence, is both desirable and undesirable. To solve this problem, we need to add modifying conditions: something is desirable under certain conditions. In this way, an action is desirable, or good, etc. just in a certain respect, that is, "*in so far as* an action has a certain characteristic" (1978, p. 98). Davidson calls the "judgements that actions are desirable in so far as they have a certain attribute, prima facie judgements" (1978, p. 98). As a prima-facie judgement, I take drinking water to be desirable in so far as it is not dangerously polluted, or salty, or hot, and the like. Such prima-facie judgements, however, cannot be directly linked to desirable actions, since it is not enough to perform an action merely because it has a desirable feature. As Davidson clarifies, "we may not know how the agent got from his desires and other attitudes – his prima facie reasons – to the conclusion that a certain action was desirable" (1978, pp. 98-99). Having prima-facie judgments about φ is not enough to intend to φ as a desirable action. When we intend to do something, hence, our relevant judgements cannot be conditional: we cannot first deliberate about the conditions under which such an action is desirable. Rather, it seems that we merely judge, unconditionally, that doing that thing is desirable. When we intend to do something, we make certain sort of judgements about the desirability of doing that thing in a way such that no further considerations make it undesirable; otherwise, we would be unable to intend anything. As Davidson puts it, we cannot make plausible the judgement as to whether doing something is desirable by adding certain conditions, such as the condition that eating something sweet is desirable if that thing is not poisonous, since there would be endless other things, which may interfere and make that action undesirable:

> It would be a mistake to try to improve the statement of my intention by saying, 'I intend to eat something sweet, provided it isn't poisonous'. … There are

> *endless* circumstances under which I would not eat something sweet, and I cannot begin to foresee them all. (1978, p. 99)

If that is true, then the judgement we are after in the case of pure intending "cannot, therefore, be a prima facie judgement; it must be an all-out or unconditional judgement which, if we were to express it in words, would have a form like 'This action is desirable'" (1978, p. 98). Davidson thereby arrives at the account he was searching for, according to which, when I intend to φ, I unconditionally, without considering any conditions, judge that doing φ is desirable for me: to intend to do something is judging, unconditionally (unless something else happens that changes the agent's mind), that doing that thing is desirable. Therefore, as Davidson suggests, in the case of pure intending, "the intention simply is an all-out judgement" (1978, p. 99).[131]

However, we should note that, according to Davidson, while pure intending is not directly explained by citing beliefs and other pro-attitudes, it does not mean that they are irrelevant to it.[132] As Davidson concluded above, in order to properly account for pure intending, we want all-out, not prima-facie, judgments, which do not lead to any contradictory outcomes. My all-out judgement about the desirability of eating something sweet does not include eating a poisonous sweet thing, since I do not believe that I will eat such things: "I do not believe I will eat a poisonous candy, and so that is not one of the actions of eating something sweet that my all-out judgement includes" (Davidson, 1978, p. 99). Having granted this background of beliefs, when I make a judgement that doing φ is desirable, I just make the judgement, without generally reasoning, deliberating, forming specific beliefs, and the like, about it. It is important to note that, for Davidson, these background beliefs are *not* part of my intending to φ: my belief that I do not eat a poisonous candy "is not part of what I intend, but an assumption without which I would not have the intention. The intention is not

[131] Hence, according to Davidson's account, while it might be possible that I make all-out judgments on the basis of having specific intention to do φ, that is, in terms of having some reason (beliefs and pro attitudes), it is also possible that I, without deliberation or being conscious about any of my specific beliefs, desires, and the like, just intend to φ. As Davidson puts it, "forming an intention, deciding, choosing, and deliberating are various modes of arriving at the judgement, but it is possible to come to have such a judgement or attitude without any of these modes applying" (1978, p. 99).

[132] The second point about Davidson's account is that, contrary to the case of intentional action, in which an action is actually performed, pure intending is about a desirable action *in the future*. According to Davidson, intending to do something is nothing but judging "that any action of mine in the immediate future that is the eating of something sweet would be desirable *given the rest of what I believe about the immediate future*" (1978, p. 99).

conditional in form; rather, the existence of the intention is conditioned by my beliefs" (1978, p. 100).[133]

5.9.3. Intending as Judgement-Dependent

In this way, we reach the point where Davidson's account of intending is completed: "To intend to perform an action is, on my account, to hold that it is desirable to perform an action of a certain sort in the light of what one believes is and will be the case" (1978, p. 100). What makes it the case that I intend to φ is my overall judgment that doing φ is desirable, which is made in the light of the beliefs and pro-attitudes (reasons) I have in general. We can say, at this point, that Davidson's account of intending looks non-reductionist and judgement-dependent. For him, it is the agent's judgement about the desirability of doing something that makes it the case that he intends to do that thing. The facts about the agent's intending to φ are, thus, constituted by, and covary with, the facts regarding his overall judgements about whether doing φ is desirable for him. What if it is the speaker's intention to be interpreted in a particular way?

If meaning something by an utterance, according to Davidson, depends on the speaker's intending to be interpreted in a particular way, then, considering Davidson's account of intending, meaning something by an utterance is nothing but the speaker's making an all-out judgment about the desirability of being interpreted as such, which is made against a background of beliefs about the world, stimuli or causes in the environment, the interpreter's abilities, and so forth. A speaker means *green* by "green" if he intends his utterance of "green" to be interpreted in a particular way, i.e. as meaning *green*, and, for him, to intend this is to make an all-out judgement that being interpreted in that way is desirable. Thus, what the speaker means by his word, he noninferentially knows, since it is nothing but making a judgement of the sort discussed above. The speaker, however, may not intend his utterance to be interpreted as meaning

[133] For instance, suppose that I intend to sit down to rest. My intending to sit is my judging that sitting is desirable, which is made against the background of my other beliefs about the things I can sit on, my environment, my abilities, my body, and so on. I will not intend to sit on the clouds as I do not have the belief that I can sit on such things, or that I can fly, or that in the clouds there is a special sort of chair waiting for me. While my judgment is made against such a background of beliefs, it does not mean that these beliefs are part of my judgement so that I first consider such conditions and then intend to do it, since it will lead me to the realization of an endless number of conditions, under which I would not intend to sit. What I do, according to Davidson, is just to make the all-out judgement, against the background of all my beliefs, that sitting is desirable.

grue, since he does not believe that the interpreter is capable of reaching such an intended interpretation. As Davidson emphasized before, such judgements are made in the light of the speaker's background beliefs about the world, the interpreter, and so on. If the speaker, among others, has the belief that his interpreter is capable of understanding his utterance in a different way, e.g. as meaning *grue*, then the speaker may intend his utterance to be interpreted in that way. This is, we can now see, perfectly compatible with Davidson's earlier claim that "In speaking or writing we intend to be understood. We cannot intend what we know to be impossible" (1991b, p. 147).

Nonetheless, one may object that Davidson's account of intending cannot be properly connected with his account of the practice of meaning something by words, since it seems that my *utterance* of a sentence is more like performing an intentional action than pure intending. It is, however, crucial to clearly distinguish between (I) explaining an *action* by describing it as caused by a belief-desire pair, e.g. the speaker's act of uttering words with a specific intention, and (II) the issue of explaining what makes it the case that the speaker intends to do one thing rather than another. What we were after in this part was to find a potential response from Davidson to KW's sceptic's argument from queerness, along the line of Wright's judgement-dependent account of intention, which avoids the problems Davidson's actual response faced. In other words, we have been searching for Davidson's non-reductionist answer to the question what makes it the case that someone intends to mean something rather than another and how he can directly and noninferentially know what he intends. When we want to describe one's action, we appeal to belief-desire pairs, which make one's actions intentional and reasonable, though Davidson believed that there is no straightforward explanation of how a speaker ends up with a specific intention via having certain beliefs and desires, i.e. how a speaker got from his prima-facie judgments to his all-out judgments. What is important here is to note that, although intentional actions are *described* by citing belief-desire pairs, what *makes it the case* that a speaker intends something rather than another is, according to Davidson, his making a certain all-out judgement. Therefore, we can take Davidson to be providing us with a judgement-dependent account of meaning as follows:

> If ideal conditions obtain, then a speaker means *green* by "green" if and only if he judges that his utterance of "green" being interpreted in a particular way, e.g. as meaning *green*, is desirable for him.[134]

Although Davidson, via his remarks on the notion of triangulation, argued that to have any intention, the speaker must have many interrelated concepts, beliefs, and the like, and that to have any such concepts and attitudes, the speaker must be in linguistic triangulation with others, he does not think that intending to do something is to have a particular belief or pro-attitude. The beliefs condition my intention, as Davidson emphasized above, but what constitutes the fact that I intend to do something is my judging that doing that thing is desirable. Therefore, although Davidson's triangulation argument, if any, may have a role to play in explaining what makes it possible for someone to have a language and propositional attitudes at all, it does not have any direct role to play in Davidson's potential judgement-dependent account of intention and, consequently, in his judgement-dependent account of meaning.

5.10. Conclusion

In this chapter, we concluded that Davidson *could* offer a sort of judgement-dependent account of meaning and intention, which has the potential to resist KW's sceptic's sceptical problem. However, obviously, it is not the way Davidson has *actually* treated KW's sceptic's sceptical problem. Davidson, instead, tried to defuse the sceptical paradox by appealing to interpretation and induction. This approach, as we argued in the previous three chapters, cannot work as a response to KW's sceptic's sceptical argument, as it could not work as a response to Quine's arguments for the indeterminacy of translation theses.

This chapter began by considering Verheggen's reading of Davidson's remarks on triangulation, which, according to her, had the potential to offer a non-reductionist response to KW's sceptic. We saw that Davidson's use of the notion of triangulation presupposes, rather than provides a reason to accept such non-reductionism. We then investigated Davidson's account of self-knowledge as a candidate for providing a

[134] Regarding Davidson's causal explanation of meaning-determination proposed in his argument from triangulation, it is still puzzling how the sort of judgement-dependent account that we have found in Davidson's works on intention can be properly related to such a causal explanation.

response to KW's sceptic's argument from queerness. Davidson based his account on the necessary conditions on the possibility of interpretation: if interpretation is to be possible at all, speakers must have noninferential knowledge of what they mean and believe. We argued that this account cannot properly explain first-person authority, since it is compatible with a weaker claim, according to which, if interpretation is to be possible, then speakers must be taken *as if* they know what they mean. This weaker claim, however, does nothing to explain the fact that the speakers really know what they mean and believe. Also, Davidson's account could not cover the cases in which the communication between the speaker and his interpreter fails, but we can intuitively take the speaker as knowing what he means and believes, e.g. the cases in which it is the interpreter who fails to correctly interpret the speaker. We also argued that Davidson's account needs to provide us with (I) an explanation of why interpretation would not be possible if speakers did not know what they mean and believe, and (II) an answer to the question how speakers have noninferential knowledge of what they mean and believe.

We also considered Wright's objection to Davidson's account, according to which Davidson's proposed explanation of the asymmetry leads to a dilemma: it either presupposes the asymmetry or leaves the asymmetry unexplained. We argued that Wright's objection to Davidson is not plausible, since it neglects the actual explanation Davidson proposes of the asymmetry and, hence, his reading of Davidson's account is problematic. We also argued that Davidson presupposes the asymmetry in order to explain it and, thus, Wright's charge that Davidson's account is question-begging is not acceptable.

As the final part of our investigation, we claimed that Davidson's actual account of intending displayed the fundamental features of a judgement-dependent account of intention capable of blocking KW's sceptic's sceptical argument, without giving any specific role to be played by his discussion of the notion of triangulation. Davidson now seems to be capable of explaining the phenomenon of self-knowledge in a different way: speakers noninferentially know what they intend or mean because the facts about what they intend or mean are constituted by the facts about their relevant judgments.

To sum up, we can conclude that Davidson's *actual* response to KW's sceptic's sceptical argument is not successful: taking interpretation and induction to be what solves the sceptical problem displays Davidson's failure to fully grasp the metaphysical/sceptical characteristic of KW's sceptic's sceptical problem. Davidson's

triangulation argument, seen in the framework of his actual response to KW's sceptic, do not help to solve the sceptical problem (any more than his actual solution on the basis of mutual interpretation could), since Davidson's causal story about cause-determination, presented in his argument from triangulation, could be simply taken for granted by KW's sceptic. Contrary to Verheggen's reading of Davidson, we argued that Davidson could provide a non-reductionist response to KW's sceptic only by providing a plausible response to KW's sceptic's argument from queerness, and we tried to show that Davidson, on the basis of his account of intention, could opt for a judgement-dependent account of meaning, in which no clear place can be found for the notion of triangulation.

6. Chapter 6: Summary and Conclusions

In this thesis, I argued that there is a similarity between the way Davidson treats Quine's arguments for the indeterminacy of translation and the way he treats KW's sceptical argument: in both cases, Davidson fails to appreciate the sceptical outcomes of the arguments. In Chapter Three, I argued that, although Davidson accepts the main premises of Quine's arguments, i.e. Quine's physicalistic view and the indeterminacy of translation thesis, he does not admit the arguments' sceptical conclusions about fine-grained meanings. Instead, he treats the indeterminacy of translation as essentially posing an epistemological problem, according to which the alternative manuals of translation equally well capture the same fact about what the speaker means by his words. We urged, however, that Davidson cannot be a Quinean and, at the same time, take the indeterminacy of translation to be an essentially epistemological thesis. We argued that such a puzzling reading of Quine leads Davidson to a dilemma: he must either give up on Quine's physicalism, in which case he has to admit his radical divergence from Quine, or remain faithful to Quine's project, in which case he has to accept the sceptical conclusions of Quine's arguments. Moreover, it was argued that the measurement scales analogy that Davidson employs to sustain his reading of Quine's indeterminacy thesis does not work: it rather points to the fact that Davidson is making the same mistake that Chomsky allegedly made in his interpretation of Quine's indeterminacy thesis, i.e. missing the vital distinction between the epistemological problem of the underdetermination of theory and the metaphysical problem of the indeterminacy of translation. Hence, it does not seem that Davidson has actually succeeded in coming to grips with the sceptical part of Quine's arguments.

In Chapter Four, our general argument was that Davidson's reading of KW's sceptic's sceptical problem suffers from a similar problem. Davidson, first of all, rejects KW's sceptic's sceptical argument by claiming that the argument is essentially reliant on the rule-following picture of meaning and then proposes his alternative solution in terms of success in the process of mutual interpretation. We argued that Davidson's objection to KW's sceptic's sceptical argument and his alternative solution to KW's sceptical solution are both mistaken.

Davidson's objection to KW's sceptic's sceptical argument failed because of his failure to fully appreciate the strength of the argument as an essentially *sceptical* one. In Chapter Two, we considered Davidson's argument against the rule-following picture of

meaning. Based on this argument, following rules is neither necessary nor sufficient for explaining a speaker's success in the practice of meaning something by an utterance. Consequently, Davidson insisted that, since KW's sceptic's sceptical argument rests on the rule-following conception of meaning, it has to be rejected. However, we argued that KW's sceptic's sceptical argument is not essentially dependent on such a conception of meaning: it was only the *first* strand of the sceptical argument that presupposed such a view. Rather, the *second* strand of the sceptical argument can still establish the sceptic's Radical Sceptical Conclusion without bringing into play the rule-following picture of meaning. According to the view presupposed in the second strand of the argument, if a speaker means something by a word, then there is a state of affairs, or a possible meaning fact, which determines the correct application of the word for the speaker. The sceptic argued that no fact about the speaker can be found to constitute his meaning one thing rather than another. The second strand of the sceptical argument, which Davidson neglects, clearly manifests the sceptical nature of KW's sceptic's argument. Therefore, Davidson's objection to KW's sceptic failed. The sceptical argument would not be ruled out merely by rejecting the rule-following picture of meaning.

On the other hand, Davidson's alternative solution was based on the claim that the speaker means something specific by a word if his utterance of the word is interpreted by his interpreter in the way the speaker intended it to be. Davidson offered this alternative solution after rejecting KW's sceptical solution: KW's sceptical solution, for Davidson, unjustifiably requires a speaker to speak as others do in order to have successful communication with others. In Chapter Two, we explained Davidson's criticism of this requirement: according to Davidson, the requirement of speaking as others do, or even requiring the speaker to speak as she did in the past, is neither necessary nor sufficient for the speaker to have successful communication with others. Rather, communication is successful if the speaker can, intentionally or otherwise, provide enough clues and evidence for the interpreter to reach the intended interpretation of the speaker's utterance. These remarks led Davidson to the conclusion that KW's sceptical solution is wrong. We argued that Davidson's objection to KW's sceptical solution is as problematic as his objection to KW's sceptic's sceptical argument. The sceptical solution appeals to the shared linguistic practices in a speech community *because* the sceptical argument has already ruled out the success of any appeal to facts about a speaker, or a group of speakers, as what can explain their success

in the practice of meaning something by words. This again demonstrates Davidson's failure to fully grasp the *sceptical* part of KW's view, this time his "sceptical" solution.

At the same time, the fact that Davidson appealed to the process of interpretation and induction as what can supposedly deal with KW's sceptic's sceptical problem supported our claim that Davidson has taken KW's problem to be essentially epistemological, rather than sceptical. On the basis of the above discussions, we argued that Davidson's alternative solution is indeed susceptible to KW's sceptic's sceptical problem, since any alleged success in interpretation would be apparent and superficial, *unless* Davidson can properly resist the sceptical problem first, i.e. the problem of explaining what makes it the case that a speaker intends to be interpreted in one way rather than another.

We also saw that Davidson's appeal to the notion of triangulation would be of no help for him to rescue his alternative solution. We argued that the sceptic can still run his sceptical argument even by presupposing the causal facts that Davidson cites in his discussion of the notion of triangulation: the fact that a specific table in view prompted the speaker to utter "table" cannot by itself determine what the speaker means by her utterance, since this fact is compatible with the sceptic's sceptical hypotheses about what the speaker means by the word, e.g. "table" meaning *tabair*, rather than *table*.

Having argued that Davidson's actual response to KW's sceptic fails, we were left with the question as to how Davidson could have responded to KW's sceptic's sceptical problem. In Chapter Five, we searched for an answer to this question in three steps. (I) We urged that non-reductionism would provide Davidson with the best chance to resist KW's sceptic's sceptical problem. We then considered Verheggen's claim that Davidson's argument from triangulation can be read as offering a non-reductionist response to KW's sceptic's sceptical problem. We argued that Verheggen's claim fails, since Davidson's triangulation argument in fact presupposes such non-reductionism, rather than establishing it. In other words, the argument from triangulation leads to KW's sceptical problem, unless we take non-reductionism to be in play in advance. We then showed that Davidson would not really need to bring in the notion of triangulation in order to provide us with a non-reductionist response to KW's sceptic; rather, such a response requires a proper account of self-knowledge, since, as discussed in Chapter One, KW's sceptic tries to rule out non-reductionism by putting forward an argument from queerness, which, in this case, leads to the problem of accounting for self-knowledge. In other words, if it is claimed that the fact that a speaker means something

by an utterance is itself a primitive fact, then the sceptic argued that it would remain mysterious what such a fact is and how we know it. Davidson could instead claim that this primitive fact is the speaker's intending to be interpreted in a particular way. The problem for Davidson is now to explain how the speaker can have noninferential, authoritative knowledge of what he intends and means.

(II) As the next step of our inquiry, we found out that Davidson has actually attempted to provide an account of first-person authority, i.e. an explanation of the asymmetry between the way the speaker knows himself and the way others know about the speaker. Davidson's explanation of the asymmetry was generally based on the idea that the speaker noninferentially knows what he means and believes because, if he does not, interpretation would not be possible. Wright raised an objection to Davidson's suggested account, according to which this account is based on the different ways available evidence is accessible to the speaker and the interpreter in the process of interpretation and thereby faces a dilemma: either the speaker is credited with noninferential knowledge of what he means and believes *before* interpretation takes place, in which case the asymmetry is presupposed, or the speaker is credited with such knowledge *after* the interpretation takes place, in which case the asymmetry disappears. We argued that Wright's objection to Davidson is implausible, since Wright's reading of Davidson's account neglects the explanation that Davidson has actually provided of the asymmetry: the "Disquotational Argument" that Wrights attributes to Davidson is not Davidson's actual "Transcendental Argument" for the asymmetry, which is based on the necessary conditions for the possibility of interpretation.

Nevertheless, we argued that Davidson's account of self-knowledge fails for other reasons. First of all, it was argued that his account is compatible with a weaker claim, according to which, in order for interpretation to be possible, it is enough to treat the speaker *as if* he knows what he means and believes. This weaker claim, however, does not explain whether the speaker *really* knows what he means and believes. Secondly, Davidson takes self-knowledge to be fallible: if the speaker fails to provide enough evidence for the interpreter to reach the speaker's intended interpretation, then the speaker does not know what she means, since there is no meaning to be known at all. I argued that failure in interpretation can be due to the interpreter's failure to properly employ the available evidence, and, in this sense, the failure in interpretation cannot

lead to the conclusion that the speaker does not know what he means.[135] Thirdly, we argued that Davidson's account as it stands is incomplete: Davidson needs to explain *why* interpretation would not be possible if speakers did not have noninferential knowledge of what they mean and believe. A potential response from Davidson could be that the application of the principle of charity implies the speaker's having noninferential knowledge of what she means by her words since, otherwise, there would be no reason for the speaker to choose one string of words rather than another to express what she believes (given that, for Davidson, beliefs and meanings are interconnected) and there would be no reason for the interpreter to take the speaker's utterance to have any meaning at all. Fourthly, we also argued that Davidson's account only amounts to the claim that the existence of the asymmetry is necessary, while he has to provide us with an account of *how* self-knowledge works, that is, how speakers noninferentially know what they mean and believe.

(III) As the last part of our investigation, we considered a completely different way of accounting for self-knowledge, which could provide us with a non-reductionist response to KW's sceptic's sceptical problem. This response was the one offered by Wright's judgement-dependent account of meaning and intention. We suggested that Davidson's account of intending demonstrates the essential features of such a judgement-dependent account, without involving the notion of triangulation. According to Davidson's account of intending, for a speaker to intend to φ is for him to make an all-out, unconditional judgement regarding the desirability of doing φ. We suggested that this account can be linked with Davidson's remarks on meaning: if idealized conditions hold, then a speaker intends her utterance to be interpreted in a particular way if and only if she judges that being interpreted in that way is desirable for her. Therefore, we put the overall conclusion of the thesis as follows: although Davidson's actual response to KW in terms of the notions of interpretation and induction, as well as his actual account of self-knowledge as a candidate response to KW's sceptic's argument from queerness, fails, he still has the resources to develop a judgement-dependent account of meaning and intention, which has the potential to withstand the sceptical problem.

[135] It is however hard to see how Davidson can develop a criterion distinguishing between the cases in which the speaker fails to provide enough evidence and the cases in which the interpreter fails to properly utilize the available evidence. Davidson's account, hence, fails to properly explain the fallibility of the speaker's self-knowledge, i.e. to cover the cases in which, although interpretation has supposedly failed, we intuitively take the speaker to know what he means and believes.

Finally, our inquiry in this thesis can give rise to important questions for future work. One interesting question is: What would Davidson's overall philosophy look like if he could be taken to be a non-reductionist about meaning and mental content? [136] Another important inquiry can be directed to the potential form the Davidsonian judgement-dependent account of meaning and intention can take. Due to limitations of space, we could not fully investigate this issue and draw a comprehensive comparison between Davidson's account and Wright's account of intention. A further question might deal with the extent to which Davidson's later works have been inspired by Kripke's reading of Wittgenstein, as well as Wittgenstein himself. In other words, an interesting investigation can concern the incompatible readings which Davidson and Kripke offer of the later Wittgenstein.[137] Finally, a very interesting question can be asked regarding how we can place Davidson's works on meaning in the extensive contemporary debate on the normativity of meaning. Are Davidson's remarks on meaning and interpretation compatible with the idea that meaning is essentially normative?[138] Clearly, there are many other fundamental questions that arise here, the discussion of which must be left for a future investigation.

---

[136] One of the issues with Davidson's philosophy of language has been to clarify his metaphysical view about meaning, namely, whether he is a factualist or non-factualist about meaning. Neither has he been clear on this matter, nor has there been a systematic investigation of this matter. For some rare discussions see (Smart, 1999), (Stoutland, 1982a, 1982b), and (Byrne, 1998). If, as we suggested, Davidson's account of meaning and intention can be treated as judgement-dependent, then he can be taken to be a non-reductionist, and hence factualist, about meaning and intention.

[137] Interestingly, Davidson's remarks on rule-following start with his paper "Communication and Convention" (1984b), written just two years after Kripke's book on Wittgenstein (1982) was published. From 1984, many of Davidson's famous papers on philosophy of language deal with problems with the view we called the "Common View" and attempt to provide an alternative to this view. See, for instance, (Davidson, 1986, 1991a, 1991b, 1992, 1994, 1997a, 1997b, 1999b, 1999c, 2001d). It is not, I believe, an innocent coincidence that Davidson then attributes the Common View to KW and takes his alternative view, put in terms of mutual interpretation and linguistic triangulation, to be an alternative solution to KW's sceptical problem. This, I think, can signal the influence Kripke's book has had on Davidson's later views about the nature of meaning and understanding. For some discussion of the relation between Davidson's and Wittgenstein's views, see (Bridges, 2016).

[138] Davidson rejected the claim that meaning something specific by a word has anything to do with using the word in accordance with one rule rather than another. He, however, was not against the claim that, if a speaker means something by a word, then there is a fact, e.g. that the speaker intends to be interpreted in a particular way, that determines the correct application of the word for her. This may be treated as a characterization of the normativity of meaning thesis. Davidson, however, believed that meanings are not stable so that a speaker may intend to mean whatever she may, provided that her utterance can be understood in the way the speaker intended it to be. In what sense, then, can meaning be considered as normative in this Davidsonian picture? For some discussion of this matter see (Verheggen & Myers, 2016), (Turner, 2011), (Glüer, 2001), (Fennell, 2000), (Pascal, 1997, 2008), and (Pagin, 2002).

Appendix: The Early Davidson[139]

Davidson's earlier work on meaning and linguistic understanding mainly concentrates on the principal characteristics of an adequate theory of meaning for a natural language. Davidson famously uses Tarski's theory of truth as a theory, which, after being properly constrained, can be treated as a theory of meaning specifying the meaning of a speaker's utterances. The most important feature of Tarski's theory of truth is its compositionality: it is capable of specifying the truth-conditions of sentences by specifying the semantic features of their parts and the way the parts contribute to the truth-conditions of the sentences as a whole. For Davidson, any adequate theory of meaning for a natural language must be compositional, since, according to Davidson, all natural languages are essentially compositional. Davidson also treats such a theory as essentially empirical so that it should be constructed and tested by an interpreter trying to interpret the actual utterances of a speaker. In this regard, we begin with introducing Davidson's remarks on the compositionality of natural languages. We will then concentrate on the characteristics of a Davidsonian theory of meaning and the constraints Davidson imposes on such a theory.

1. Learnability Argument and Compositionality

In "Theories of Meaning and Learnable Languages" (1965), Davidson proposes an argument in favor of the idea that natural languages are compositional and that, indeed, it is this feature of these languages that makes them learnable. Davidson's target is to lead us to the conclusion that natural languages are essentially compositional, that is, that they contain a limited number of primitive terms and a finite number of rules for the way these terms can be combined to construct more complex expressions, such as sentences. The consequence of this argument is that an adequate theory of meaning for such languages must be compositional as well: it must be capable of showing how the meaning of the complex expressions of these languages is a function of the meaning, or the semantic features, of their less complex expressions. As Davidson states,

> I propose what seems to me clearly to be a necessary feature of a learnable language: it must be possible to give a constructive account of the meaning of the sentences in the language. Such an account I call a theory of meaning for the

[139] This appendix can be skipped by readers already familiar with the early Davidson's theory of meaning.

> language, and I suggest that a theory of meaning that conflicts with this condition … cannot be a theory of a natural language; and if it ignores this condition, it fails to deal with something central to the concept of a language. (1965, p. 3)

First of all, what does Davidson mean by "a constructive account of meaning"? A constructive account of meaning is an account specifying the meaning of complex expressions, such as sentences, on the basis of the meaning of simpler expressions, such as names and predicates, and the way such expressions are combined to form the complex ones. This account, hence, appreciates the compositionality of natural languages. Davidson called such a constructive account a "theory of meaning" for a language. If Davidson can show that natural languages are compositional, then he simultaneously shows that an adequate theory for such languages must be constructive, and, in this sense, compositional. Consequently, we can call such adequate theories "compositional theories of meaning" and any language, for which such an account can be offered, "a compositional language".[140] However, what is Davidson's argument for the compositionality of natural languages?

Davidson starts with drawing our attention to some important empirical facts about how language learning takes place: our language acquisition starts with learning simple names and predicates that are applicable to observable objects in our environment; we then continue learning more complex predicates and terms, which may not be directly applicable to observable objects and events in the world; for the next step, we gradually learn the terms that are more theoretical and abstract. This means that the process of learning a natural language starts with learning a small number of simple terms and predicates and extends to more complex ones. Moreover, in a limited amount of time, we become competent speakers of our languages, capable of generating and understanding a variety of sentences of these languages.

These empirical observations help Davidson to offer his more sophisticated argument in favor of the claim that natural languages are necessarily compositional. More specifically, his argument concerns the question: how can we "describe the skill or ability of a person who has learned to speak a language"? (1965, pp. 7-8). In other words, how can we explain the speaker's ability to produce and understand a potentially infinite number of new sentences and expressions, which she has never learnt or heard

[140] For more discussion of this point see (Ludwig, 2003, p. 13), (Ludwig & Lepore, 2007, pp. 17-19), and (Miller, 2007a, pp. 261-262, 274).

before? How can "an infinite aptitude [to speak and understand new sentences] … be encompassed by finite accomplishments"? (Davidson, 1965, p. 8). Davidson's argument goes as follows.

A human being, e.g. an infant, is a finite creature living a finite amount of time and is born with no language. He, however, has the ability to learn a language in a finite amount of time and becomes competent in speaking and understanding that language. Such a natural language, nonetheless, contains a potentially infinite number of sentences. Indeed, a competent speaker of such languages has the ability to produce and understand a potentially infinite number of new sentences. However, we know that "we do not at some point suddenly acquire an ability to intuit the meanings of sentences on no rule at all; that each new item of vocabulary, or new grammatical rule, takes some finite time to be learned; that man is mortal" (Davidson, 1965, p. 9). In other words, learning the meaning of any new term requires learning the rule determining the meaning of that term and learning each rule takes some time. If we accept that understanding any new sentence in such languages requires learning the rule determining its meaning and that natural languages potentially contain an infinite number of sentences, then a competent speaker of such languages would need an infinite amount of time to learn, or say, master, such languages. In this sense, such a language would not be learnable at all, since "no matter how many sentences a would-be speaker learns to produce and understand, there will remain others whose meanings are not given by the rules already mastered. It is natural to say such a language is *unlearnable*" (Davidson, 1965, p. 8). If natural languages were not compositional, that is, if they did not contain a finite number of primitive terms and a limited number of rules for their combination to construct new complex expressions, learning such languages would not be possible. Nonetheless, we learn, and become competent speakers of, such languages in a short period of time. Therefore, if natural languages are learnable, then they have a finite set of basic vocabulary and a finite set of general rules specifying how these terms can be combined to form new complex expressions so that, once the learner learns such terms and rules, she will be capable of producing and understanding a potentially infinite number of sentences and complex expressions.[141] In this sense, for Davidson, "to say that a language is learnable is just to say that speakers

[141] For more discussion, and a more detailed characterization, of this argument see (Ludwig & Lepore, 2005, pp. 28-29, 33-34)

can understand novel utterances, without explicit training in their use" (Miller, 1997, p. 148). However, what are such primitive terms and rules?

Davidson introduces a "semantical primitive" as an expression such that "the rules which give the meaning for the sentence in which it does not appear do not suffice to determine the meaning of the sentences in which it does appear" (1965, p. 9). The primitive expression, hence, is the expression, to know the meaning of which we need to learn a new rule. We cannot understand the sentences containing it on the basis of understanding the sentences which do not contain it. According to Davidson, "learnable language has a finite number of semantical primitives" (1965, p. 9), so that to learn such a language, it is enough that we learn the primitive expressions of that language and the rules specifying how these terms contribute to the meaning of the complex expressions, in which they occur. If a language lacks this feature, i.e. if it is not compositional, then there would be "no explaining the fact that, on mastering a finite vocabulary and a finitely stated set of rules, we are prepared to produce and to understand any of a potential infinitude of sentences" (Davidson, 1967, p.17).

Davidson thinks that natural languages are, hence, essentially compositional and that a compositional theory of meaning can be given for such languages. Such a compositional theory of meaning for a natural language must be capable of specifying the meaning of each sentence of this language by specifying the semantic features of the primitive parts contained in that sentence and the way those parts are combined to form that sentence.[142] However, what form should such a theory have?

2. Davidson's Theory of Meaning

Davidson's famous paper "Truth and Meaning" (1967) is dedicated to a detailed exposition of the form such an adequate theory of meaning must have. Davidson's learnability argument proposed one of the most important conditions that an adequate theory of meaning for a natural language must meet: "A satisfactory theory of meaning must give an account of how the meanings of sentences depend upon the meanings of words. Unless such an account could be supplied for a particular language, it is argued,

[142] As Ludwig and Lepore characterize this view, "a compositional meaning theory for a language L is a formal theory that enables anyone who understands the language in which the theory is stated to understand the primitive expressions of L, and the complex expressions of L on the basis of understanding the primitive ones" (2005, pp. 26-27).

there would be no explaining the fact that we can learn the language" (Davidson, 1967, p. 17). Now, Davidson's target is to clarify "what it is for a theory to give an account of the kind adumbrated" (1967, p. 17).

Davidson is against any account of meaning which takes meanings to be entities of some sort. Such a problematic proposal, for instance, begins "by assigning some entity as meaning to each word (or other significant syntactical feature) of the sentence; thus we might assign Theaetetus to 'Theaetetus' and the property of flying to 'flies' in the sentence 'Theaetetus flies'" (Davidson, 1967, p. 17). In this sort of view, the meaning of a name is the entity, to which the name refers, such as Theaetetus, and the meaning of a predicate is the property, which is instantiated by an object, such as the property of flying as instantiated by Theaetetus. According to Davidson, treating meanings in this way fails to accommodate the compositional nature of natural languages, since "the problem then arises how the meaning of the sentence is generated from these meanings" (1967, p. 17). If we take meaning to be some sort of entity and assign a meaning to each part of a sentence, then we face "an infinite regress" (Davidson, 1967, p. 17), i.e. an infinite regress of introducing new entities. The reason is that, if meanings are entities, then when we assign a meaning to each sub-sentential part of a sentence, we must assign another entity as the meaning of the sentence as a whole. For instance, if we accept that "Theaetetus flies" is true if and only if Theaetetus instantiates the property of flying, then whether or not the above sentence is true would depend on the truth-value of the sentence "Theaetetus instantiates the property of flying". However, the question is: When is the sentence "Theaetetus instantiates the property of flying" true? This sentence is true if and only if Theaetetus instantiates the property of *instantiating the property of flying* (see Miller, 2013b). This, obviously, leads us to an infinite regress of attributions of entities. We need to presume an infinite number of entities. Consequently, we will be left with no answer to the question: "How does the meaning of the whole depend on the meaning of the parts?" (Davidson, 1967, pp. 17-18). Hence, Davidson is after a different way of specifying the meaning of sentences, which does not view meanings as entities.

According to Davidson, what is fundamental to construct an adequate theory of meaning is "not to assume that parts of sentences have meanings except in the ontologically neutral sense of making a systematic contribution to the meaning of the sentences in which they occur" (1967, p. 22). In other words, a satisfactory theory of meaning should not "suggest that individual words must have meanings at all, in any

sense that transcends the fact that they have a systematic effect on the meanings of the sentences in which they occur" (1967, p. 18). Such a theory must assign a meaning to a word just as the result of its contribution to the meaning of the sentence, in which it occurs. This is directly related to the "holistic" feature of meanings, according to which if the meaning of a word depends on the role it plays in the meaning of the sentence as a whole, and if each word can occur in many different sentences, then the meanings of sentences are interdependent. Similarly, our understanding of the meaning of words will depends on our understanding of the meaning of the sentences, in which they occur. This feature of meaning, however, has an implication for an adequate theory of meaning: the meaning of a sub-sentential part can be determined only by determining the meaning of the sentence in which it appears, and the meaning of a sentence can be determined only via determining the meaning of many other sentences of that language. As Davidson emphasizes,

> if sentences depend for their meaning on their structure, and we understand the meaning of each item in the structure only as an abstraction from the totality of sentences in which it features, then we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (and word) in the language. (1967, p. 22)

This idea is Fregean, according to Davidson. Frege suggested what is now called "the context principle", according to which a word has a meaning only with respect to its contribution to the meaning of a sentence as a whole. Davidson extends this thesis to a language: "Frege said that only in the context of a sentence does a word have meaning; in the same vein he might have added that only in the context of the language does a sentence (and therefore a word) have meaning" (1967, p. 22). Hence, an adequate theory of meaning is a compositional theory for an entire language: it must be capable of systematically specifying the meaning of all the potential sentences of the language under study.

What we expect from such a theory may be initially put as the requirement that our theory must be a theory that has as consequences all sentences of the form "*s* means *m*", "where '*s*' is replaced by a structural description of a sentence and '*m*' is replaced by a singular term that refers to the meaning of that sentence; a theory, moreover, that provides an effective method for arriving at the meaning of an arbitrary sentence structurally described" (Davidson, 1967, p. 20). According to Davidson, the structural description of a sentence or an expression is "a concatenation of elements drawn from a fixed finite list (for example of words or letters)" (1967, p. 18, fn. 2). Take "$\frown$" to be the

concatenation symbol. Now, for instance, a structural description of the complex expression "the father of Michael" will be the ⌒father⌒of⌒Michael; similarly, the structural description of the word "father" will be f⌒a⌒t⌒h⌒e⌒r. It is, however, easier to take '*s*' to be the name of the sentence in question. To make a name of a sentence, we can put the sentence in quotation marks. Hence, "Snow is white" is the name of this sentence. In this situation, it is said that the sentence is mentioned, by being put in the quotation marks, or by being given a structural description, on the left -hand side of the sentence'*s* means *m*'. Davidson, however, has already criticized and rejected any view, which treats meanings as entities, and if "*m*", in the proposal discussed above, is supposed to be a singular term referring to the meaning of the sentence mentioned on the left, then meanings are treated as entities. How can this problem be solved?

To solve this problem, Davidson suggests, "nothing could be easier: just write '*s* means that *p*', and imagine '*p*' replaced by a sentence" (1967, p. 22). Instead of using a singular term to refer to the meaning of the sentence named on the left, Davidson suggests that we use a sentence instead, in which case "the theory will have done its work if it provides, for every sentence *s* in the language under study, a matching sentence (to replace '*p*') that, in some way yet to be made clear, 'gives the meaning' of *s*" (1967, p. 23). Sentences are not like singular terms, which refer to some objects or entities, and if our theory can find, for each sentence of the language under study, a matching sentence, which gives the meaning of the mentioned sentence, then we can say that our theory has been successful to specify the meaning of each sentence of that language. However, what are these matching, or meaning-specifying, sentences? Which sentence can be used to specify the meaning of the sentence "Snow is white"? Davidson thinks that the answer is relatively clear, though it depends on whether the theory is given in the "object-language" or in the "meta-language". We can take the object-language to be the language under study, that is, the language the meaning of the sentences of which is to be specified by the theory. The meta-language, hence, is the language in which the theory is specified. Thus, the object-language is the language *for* which the theory is given, and the meta-language is the language *in* which the theory is constructed. For example, we may talk about the language of propositional logic, or any formal language, by using English sentences. The formal language, about which we talk, is the object-language and the language, in which we talk about that formal language, is the meta-language. Now, according to Davidson, "one obvious candidate for matching sentence is just *s* itself, if the object language is contained in the

metalanguage; otherwise a translation of *s* in the metalanguage" (1967, p. 22). Therefore, our theory can use the same sentence named on the left to specify the meaning of that sentence, *if* the object-language and the meta-language are the same, or use the translation of the named sentence on the left to specify its meaning, *if* the object-language and the meta-language are different (and, hence, our theory is successful if it can find the right translation of the object language's sentence in the meta-language). For instance, suppose that the object language and the meta-language are the same, e.g. English. Our theory generates the following sentence, or say a "theorem", to specify the meaning of the sentence "Snow is white":

"Snow is white" means that snow is white.

If the object-language is German, and the theory is given in English, then our theory must have the following theorem specifying the meaning of the German sentence "Schnee ist weiss":

"Schnee ist weiss" means that snow is white.

Davidson, however, is not yet satisfied with the way the theory specifies the meaning of sentences, since the predicate "means that" is intensional and obscure. Why is it a problem for Davidson?

Davidson has so far introduced two conditions that an adequate theory of meaning must meet. (I) The first condition was the compositionality condition, according to which the theory of meaning must be compositional, in the sense that it must be capable of specifying the meaning of complex expressions, such as sentences, in terms of specifying the meaning of their parts and the way they contribute to the meaning of the sentences as a whole. (II) The second condition was the requirement that an adequate theory of meaning must be capable of "giving the meaning" of each sentence of the language under study; as Davidson stated, "the theory will have done its work if it provides, for every sentence *s* in the language under study, a matching sentence … that … 'gives the meaning' of *s*" (1967, p. 23). Although the suggestion that the theory must entail sentences in the form "*s* means that *p*" satisfies the above two conditions, it is not satisfactory for Davidson. The reason is that the predicate "means that" creates an intensional context, "a context in which the substitution of expressions having the same semantic value need not preserve the semantic value (truth-value) of the original sentence" (Miller, 2007a, pp. 276-277). The problem is that we have no alternative way

but to presuppose the notion of meaning to explain the difference between the following two meaning-giving theorems:

(M1) “Snow is white” means that snow is white.

(M2) “Snow is white” means that grass is green.

The sentences appearing on the right-hand side of (M1) and (M2) have the same reference, or semantic value, that is, “Snow is white” and “Grass is green” are both true. But, obviously (M2) does not give the meaning of “Snow is white”. The only difference between (M1) and (M2) is that we substituted the sentence on the right-hand side of (M1) with the sentence that has the same reference, that is, the same truth-value. However, how can we explain the difference between (M1) and (M2)? How can we say that (M2) is wrong? The best explanation we can give is to say that “Snow is white” does not mean that *grass is green*. But, once we do this, we are presupposing the notion of meaning, and we thereby fail to provide an adequate theory, which specifies meanings, rather than presupposing them. Sentences such as (M1) and (M2), and, generally, the sentences in the form “*s* means that *p*”, create intensional contexts, and the theory, which generates such sentences, cannot provide a satisfactory explanation of the notion of meaning.

Hence, Davidson’s desired theory must avoid generating such sentences. As a consequence, Davidson wants an extensional predicate, rather than the intensional predicate “means that”, to be used in the theorems of his theory of meaning. As Davidson clarifies, “as a final bold step, let us try treating the position occupied by ‘*p*’ extensionally: to implement this, sweep away the obscure ‘means that’, provide the sentence that replaces ‘*p*’ with a proper sentential connective, and supply the description that replaces ‘*s*’ with its own predicate” (1967, p. 23). The result is a schema in the following form:

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

We can call this, (T)-schema, and the instances of this schema, T-sentences. The predicate “is *T*” is not supposed to create intensional contexts and, hence, does not presuppose the notion of meaning; rather, if “*p*” is substituted by sentences, which have the same reference, the T-sentences will preserve the truth-value. However, if the theory is to meet the second condition introduced above, then all of the instances of (T), i.e. T-sentences, must somehow “give the meaning” of the sentences of the object-language.

In this regard, we can call this condition "The Extensional Adequacy Condition" (Miller, 2007a, p. 274). According to this condition, our theory must give the meaning of all sentences of the object language by producing proper T-sentences for those sentences. What is left to do, hence, is to place enough restrictions on the predicate "is *T*" so that the T-sentences entail, or give, the meaning of the object language's sentences:

> What we require of a theory of meaning for a language *L* is that without appeal to any (further) semantical notions it place enough restrictions on the predicate 'is *T*' to entail all sentences got from schema *T* when '*s*' is replaced by a structural description of a sentence of *L* and '*p*' by that sentence. (Davidson, 1967, p. 23)

We need to impose enough constraints on our theory so that, for instance, it produces the following T-sentences, in the way that we can be sure that the sentences used on the right-hand side of (T) have the same meaning as the sentences mentioned on the left-hand side have:

(T) "Snow is white" is *T* if and only if snow is white.

(T) "Schnee ist weiss" is *T* if and only if snow is white.

At this point, Davidson mentions that, obviously, the predicate "is *T*" is indeed nothing but the predicate "is true": "It is clear that the sentences to which the predicate 'is *T*' applies will be just the true sentences of *L*" (1967, p. 23). For instance, obviously "Snow is white", or "Schnee ist weiss", are true just in those situations, in which snow is white. Hence, we can simply say that "'Snow is white' is true if and only if snow is white". In this case, the second condition, i.e. the Extensional Adequacy Condition, will be the condition stating that our theory must produce all true instances of (T), which give the meaning of the object language's sentences. Obviously again, this is nothing but Tarski's condition for the adequacy of any acceptable definition of truth. According to Tarski's adequacy condition, any acceptable definition of truth should have all instances of the (T) schema as its consequence, where (T) is characterized as follows:

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

In this schema, "*s*" must be replaced by the name of the object-language's sentence, by putting it in quotation marks, and "*p*" by the sentence itself (if the object-language contains the meta-language), or by the translation of it (if the object-language does not contain the meta-language). As we will see in the next section, this condition is called

"Material Adequacy" or "Convention T". Hence, what Davidson has reached is indeed Tarski's definition of truth because "the condition we have placed on satisfactory theories of meaning is in essence Tarski's Convention *T* that tests the adequacy of a formal semantical definition of truth" (1967, p. 23). As Davidson emphasizes, the role of Convention T is "that of providing a test of the adequacy of a theory of truth: An acceptable theory must entail a true sentence with the form of [(T)] no matter what sentence of English is described by the [structural description] that replaces '*s*'" (1969, p. 46). Therefore, any adequate theory of meaning must generate all the instances of (T) for the language under study, such as:

(T1) "Snow is white" is true if and only if snow is white.

(T2) "Schnee ist weiss" is true if and only if snow is white.

Hence, it seems that what Davidson has come up with is indeed Tarski's theory of truth, which is supposed to work as a theory of meaning, that is, a theory, which specifies the meaning of the object-language's sentences. T-sentences, as we saw, can be taken to be specifying the conditions under which the sentences named on the left-hand side of the theorems, that is, the object-language's sentences, are true. In other words, they specify the truth-conditions of the object-language's sentences. As Davidson has said, the meanings of sentences can be given by specifying their truth-conditions: "The definition works by giving necessary and sufficient conditions for the truth of every sentence, and to give truth conditions is a way of giving the meaning of a sentence" (1967, p. 24). This is, indeed, an endorsement of Frege's original thought that the sense of a sentence can be given by specifying the conditions under which the sentence is true. Davidson's suggested theory, hence, if meets the conditions he has so far introduced, that is, the Extensional Adequacy and Compositionality conditions, turns out to be just Tarski's theory of truth. However, for a better understanding of Davidson's use of Tarski's theory, it is important to briefly look at Tarski's definition of truth and his Convention T.

2.1. Tarski's Theory of Truth

Tarski in his famous paper "The Semantic Conception of Truth and the Foundations of Semantics" (1944), aimed at providing a satisfactory definition of truth for a language. Tarski gives an example to clarify his idea of what such a definition must specify:

"Consider the sentence *'snow is white*.' We ask the question under what conditions this sentence is true or false" (1944, p. 343). He continues,

> it seems clear that if we base ourselves on the classical conception of truth, we shall say that the sentence is true if snow is white, and that it is false if snow is not white. Thus, if the definition of truth is to conform to our conception, it must imply the following equivalence: *The sentence "snow is white" is true if, and only if, snow is white*. (1944, p. 343)

Tarski takes the sentence that occurs on the left-hand side of the equivalence to be the name of the sentence of the language under study, and the one on the right-hand side, without the quotation mark, to be the sentence itself, if the object-language contains the meta-language, and, in the cases in which the object-language does not contain the meta-language, the sentence on the right-hand side will be the translation of the sentence named on the left. Tarski, then, considers an arbitrary sentence, "P". Suppose that the name of this sentence is replaced by "X". According to Tarski, the sentence "X is true" and "P" are equivalent in the following form:

(T) X is true if and only if P.

He calls "any such equivalence (with '*p*' replaced by any sentence of the language to which the word '*true*' refers, and '*X*' replaced by a name of this sentence) an '*equivalence of the form* (T)'" (1944, p. 344). We called this equivalence, a T-sentence. Now, Tarski's concern is to clarify the conditions under which the definition of the predicate "… is true", or the term "true", is adequate and accurate. "Material Adequacy" is introduced by Tarski as one of these conditions: "We wish to use the term "*true*" in such a way that all equivalences of the form (T) can be asserted, and *we shall call a definition of truth 'adequate' if all these equivalences follow from it*" (1944, p. 344). This is what Davidson called "Convention T". Every instance of (T) is indeed just a partial definition of truth. To have a general definition of truth for a language, we need all instances of (T): "The general definition has to be, in a certain sense, a logical conjunction of all these partial definitions" (Tarski, 1944, p. 244). If we have a procedure, through which all the true instances of (T) can be generated, then we will have a general definition of truth. These instances, as we saw, specify the conditions under which the sentence named on the left-hand side is true. Hence, we need a procedure via which we can systematically generate the truth-conditions of all sentences of the language under study. We now see how such a procedure can be achieved. In the

following part, I will follow Miller's exposition of Tarski's theory of truth (Miller, 2007a).[143]

Tarski's theory deals with formal languages. Suppose a simple formal language, which contains three simple or atomic sentences, such as P, Q, and R, and some complex sentences, which can be constructed via combining sentential connectives, such as '&', 'V' and '~', with the atomic sentences. The complex sentences will be sentences in the form P & Q, P ∨ Q, ~P, and the like. Tarski's theory is a theory showing how the truth-conditions of the complex sentences of our simple language can be systematically entailed from the truth-conditions of the simpler sentences. Using Miller's example (2007a, pp. 283-284), we can state this theory as follows:

(1a) "P" is true if and only if snow is white.

(1b) "Q" is true if and only if grass is green.

(1c) "R" is true if and only if penguins waddle.

(2a) For any sentences A, B, "A & B" is true if and only if "A" is true and "B" is true.

(2b) For any sentences A, B, "A ∨ B" is true if and only if "A" is true or "B" is true.

(2c) For any sentence A, "~A" is true if and only if "A" is not true.

The above sentences can be called the "axioms" of our theory of truth. Based on such axioms, we can generate "theorems", which specify the truth-condition of any complex sentence of this simple language. For instance, take the complex sentence "P & ~Q". By using (2a), we will reach the following theorem:

(1) "P & ~Q" is true if and only if "P" is true and "~Q" is true.

For the complex sentence "~Q", we have (2c) above, which leads us to (2):

(2) "P & ~Q" is true if and only if "P" is true and "Q" is not true.

Now, by using (1a) and (1b), we will come up with the following theorem specifying the truth-conditions for our complex sentence:

(T) "P & –Q" is true if and only if snow is white and grass is not green.

[143] For more discussion of Tarski's theory of truth, see (Kirkham, 2001) and (Lynch, 2001)

The truth-conditions of these complex sentences are, hence, a function of the semantic properties of their parts, that is, the truth-conditions of the atomic sentences, P and Q. Davidson, however, believed that such a theory can be used in the case of natural languages too. Natural languages contain sentences, which are composed of sub-sentential components, such as names and predicates. These languages also contain existential and universal quantifiers. To provide a theory of truth for even a small fragment of such natural languages, we need to bring in the language of first-order predicate logic, which helps us to formalize, and provide the logical forms of, the sentences of this fragment. The language of the predicate logic contains sentences such as:

(1) $(\forall x)(Fx \,\&\, Gx)$.

To generate the truth-conditions of the sentences like (1), we need to see how the semantic features of their components contribute to the truth-conditions of the sentences as a whole. The constituents of the above complex sentence "are the open sentences "Fx", "Gx", and because the variables in these sentences are unbound, "Fx" and "Gx" are neither true nor false" (Miller, 2007a, p. 285). "Fx" and "Gx" are called open sentences because they contain variables which are not bound by any quantifier, and hence they cannot be said to be true or false. The closed sentences, such as $(\forall x)Fx$, are those in which the variables are bound by specific quantifiers. Tarski suggests that to deal with open sentences, we can use the notion of "satisfaction": we can say that an open sentence is satisfied by certain things, and not by certain other things. For example, if we take "F" to represent the predicate ". . . is a musician", then "Fx" is satisfied by Beethoven and Bach, and not by Stalin. Hence, Tarski's general approach to deal with complex languages, such as the language of predicate logic, is to define the notion of satisfaction for the open sentences of these languages and, then, define the notion of truth in terms of this notion.

Tarski takes the satisfaction relation as a relation "between open sentences and *infinite ordered sequences of objects*" (Miller, 2007a, p. 285). Such an ordered sequence can be, for instance, <Beethoven, Bach, Stalin, John, …>. An open sentence that contains *n* variables is said to be satisfied by a sequence of objects *if* it is satisfied by the first *n* numbers of that sequence. The one variable open sentence "Fx", hence, will be satisfied by the above sequence of objects *if* it is satisfied by the first member of that sequence, that is, Beethoven. Hence, if we call the above sequence, i.e. <Beethoven, Bach, Stalin,

John, …>, the sequence X, then the open sentence "FX_1" will be satisfied by the sequence X if it is satisfied by X_1, that is, by the first member of this sequence. That is to say, "FX_1" is satisfied if X_1 is a musician. This open sentence will also be satisfied by the first member of the following sequence: <Bach, Beethoven, Stalin, John, …>, and, thereby, by any other sequences the first member of which is replaced by Beethoven. Indeed, any sequence, the first member of which is replaced with the first member of X, that is, with Beethoven, satisfies our sentence. This is the reason why "a closed sentence is satisfied by all sequences if and only if it is satisfied by at least one sequence" (Miller, 2007a, p. 287). For a better understanding of this idea, we should first consider how Tarski's definition for satisfaction can be given for our simple language. If we take X and Y to represent ordered sequences of objects, and A and B to represent open sentences of our language L*, and if we take "F" to be representing the predicate "… is a philosopher" and "G" the predicate "… is Greek", then following Miller (2007a, p. 286), we can define satisfaction relation as follows:

(1) ∀X: X satisfies "F X_1" if and only if X_1 is a philosopher;

(2) ∀X: X satisfies "G X_1" if and only if X_1 is Greek;

(3) ∀X, A: X satisfies "~A" if and only if X does not satisfy "A";

(4) ∀X, A, B: X satisfies "A & B" if and only if X satisfies "A" and X satisfies "B";

(5) ∀X, A: X satisfies "($\exists X_1$)A" if and only if there is a sequence Y, differing from X in at most the 1st place, such that Y satisfies "A";

(6) ∀X, A: X satisfies "($\forall X_1$)A" if and only if every sequence Y, differing from X in at most the 1st place, is such that Y satisfies "A";

Now, for closed sentences, such as (∀X)FX, truth can be defined as follows: "A closed sentence of L* is true if and only if it is satisfied by *all* sequences" (Miller, 2007a, p. 287). The reason for this claim is that if a closed sentence is satisfied by one sequence, then it will be satisfied by all sequences. For instance, consider the closed sentence "(∀X)FX". The definition (6) above gave the definition of satisfaction for this sentence: it will be satisfied by the sequence X if and only if every sequence Y, differing from X in its first place, satisfies the open sentence "FX". If Y is supposed to satisfy "FX", then the first member of this sequence has to be a philosopher, such as Kant, since we took "F" to be representing the predicate "… is a philosopher". But, Y

can be any sequence, the first member of which is replaced by Kant. In other words, all sequences can be seen to be satisfying this open sentence simply if we replace their first member with Kant. For instance, take the sequence Z. Does Z satisfy "FX"? For Z to satisfy "FX", there must be a sequence, differing from Z in its first place, satisfying "FX". Such a sequence can be any sequence its first place is replaced by the first member of Y, that is, Kant. In this regard, "FX" is satisfied by every sequence Y. Hence, our closed sentence "$(\forall X)FX$" is satisfied by every sequence. Now, the question is: How can this procedure, or steps, lead us to a specification of the truth-conditions of our language's sentences?

Take one of the closed sentences of our language to be "$(\forall X_1)G\ X_1$". According to our definition of the predicate "G", we know that "$(\forall X_1)G\ X_1$" is true if and only if everything is Greek. Miller (2007a, pp. 287-288) spells out the steps, via which this truth-condition can be generated:

"$(\forall X_1)(G\ X_1)$" is true if and only if it is satisfied by all sequences (from the definition of truth for closed sentences);

"$(\forall X_1)(G\ X_1)$" is true if and only if it is satisfied by at least one sequence (as discussed above);

"$(\forall X_1)(G\ X_1)$" is satisfied by a sequence X if and only if every sequence X*, differing from X in at most the first place, satisfies "$G\ X_1$" (from the step (6) above).

Now, every sequence X* satisfies "$G\ X_1$" if and only if everything is Greek. The reason is that every sequence satisfies this open sentence if its first member, X_1, is Greek and any sequence the first member of which is replaced by a Greek, say Plato, will satisfy this open sentence. Hence, we end up with the following sentence specifying the truth-condition of our closed sentence:

"$(\forall X_1)(G\ X_1)$" is true if and only if everything is Greek.[144]

This procedure is called a "recursive procedure", in which, as Tarski states, "we first describe sentential functions of the simplest structure …, and then we indicate the operations by means of which compound functions can be constructed from simpler ones" (1944, p. 352). In other words, it is the procedure, through which the truth-

[144] For more examples, see (Miller, 2007a, pp. 287-289).

conditions of complex sentences can be derived from a specification of the semantic properties of their constituents.

Therefore, we saw how Tarski's theory of truth can generate the truth-conditions of the complex sentences of a language in terms of specifying the semantic properties of their parts and the way those parts are combined to form the sentences as a whole. We can now continue investigating how Davidson uses this theory in his project.

2.2. Davidson's Use of Tarski's Theory of Truth

Davidson declared that what he reached was nothing but Tarski's T-schema, all the true instances of which, i.e. the T-sentences, must be generated by an adequate theory of meaning. What are the consequences of using this theory for Davidson?

First of all, the T-sentences, in Tarski's theory, specify the truth-conditions of the object-language's sentences. Since Davidson believed that the meanings of sentences can be specified by determining their truth-conditions, this theory can be said to be specifying the meaning of the sentences. The theory of truth in its Tarski-style can be put to use as a theory of meaning. This theory satisfies the conditions Davidson has placed on an adequate theory of meaning: his Extensional Adequacy Condition now turns out to be identical to Tarski's Convention T and, hence, Tarski's theory meets Davidson's first condition. Also, Tarski's theory of truth respects the Compositionality Condition as well, since it produces the truth-conditions of sentences by specifying the semantic properties of their parts and their mode of combination. As a result, it turns out that "a Tarski-type truth definition supplies all we have asked so far of a theory of meaning" (Davidson, 1967, p. 24).[145]

However, it is not easy for Davidson to take Tarski's truth-theory as a theory of meaning. One problem is that Tarski used the notion of translation, or sameness in meaning, in his definition of truth and he could presuppose this notion without facing any difficulty: "Tarski spelled out his condition of material adequacy in the course of attempting to give a definition of 'true': so there is no problem about his stipulating that in the instances of the (T) schema the sentence replacing 'p' should be a translation of the sentence named by 's'" (Miller, 2007a, p. 290). However, for Davidson, who wishes

[145] For more on Davidson's remarks on Tarski-style theory of truth, see (Davidson, 1973c, 1968)

to provide a theory of *meaning*, without presupposing this notion, appealing to a theory, which relies on the notion of translation, or sameness in meaning, is problematic, since the notion of meaning is already implicit in the notion of translation or sameness in meaning. Davidson is aware of this difficulty and responds to it as follows:

> 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 …. 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. (1973a, p. 134)

Davidson agrees that he cannot use the notion of translation in his theory, or, indeed, in his Extensional Adequacy Condition. Rather, all that he demands is the requirement that the theory generates all true instances of (T). However, the problem is: How can Davidson assure us that such instances give the "meaning" of the object-language's sentences?

The next problem with Davidson's account, so far sketched, is that the biconditional sentences, i.e. the T-sentences, which contain the phrase "if and only if", are true simply if the sentence appearing on the right-hand side has the same truth-value as the sentence on the left-hand side. In other words, the following two T-sentences are both true:

(T1) "Snow is white" is true if and only if snow is white.

(T2) "Snow is white" is true if and only if grass is green.

The theory, Φ1, producing (T1), and the theory, Φ2, producing (T2), both are producing true T-sentences regarding the object-language sentence named on the left-hand side, while obviously (T2) does not give the meaning of the sentence "Snow is white". In this situation, appealing to the Compositionality Condition and the Extensional Adequacy Condition is not enough to rule out one of these theories as incorrect, since both Φ1 and Φ2 are respecting the two conditions: both can derive the truth-condition of the object language's sentence from some basic axioms specifying the semantic properties of the sentence's parts, and both produce true T-sentences.

These problems show that Davidson has to introduce more constraints for his theory if the theory is supposed to work as a theory of meaning, which correctly specifies the meanings of the object-language's sentences.

3. Radical Interpretation

Davidson offers a third condition that an adequate theory of truth in its Tarski-style must meet if it is to work as a theory of meaning specifying the correct truth-conditions of the sentences of the language under study: "A theory of meaning … is an empirical theory, and its ambition is to account for the workings of a natural language. Like any theory, it may be tested by comparing some of its consequences with the facts" (Davidson, 1967, p. 24). If the sentences appearing on the right-hand side of the T-sentences are supposed to specify the correct truth-conditions, i.e. the meaning, of the object-language's sentences, and if we are not allowed to presuppose the notion of translation in advance, then we need to put a further constraint on the theory so that we can be assured that the T-sentences of this theory correctly give the meaning of the sentences. As Davidson explains, "we only need to ask, in sample cases, whether what the theory avers to be the truth conditions for a sentence really are. A typical test case might involve deciding whether the sentence 'Snow is white' *is* true if and only if snow is white" (1967, p. 25). This new constraint is introduced by Davidson by treating the theory of meaning as an empirical theory, which must be constructed, checked, and confirmed by an interpreter who tries to interpret, i.e. give the meaning of, the sentences a speaker utters. This interpreter must check whether the T-sentences the theory generates are *interpretative*, that is, whether they really produce the correct interpretation of the object language's sentences. If the theory is confirmed by the interpreter as producing acceptable theorems, then our theory is correct, though we have not presupposed the semantical notion of translation or interpretation.

If a theory of meaning is to be constructed and checked by an interpreter, then some new features must be added to its axioms and theorems, since natural languages are spoken languages and contain indexicals, such as "I", "now", and "here". In other words, if the interpreter is supposed to specify the truth-conditions of the actual uttered utterances of a speaker, then it matters when, where, and by whom, a sentence is uttered. The reason is that the sentence "It is raining here now" as uttered by Alex in Dunedin on 12 June 2016 is true, while it may be false if uttered on 15 June 2016. Davidson deals with this feature of natural languages by characterizing "truth as a

relation between a sentence, a speaker, and a time" (Miller, 2007a, p. 293).[146] Hence, the theorems of the theory of meaning will have the following form:

(T) "It is raining here now" as uttered by the speaker, S, at time *t*, is true if and only if it is raining in the vicinity of S at time *t*.

If we call the theory, which correctly generates the above T-sentence, "interpretative", then the question is: When can we say that our theory is interpretative? Davidson's answer to this question is that our theory is interpretative if it satisfies certain constraints on the procedure of interpretation. What are these constraints? What restrictions does radical interpretation impose on the Davidsonian theory of meaning?

Davidson, in his paper "Radical Interpretation" (1973a), discusses this issue by considering the problem of interpreting a speaker who speaks an unknown language. The best scenario for constructing and testing a theory of meaning in the form Davidson has in mind is the scenario of radical interpretation: an interpreter who faces a native, whose language is entirely unknown to the interpreter. In this situation, "the theory is supposed to supply an understanding of particular utterances that is not given in advance, so the ultimate evidence for the theory cannot be correct sample interpretations" (Davidson, 1973a, p. 128). In this sense, the interpretation is radical, since the theory must be constructed and checked from scratch, without presupposing any knowledge of the meaning of the utterances the native makes. Suppose that the native speaker utters the words "Es regnet". According to Davidson, if the speaker uttered the words under the right conditions and the interpreter can properly take into consideration such conditions, then she will be able to understand what the speaker means by his utterance, e.g. that the speaker has said that it is raining. By engaging in radical interpretation, hence, "we can say what his words, on that occasion, meant" (Davidson, 1973a, p. 125).

Davidson is particularly concerned with two important issues regarding the process of interpretation: (I) "What could we know that would enable us to do [interpretation]?" and (II) How could we come to know it?" (Davidson, 1973a, p. 125). Davidson deals with the first question first. What is essential to the process of interpretation is that "the

[146] As Davidson says, "truth (in a given natural language) is not a property of sentences; it is a relation between sentences, speakers (or utterers), and dates" (1969, p. 43). Theories of truth, thus, must "characterize or define a three-place predicate 'T s,u,t' " (1969, p. 44). For instance, a sentence, S, as uttered by the speaker U at time *t* is true if and only if *P*.

interpreter must be able to understand any of the infinity of sentences the speaker might utter. If we are to state explicitly what the interpreter might know that would enable him to do this, we must put it in finite form" (Davidson, 1973a, pp. 127-128). This is, as we saw, directly related to the compositionality of natural languages: if the interpreter is supposed to "understand" the native's language, then she should be able to understand all the potential sentences that can be uttered in this language. This means that the interpreter must learn the meaning of the primitive terms and predicates of this language and the way they can be combined to form more complex expressions. As Davidson emphasizes, "a satisfactory theory for interpreting the utterances of a language, our own included, will reveal significant semantic structure: the interpretation of utterances of complex sentences will systematically depend on the interpretation of utterances of simpler sentences" (1973a, p. 130). As we already know, Davidson's suggestion for a theory that can accomplish this task is a Tarski-style theory of truth: "We have such theories, I suggest, in theories of truth of the kind Tarski first showed how to give" (1973a, p. 130). Hence, what the interpreter could know, which would enable her to interpret the native's utterances, turns out to be a theory of truth in its Tarski-style: "Someone who knows the theory can interpret the utterances to which the theory applies" (Davidson, 1973a, p. 128). Hence, knowledge of a theory of truth, which meets specific conditions, can enable the interpreter to understand what a speaker utters. In this way, according to Davidson, a "theory of truth, modified to apply to a natural language, can be used as a theory of interpretation" (1973a, p. 131).

However, the second question Davidson had to find an answer to was: How could an interpreter come to know such a theory? Radical interpretation is introduced by Davidson as the answer to this question. Indeed, an interpreter who starts interpreting a speaker speaking an unknown language can be said to start constructing, checking, and confirming a theory of truth, which is supposed to specify correct interpretation of the native's utterances. If that is true, then the most important question about radical interpretation turns to a question about the evidence that the interpreter can use to start her interpretation. In other words, if the interpreter is completely ignorant about the native's language, then how can the interpreter even begin the process of interpretation? As Davidson asks, "given a theory that would make interpretation possible, what evidence plausibly available to a potential interpreter would support the theory to a reasonable degree?" (1973a, p. 125). Suppose that Kurt utters "Es regrent" and the interpreter hopes to interpret the speaker's utterance. What does the interpreter have

access to for constructing her interpretation? According to Davidson, “we have still to say what evidence is available to an interpreter – evidence, we now see, that T-sentences are true” (1973a, p. 134). The first problem is that the evidence, to which the interpreter can have access, is very limited: essentially, it is nothing but observing the speaker’s behaviour and the goings-on in the speaker’s environment, since the language is entirely unknown to the interpreter. Davidson cannot allow the interpreter to know detailed information about the speaker’s propositional attitudes, e.g. what the speaker believes, desires or intends, since this would presuppose that the interpreter already knows what the speaker means by his utterance. As Davidson states, “the evidence cannot consist in detailed descriptions of the speaker’s beliefs and intentions, since attributions of attitudes, at least where subtlety is required, demand a theory that must rest on much the same evidence as interpretation” (1973a, p. 134). This issue leads us to a very important remark from Davidson: beliefs and meanings are deeply interdependent.

3.1. The Interdependence of Belief and Meaning and Holding True Attitudes

According to Davidson, the interpreter cannot use any evidence which contains detailed information about the speaker’s beliefs because “meaning and belief are interlocked” (1974c, p. 154). Generally speaking, Davidson thinks that “the dependence of speaking on thinking is evident, for to speak is to express thoughts” (1975, p. 155). When a speaker utters a sentence, he has chosen those specific words, which mean something specific for the speaker, to express what he believes to be the case. In this sense, the speaker utters the words he does because of the meaning they have for him and the belief he has about the events and objects in his environment. As Davidson states, “the interdependence of belief and meaning is evident in this way: a speaker holds a sentence to be true because of what the sentence (in his language) means, and because of what he believes” (1973a, p. 134). For instance, when a speaker utters “Snow is white”, he utters those words, i.e. he holds the sentence to be true, on that occasion because of the belief he intends to express by his utterance, e.g. the belief that snow is white, and the meaning, which the sentence has for the speaker, e.g. that *snow is white*. As Davidson explains,

> someone who utters the sentence 'The candle is out' as a sentence of English must intend to utter words that are true if and only if an indicated candle is out at the time of utterance, and he must believe that by making the sounds he does he is uttering words that are true only under those circumstances. (1973a, p. 155)

If the meaning of a sentence can be given by specifying its truth-conditions, then a speaker holds a sentence to be true because of the truth-conditions that the sentence he utters has for him and the belief he has about the happenings in his environment. Because of that, "it is not reasonable to suppose we can interpret verbal behaviour without fine-grained information about beliefs and intentions" (Davidson, 1974c, p. 147). In the case of radical interpretation, the evidence cannot have any reference to the speaker's beliefs, intentions, and the like: "Radical interpretation should rest on evidence that does not assume knowledge of meanings or detailed knowledge of beliefs" (Davidson, 1973a, p. 135). This problem, hence, is the problem with "the nature of the evidence for the adequacy of a theory of interpretation. The evidence must be describable in non-semantic, non-linguistic terms …; it must also be evidence we can imagine the virgin investigator having without his already being in possession of the theory it is supposed to be evidence for" (Davidson, 1974c, p. 143). In other words, as Davidson explains,

> making detailed sense of a person's intentions and beliefs cannot be independent of making sense of his utterances. If this is so, then an inventory of a speaker's sophisticated beliefs and intentions cannot be the evidence for the truth of a theory for interpreting his speech behaviour. … we cannot hope to interpret linguistic activity without knowing what a speaker believes, and cannot found a theory of what he means on a prior discovery of his beliefs and intentions. (1974c, p. 144)

The consequence of this claim is that we cannot construct a single theory, which only deals with meaning, without taking into account the speaker's beliefs. If we cannot interpret the linguistic behaviour of a speaker without knowing what the speaker believes, and *vice versa*, then "in interpreting utterances from scratch – in *radical* interpretation – we must somehow deliver simultaneously a theory of belief and a theory of meaning" (Davidson, 1974c, p. 144). Hence, the interpreter must specify the meaning of the native's utterances and, at the same time, attribute certain beliefs to the speaker. She has to do both at the same time. In this sense, "we should think of meanings and beliefs as interrelated constructs of a single theory" (Davidson, 1974c, p. 146). However, the question remains: What sort of evidence is the interpreter allowed to use in order to construct such a theory and interpret the speaker's utterances and beliefs?

Davidson suggests that we can allow the interpreter to know some general, but not detailed, information about the speaker's holding true attitudes in order to help her to start her interpretation. We can assume that the interpreter knows that the speaker, when he utters a sentence, holds that sentence to be true on that specific occasion. As Davidson states, "the attitude is that of holding true, relativized to time. We may as well suppose we have available all that could be known of such attitudes, past, present, and future" (1974c, p. 144). We saw that, according to Davidson, a speaker holds a sentence to be true because of what he means by the sentence and what he believes to be the case. If the native utters a sentence on an occasion, the interpreter can now consider as evidence the fact that the speaker holds the sentence to be true on that occasion. Hence, the evidence, which is characterized without any explicit reference to the speaker's beliefs, can be put as follows:

(E) The speaker, S, holds "Es schneit" to be true at *t* if and only if it is snowing near S at *t*.[147]

The interpreter is supposed to use this evidence for constructing and checking the following T-sentence, or theorem, of her theory of truth:

(T) "Es schneit", as uttered by S at *t*, is true if and only if it is snowing near S at *t*.

This is the evidential base for the interpreter's interpretation. The evidence, however, is semantically characterized, as the interpreter presupposes the speaker's holding true *attitude*, though the evidence does not contain any detailed information about this attitude of the speaker. As Davidson explains,

> a good place to begin is with the attitude of holding a sentence true, of accepting it as true. This is, of course, a belief, but it is a single attitude applicable to all sentences, and so does not ask us to be able to make finely discriminated distinctions among beliefs. It is an attitude an interpreter may plausibly be taken to be able to identify before he can interpret, since he may know that a person intends to express a truth in uttering a sentence without having any idea *what* truth. (1973a, p. 135)

[147] Davidson gives an example to clarify this: "Typical of the sort of evidence available then would be the following: a speaker holds 'Es schneit' true when and only when it is snowing. I hope it will be granted that it is plausible to say we can tell when a speaker holds a sentence to be true without knowing what he means by the sentence, or what beliefs he holds about its unknown subject matter, or what detailed intentions do or might prompt him to utter it. It is often argued that we must assume that most of a speaker's utterances are of sentences he holds true: if this is right, the independent availability of the evidential base is assured" (1974c, p. 144).

However, how can the evidence in the form of (E) lead us to a correct interpretation of the speaker's utterance, i.e. a correct T-sentence?

There is still a problem with the process of interpretation, which has its root in the interdependence of belief and meaning. The interpreter, employing the sort of evidence mentioned above, needs to deal with two variables: meanings and beliefs. And, meanings and beliefs are claimed to be interlocked. Now, suppose that, while the speaker utters that "It is snowing", he, for some reason, does not believe that it is snowing. In this case, we cannot say that the speaker holds the sentence to be true on that occasion, despite the fact that he indeed uttered that sentence. Also, suppose that the speaker does not intend to mean that *snow is white* by his utterance of "Snow is white"; rather, to mean *grass is green*. In this case, the evidence that the speaker holds "Snow is white" to be true cannot be taken as evidence for the claim that he means that *snow is white* by "Snow is white", that is, the evidence for the theorem that "'Snow is white', as uttered by S at *t* is true if and only if snow is white near S at *t*". According to Davidson,

> the problem, then, is this: we suppose we know what sentences a speaker holds true, and when, and we want to know what he means and believes. Perhaps we could crack the case if we knew enough about his beliefs and intentions, but there is no chance of this without prior access to a theory of interpretation. Given the interpretations, we could read off beliefs from the evidential base, but this assumes what we want to know. (1974c, p. 145)

Beliefs and meanings are interdependent so that, to know what the speaker means by his utterance, we need to know what beliefs he holds when he makes that utterance. In this regard, "knowing that he holds the sentence to be true, and knowing the meaning, we can infer his belief; given enough information about his beliefs, we could perhaps infer the meaning" (Davidson, 1973a, pp. 134-135). If interpretation is to be possible at all, not only do we need to fix one of these variables, that is, either meaning or belief, but we also need to find a way, in which we can treat the speaker as behaving intelligibly, that is, as meaning something intelligible by his words and holding intelligible beliefs. Davidson offers a solution to this problem: if we can fix the speaker's beliefs in an intelligible way, we can determine what he means by his words, and *vice versa*. [148]

[148] As Davidson says, the interpreter's "knowledge of the circumstances under which someone holds sentences true is central to interpretation" since it "relates belief and interpretation in a fundamental way. We can know that a speaker holds a sentence to be true without knowing what he means by it or what belief it expresses for him" (1975, p. 162). Davidson's hope is that "if we know what belief a sentence held true expresses, we know how to interpret it" (1975, p.162).

Davidson's suggestion is that the interpreter must apply the "Principle of Charity" in order to fix the speaker's beliefs.

3.2. The Principle of Charity

If the interpreter is to be able to start her interpretation of the speaker's speech at all, she must do her best to attribute intelligible beliefs to the speaker and, thereby, treat the speaker's behaviour as intelligible. To be successful in doing so, we must "assign to [the speaker] the beliefs that *we* (the interpreters) would have in the circumstances in question, and then go on to move from these beliefs, together with the facts about what sentences [the speaker] holds true, to conclusions concerning what those sentences mean" (Miller, 2007a, p. 296). For instance, suppose that it is snowing in the speaker's vicinity. The speaker utters "Es schneit". Davidson allowed that I, as the interpreter of the speaker's utterance, know that the speaker holds this sentence to be true on this occasion. To understand what the speaker means by his utterance, however, I am required to attribute to him the most intelligible belief that I would have under the similar circumstance. Under the circumstance in question, I would have the belief that it is snowing. In this case, since the speaker holds the sentence "Es schneit" to be true because of what he means and what he believes, then by knowing what he believes, I can infer what he means by his utterance. The native speaker probably means that *it is snowing*. As Davidson says,

> this method is intended to solve the problem of the interdependence of belief and meaning by holding belief constant as far as possible while solving for meaning. This is accomplished by assigning truth conditions to [the native speaker's] sentences that make native speakers right when plausibly possible, according, of course, to our own view of what is right. (1973a, p. 137)

This requirement, or principle, hence, puts another empirical constraint on the procedure of radical interpretation, through which the theory of meaning is constructed, tested, and confirmed: the native's utterances must be interpreted in the light of the intelligible beliefs which the interpreter attributes to the native. The principle of charity has been characterized in different ways by Davidson.[149] The point of this principle is that the

[149] Davidson sometimes prefers to talk about the "principles of charity", rather than a single principle. For instance, he states: "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

interpreter must maximize the intelligibility, or minimize the unintelligibility, of the speaker's behaviour: "Charity prompts the interpreter to maximize the intelligibility of the speaker, not sameness of belief" (Davidson, 2001b, p. xix). To interpret the speaker, the interpreter does not have to always attribute to the speaker the same beliefs that she would have on that occasion; rather, she is only required to attribute *intelligible* beliefs to the speaker. For instance, suppose that the speaker, who uttered "Es schneit" when it was snowing, now utters the same sentence when he and the interpreter are at home and the weather seems nice and clear. Perhaps, the speaker utters such a sentence because, for instance, he is hallucinating, drunk, or too much reliant on the on-line weather forecast. In this case, the interpreter has two options. On the one hand, the interpreter could attribute to the speaker the belief that she thinks is true, that is, the belief that she would have on that occasion: it is not snowing outside. If she does this, then she has to attribute a different meaning to the speaker's utterance, e.g. that the speaker means that *it is not snowing* by uttering "Es schneit", since the speaker is supposed to hold the sentence to be true because of what he means by that sentence and what he believes to be the case. However, on the other hand, if the interpreter has checked the speaker in enough cases and concluded that the speaker has used this utterance only in the situations, in which it is snowing, then it does not seem to be the best option to attribute the true belief to the speaker and assume that the speaker has changed the meaning of his utterance. Rather, the more intelligible way to treat the speaker is to attribute to him a false belief about what is going on outside: the speaker still means that *it is snowing* by his utterance, while he is wrong in his belief that it is snowing outside. As Davidson emphasizes, "the aim of interpretation is not agreement but understanding" (2001b, p. xix).

Hence, the application of the principle of charity is necessary if the interpretation is to be possible: "Charity is not an option, but a condition of having a workable theory" (Davidson, 1974a, p. 197). To have a theory, which produces the correct interpretation of the native's utterances, we have no choice but to fix the beliefs of the speaker and interpret his speech, and doing so essentially relies on the application of the principle of charity. In this sense, "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" (Davidson, 1974a, p.

---

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" (1991a, p. 211).

197). If we fail to make sense of the speaker's behaviour as intelligible, we would not have any good reason to count the utterances he makes as having any meaning at all. Hence, "charity in interpreting the words and thoughts of others is unavoidable" (Davidson, 1967, p. 27).[150]

Nonetheless, there is still a problem to be dealt with. Suppose that it is for the first time our radical interpreter confronts the native. The native utters "Schnee ist weiss" on the occasion, in which there is white snow in view. The interpreter starts relating the goings-on in the speaker's environment to the speaker's utterance. She uses the following evidence:

(E1) The native holds "Schnee ist weiss" to be true at *t* if and only if snow is white in S's vicinity at *t*.

She, then, takes the above evidence to be supporting the following theorem:

(T1) "Schnee ist weiss", as uttered by S at *t*, is true if and only if snow is white in S's vicinity at *t*.

But, suppose that, at the time the speaker utters that sentence, there is a dog barking in the neighborhood. Our interpreter, however, has the following evidence available too:

(E2) The native holds "Schnee ist weiss" to be true at *t* if and only if a dog barks in S's vicinity at *t*,

This evidence supports a different theorem regarding the utterance of the speaker, that is,

(T2) "Schnee ist weiss", as uttered by S at *t*, is true if and only if a dog barks in S's vicinity at *t*.

Obviously, on the basis of that situation alone, there is no clue as to whether (T1) is correct or (T2). In other words, both the T-sentences are true, while they deliver different truth-conditions, i.e. meanings, for the native's utterance. How can this problem be solved?

[150] As Davidson emphasizes, "the point of the principle is to make the speaker intelligible, since too great deviations from consistency and correctness leave no common ground on which to judge either conformity or difference. From a formal point of view, the principle of charity helps solve the problem of the interaction of meaning and belief by restraining the degrees of freedom allowed belief while determining how to interpret words" (1983, pp. 148-149).

3.3. Holism and T-Sentences

Davidson, to solve this problem, draws our attention again to the holistic feature of meaning:

> If sentences depend for their meaning on their structure, and we understand the meaning of each item in the structure only as an abstraction from the totality of sentences in which it features, then we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (and word) in the language. (1967, p. 22)

According to Davidson, it is not possible to interpret a single utterance of the native, without interpreting a large number of other sentences that the native may utter on different occasions. Emphasizing on the holistic feature of the procedure of interpretation is, indeed, another constraint that Davidson imposes on such a process:

> [T]his criterion is that the totality of T-sentences should … optimally fit 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. (1973a, p. 139)

To determine which T-sentence is interpretative, that is, correctly gives the meaning of the native's utterance, the interpreter needs to consider different situations, in which the speaker utters that sentence. Moreover, as the holistic feature of meaning and the compositionality of natural languages imply, the interpreter needs to break the native's sentences down to their parts and investigate how the meaning of the sentences as a whole depends on the meaning of their parts and the way they are combined to form those sentences. For instance, concerning the example mentioned above, suppose that our interpreter continues her interpretation. Next time, when it is snowing and everything is covered by snow, and there is no dog barking in the vicinity, the interpreter checks whether or not the native utters the same sentence. Suppose that this time again the native utters "Schnee ist weiss". The interpreter now has the following evidence:

(E) The native holds "Schnee ist weiss" to be true at t_2 if and only if snow is white in S's vicinity at t_2.

The interpreter, then, has stronger evidence to confirm the following T-sentence as what specifies the meaning of the speaker's utterance:

(T) “Schnee ist weiss”, as uttered by S at t_2, is true if and only if snow is white in S’s vicinity at t_2.

The interpreter continues checking other situations, in which the native uses some parts of this sentence in his utterance of different sentences. Suppose that the interpreter has broken the above native sentence down to its parts and come up with the following axioms:

(A1) “Schnee”, for the speaker S at *t*, refers to snow.

(A2) “*x* ist weiss”, for the speaker S at *t*, is satisfied by *x* if and only if *x* is white.

Next time, the native utters “Der Schwan ist weiss”, while he is pointing at a white swan in a lake. The interpreter, hence, has strong evidence to conclude that:

(T3) “Der Schwan ist weiss”, as uttered by S at *t*, is true if and only if there is a white swan in S’s vicinity at *t*.

She can use (A2) above, together with the following (A3) to infer (T3):

(A3) “Schwan“, for the speaker S at *t*, refers to swan.

As the interpreter continues her interpretation, she gradually constructs her theory of truth, and checks and confirms the theory’s theorems, specifying the interpretation of the native’s utterances via specifying the semantic properties of the sub-sentential parts and the way they contribute to the meaning of the sentences as a whole. If the theory is constructed in a way that meets all the constraints Davidson has introduced, and if the interpreter applies the principle of charity, then she will come up with a correct theory of truth, which systematically produces the correct interpretation of the native’s language’s sentences. The interpreter who knows such a theory and knows that the theory meets the formal and empirical constraints can be said to be capable of understanding the native’s language. In this way, a correct theory of truth in its Tarski-style can be said to be modelling the speaker’s and the interpreter’s linguistic abilities to speak and understand, that is, to be describing what they know, which enables them to speak to each other and understand each other. As Davidson says,

> the theory may be used to describe an aspect of the interpreter’s competence at understanding what is said. We may, if we please, also maintain that there is a mechanism in the interpreter that corresponds to the theory. If this means only that there is some mechanism or other that performs that task, it is hard to see how the claim can fail to be true. (1974c, p. 141)

This claim, however, does not imply that the interpreter explicitly knows such a theory; rather, the theory can be taken to be providing us with a way of modeling or describing the interpreter's and the speaker's linguistic abilities. In this sense, "the theory can be used to describe what every interpreter knows, namely a specifiable infinite subset of the truths of the theory" (Davidson, 1974c, p. 142). As a conclusion, hence, "a T-sentence of an empirical theory of truth can be used to interpret a sentence, then, provided we also know the theory that entails it, and know that it is a theory that meets the formal and empirical criteria" (Davidson, 1973a, p. 139). In this way, a Tarski-style theory of truth can work as a theory of interpretation, systematically specifying the meaning of a potentially infinite number of the sentences of the native's language.[151]

4. The Indeterminacy of Interpretation

The indeterminacy of interpretation thesis is a very important theme in Davidson's earlier works on meaning and linguistic understanding, and, hence, it is important to end our exposition of Davidson's earlier view by briefly looking at his remarks on the indeterminacy of interpretation, that what the indeterminacy of interpretation is and how it appears as a consequence of the procedure of radical interpretation. Doing so facilitates our discussion of Davidson's grasp of Quine's indeterminacy of translation thesis in Chapter Three.

There are two kinds of indeterminacy that Davidson considers. The first sort of indeterminacy emerges with regard to the fact that the available evidence, to which the interpreter has access, might not be sufficient to fix one correct interpretation of the native's utterances. Generally speaking, it is possible that the available evidence for interpreting the native's utterance fails to determine to which object the native's word, in the native's language, refers. Suppose that a rabbit scurries and the native utters "Lo, gavagai". Based on the available evidence, i.e. that the native holds "Lo, gavagai" to be true if and only if there is a rabbit in view, it is reasonable to interpret the native's utterance as follows:

(T1) "Lo, gavagai" is true if and only if there is a rabbit in view.

[151] For more on Davidson's theory of meaning and radical interpretation, see (Kemp, 2012), (Ludwig, 2003), (Joseph, 2004), (Ludwig & Lepore, 2005, 2007), (Miller, 2007a, 2013b), (Glock, 2003), and (Ramber, 1989)

However, what is the referent of the word “gavagai”, which occurs in the native’s uttered sentence? It is possible that this word, in the native’s language, refers to something different from what “rabbit” refers to, in the interpreter’s language. For instance, “gavagai” may be taken by the native to be referring to an undetached rabbit part, to a stage of a rabbit, to an instantiation of the concept of rabbithood, and the like. The available evidence, however, is compatible with all such alternative interpretations of the native’s word. These alternative interpretations will give rise to different T-theorems for the native’s utterance, for instance:

(T2) “Lo, gavagai” is true if and only if there is an undetached rabbit part in view.

(T3) “Lo, gavagai” is true if and only if there is a stage of a rabbit in view.

And, so on.

In this way, the reference of the native’s word is said to be inscrutable: the available evidence is not sufficient to help the interpreter to fix one of these references as the correct reference of the native’s words. As a result, some degree of indeterminacy happens in the interpretation of the native’s sentence: “The inscrutability of reference says there is no way to tell which way of connecting words with things is the right way; if one way works, there will be countless others that do as well” (Davidson, 1997a, p. 78). There are always alternative interpretations of the native’s utterances because of the inscrutability of the reference of the words occurring in them. This is a consequence of the fact that we can always substitute a word (or a satisfaction relation) with another word (or another satisfaction relation), which has the same reference, without making any change in the truth-conditions of the sentence in which they occur. Davidson gives an example to clarify such indeterminacy. Suppose a satisfaction relation *s* represents, or maps, “Rome” onto Rome and “is a city in Italy” onto cities in Italy. Another satisfaction relation *s′*, however, maps “Rome” onto an area 100 miles to the south of Rome and “is a city in Italy” onto areas 100 miles south of cities in Italy. Having taken for granted the relation *s*, the T-sentence for the sentence “Rome is a city in Italy” will be:

(T1) ‘Rome is a city in Italy’ is true if and only if Rome is a city in Italy,

while, if we choose *s′*, we will have:

(T2) ‘Rome is a city in Italy’ is true if and only if the area 100 miles south of Rome is an area 100 miles south of a city in Italy.

The problem is that the truth-conditions of the above sentences are the same, that is, if Rome is a city in Italy, then an area 100 miles south of Rome is an area 100 miles south of a city in Italy. Both (T1) and (T2) are true as well. According to the inscrutability of reference thesis, “there can be no evidence that *s* is any better than *s′* for interpreting the sentence” (Davidson, 1997a, p. 78). The reason is that there is no clue to determine which reference of the word “Rome” is the correct reference of this word. All the available evidence fails to pick up one of these two options as correct. Hence, “there is no saying which one his word “Rome” refers to” (Davidson, 1997a, p. 79). However, one might object that it is obvious that the reference of “Rome” is just Rome, not anything else; “Rome” does not *mean* an area 100 miles south of Rome. But, according to Davidson, this claim neglects the compositionality feature of natural languages and the holistic features of meaning: “Individual words don’t have meanings. They have a *role* in determining the truth conditions of sentences, and this role is captured by *s′* as surely it is by *s*” (1997a, p. 79). The fact that, in our speech community, there are some conventions regarding the use of words which we usually follow, or the fact that there are some standard methods of word-for-word translation “does not mean that other interpretations are wrong: *s′* provides as correct an interpretation of ‘Rome is a city in Italy’ as *s*” (Davidson, 1997a, p. 79). The inscrutability of reference and the resulted indeterminacy of interpretation, together with the holistic feature of the process of interpretation, lead us to the important conclusion that there can always be different theories of meaning, which are all compatible with all the available evidence the interpreter can have access to. For instance, it is possible that one interpreter comes up with a theory of meaning containing (T1) and another interpreter with a different theory containing (T2). According to Davidson, “if there is indeterminacy, it is because when all the evidence is in, alternative ways of stating the facts [regarding what the native means by his words] remain open” (1974c, p. 154).

The second sort of indeterminacy appears as a result of the interdependence of belief and meaning. As Davidson states, “a theory of interpreting the utterances of a single speaker, based on nothing but his attitudes towards sentences, would, we may be sure, have many equally eligible rivals, for differences in interpretation could be offset by appropriate differences in the beliefs attributed” (1974c, p. 153). We saw that the available evidence for the interpreter is the fact that the speaker holds a sentence to be

true on a specific occasion. A speaker holds a sentence to be true because of what he means by the sentence and what he believes. This interdependence of belief and meaning, however, brings about another sort of indeterminacy, for which Davidson gives an example:

> I find that I very often disagree with other people over whether to call the color of some object green or blue. The disagreement is consistent: there is a fairly definite range of cases where I say green and they say blue. We can account for this difference in two ways: it may be that I (or most other people) are wrong about the color of certain objects, or it may be that I don't use the words "blue" and "green" in quite the way others do. There may be no way to decide between these two accounts; by making compensatory adjustments elsewhere in one's interpretation of my sentences and beliefs one can accommodate either story. But on one account certain of my pronouncements about colors are false, while on the other they are true. (1997a, pp. 80-81)

Contrary to the first sort of indeterminacy, this second type of indeterminacy affects the truth-conditions and truth-values of T-sentences. According to this indeterminacy, the interpreter always has a choice in her attribution of beliefs to the speaker and interpreting the speaker's utterances. For instance, when Davidson uses "green" to talk about the things, which other people take to be blue, his interpreters have a choice as follows: (I) they can take Davidson to have the similar beliefs that they have about the colour of those things, but to mean something different from what they mean by "green", for instance, *green or blue*. (II) They can take Davidson to mean the same thing they do by "green", that is, *green*, but to have a different, wrong belief about the colour of the things to which he applies "green". When a disagreement in interpretation occurs, the interpreters always have a choice: either fix the meaning and attribute a different belief to the speaker, or fix the belief and attribute a different meaning to his words. In this case of indeterminacy, however, truth-values change: according to one interpretation, what the speaker utters is true, while, according to the other, it is false. Again, by taking into consideration the holistic feature of the procedure of interpretation, the consequence of this indeterminacy is that there will always be alternative theories of interpretation for interpreting the speaker's utterances.[152]

However, Davidson thinks that the formal and empirical constraints he has placed on theories of meaning and the process of interpretation will reduce the degree of such indeterminacies to an acceptable level: "I believe the range of acceptable theories of

[152] As we discuss this in detail in Chapter Three, the first sort of indeterminacy is called by Quine "the Inscrutability of Reference", while the second is properly called "the Indeterminacy of Translation".

truth can be reduced to the point where all acceptable theories will yield T-sentences that we can treat as giving correct interpretations, by application of further reasonable and non-question-begging constraints" (1974c, p. 152). The point is that the interpreter must choose and stick "to some one method of interpretation" (Davidson, 1997a, p. 79). Obviously, if she shifts between alternative interpretations, she would never come up with a coherent theory of interpretation. Also, the interpreter must interpret a large number of the speaker's utterances: "Each interpretation and attribution of attitude is a move within a holistic theory" (Davidson, 1974c, p. 154). Under these circumstances, if the interpreter applies the principle of charity, then the theory she will come up with will be an acceptable one, since it will generate a correct interpretation of the speaker's utterances by representing the behaviour of the speaker as intelligible via attributing to him intelligible beliefs. The degree of the indeterminacy which remains would not be considerable and, hence, would not supposedly bring about any serious problem for Davidson's account. In this way, Davidson thinks that "indeterminacy is important only for calling our attention to how the interpretation of speech must go hand in hand with the interpretation of action generally, and so with the attribution of desires and beliefs" (Davidson, 1974c, p. 154).[153] It is, however, the purpose of Chapter Three to discuss and evaluate Davidson's reasons for this claim.

[153] For Davidson's discussion of the indeterminacy see e.g. (Davidson, 1979, 1974c, 1997b).

References

Baker, G. P., & Hacker, P. M. S. (1984). *Scepticism, Rules and Language*. Oxford: Basil Blackwell.

Boghossian, P. (1989). The Rule-Following Considerations. *Mind, 98(392)*, 507–549. Reprinted in Miller, A. and Wright, C. (2002), pp. 141-187, to which page references apply.

Bridges, J. (2016). The Search for "the Essence of Human Language" in Wittgenstein and Davidson. Forthcoming in Claudine Verheggen (Ed.), *Wittgenstein and Davidson on Thought and Action*. Cambridge University Press.

Byrne, A. (1996). On Misinterpreting Kripke's Wittgenstein. *Philosophy and Phenomenological Research*, 56 *(2)*, 339 – 343.

Byrne, A. (1998). Interpretivism. *European Review of Philosophy*, *3*, 199 – 223.

Chomsky, N. (1968). Quine's Empirical Assumptions. *Synthese, 19 (1/2)*, 53 – 68.

Davidson, D. (1963). Actions, Reasons, and Causes. *The Journal of Philosophy, 60*, 685–99. Reprinted in Davidson, D. (2001c), pp. 3-20, to which page references apply.

Davidson, D. (1965). Theories of Meaning and Learnable Languages. In Y. Bar-Hillel (Ed.), *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, (pp. 383–94). Amsterdam: North Holland. Reprinted in Davidson, D. (2001a), pp. 3-16, to which page references apply.

Davidson, D. (1969). True to the Facts. *The Journal of Philosophy, 66*, 748–64. Reprinted in Davidson, D. (2001a), pp. 43-54, to which page references apply.

Davidson, D. (1968). On Saying That. *Synthese, 19*, 130–46.

Davidson, D. (1967). Truth and Meaning. *Synthese, 17*, 304–23. Reprinted in Davidson, D. (2001a), pp. 17-42, to which page references apply.

Davidson, D. (1970). Mental Events. In L. Foster and J. Swanson (Eds.), *Experience and Theory* (pp. 79–102). Amherst: University of Massachusetts Press. Reprinted in Davidson, D. (2001c), pp. 207-224, to which page references apply.

Davidson, D. (1973a). Radical Interpretation. *Dialectica, 27*, 314–28. Reprinted in Davidson, D. (2001a), pp. 125-140, to which page references apply.

Davidson, D. (1973b). The Material Mind. It was originally presented at the Fourth International Congress for Logic, Methodology, and Philosophy of Science, Dordrecht, Holland. Reprinted in Davidson, D. (2001c), pp. 245-260, to which page references apply.

Davidson, D. (1973c). In Defence of Convention T. In H. Leblanc (Ed.), *Truth, Syntax and Modality*, (pp. 76–86). Dordretch: North-Holland.

Davidson, D. (1974a). On the very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association, 47*, 5–20. Reprinted in Davidson, D. (2001a), pp. 183-198, to which page references apply.

Davidson, D. (1974b). Comments and Replies. In S. C. Brown (Ed.), *Philosophy of Psychology*, The Macmillan Press and Barnes, Noble, Inc. Reprinted in Davidson, D. (2001c), pp. 239-244, to which page references apply.

Davidson, D. (1974c). Belief and the Basis of Meaning. *Synthese, 27*, 309–23. Reprinted in Davidson, D. (2001a), pp. 141-154, to which page references apply.

Davidson, D. (1975). Thought and Talk. In S. Guttenplan (Ed.), *Mind and Language* (pp. 7–23). Oxford: Oxford University Press. Reprinted in Davidson, D. (2001a), pp. 155-170, to which page references apply.

Davidson, D. (1977). Reality without Reference. *Dialectica, 31*, 247–53. Reprinted in Davidson, D. (2001a), pp. 215-226, to which page references apply.

Davidson, D. (1978). Intending. In Y. Yovel (Ed.), *Philosophy of History and Action*, (pp. 41–60). Dordretch: D. Reidel and The Magnes Press. Reprinted in Davidson, D. (2001c), pp. 83-102, to which page references apply.

Davidson, D. (1979). The Inscrutability of Reference. *The Southwest Journal of Philosophy, 10*, 7–20. Reprinted in Davidson, D. (2001a), pp. 227-242, to which page references apply.

Davidson, D. (1982). Rational Animals. *Dialectica, 36*, 317–28. Reprinted in Davidson, D. (2001b), pp. 95-106, to which page references apply.

Davidson, D. (1983). A Coherence Theory of Truth and Knowledge. In D. Henrich (Ed.), *Kant oder Hegel?* (pp. 423–38). Stuttgart: Klett-Cotta. Reprinted in Davidson, D. (2001b), pp. 137-153, to which page references apply.

Davidson, D. (1984a). First Person Authority. *Dialectica, 38,* 101–12. Reprinted in Davidson, D. (2001b), pp. 3-14, to which page references apply.

Davidson, D. (1984b). Communication and Convention. *Synthese, 59,* 3-17. Reprinted in Davidson, D. (2001a), pp. 265-280, to which page references apply.

Davidson, D. (1984c). Expressing Evaluations. Lawrence, Kan.: Lindley. The paper was delivered as the Lindley Lecture and published as a Lindley Lecture monograph at the University of Kansas. Reprinted in Davidson, D. (2004), pp. 19-38, to which page references apply.

Davidson, D. (1986). A Nice Derangement of Epitaphs. In E. Lepore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (pp. 433–46). Cambridge: Blackwell. Reprinted in Davidson, D. (2005), pp. 89-108, to which page references apply.

Davidson, D. (1987). Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association, 60*(3), 441–58. Reprinted in Davidson, D. (2001b), pp. 15-38, to which page references apply.

Davidson, D. (1988). The Myth of the Subjective. In M. Benedikt and R. Berger (Eds.), *Bewusstsein, Sprache und die Kunst.*Vienna: Edition S. Verlag der Österreichischen Staatsdruckerei. Reprinted in Davidson, D. (2001b), pp. 39-52, to which page references apply.

Davidson, D. (1991a). Three Varieties of Knowledge. In A. P. Griffiths (Ed.), *A. J. Ayer: Memorial Essays* (pp. 153–66). New York: Cambridge. Reprinted in Davidson, D. (2001b), pp. 205-220, to which page references apply.

Davidson, D. (1991b). James Joyce and Humpty Dumpty. *Midwest Studies in Philosophy, 16,* 1–12. Reprinted in Davidson, D. (2005), pp. 143-158, to which page references apply.

Davidson, D. (1991c). Epistemology Externalized. *Dialectica, 45* (*2-3*), 191-202.

Davidson, D. (1992). The Second Person. *Midwest Studies in Philosophy, 17*, 255–67. Reprinted in Davidson, D. (2001b), pp. 107-122, to which page references apply.

Davidson, D. (1994). The Social Aspect of Language. In B. McGuinness (Ed.), *The Philosophy of Michael Dummett,* (pp. 1–16). Dordrecht: Kluwer. Reprinted in Davidson, D. (2005), pp. 109-126, to which page references apply.

Davidson, D. (1995a). Could There Be a Science of Rationality?. *International Journal of Philosophical Studies, 3*, 1-16. Reprinted in Davidson, D. (2004), pp. 117-134, to which page references apply.

Davidson, D. (1995b). The Problem of Objectivity. *Tijdschrift voor Filosofie, 57* (2), 203–20. Reprinted in Davidson, D. (2004), pp. 3-18, to which page references apply.

Davidson, D. (1997a). Indeterminism and Antirealism. In C. B. Kulp (Ed.), *Realism/Antirealism and Epistemology* (pp. 109–22). Lanham, Md.: Rowman and Littlefield. Reprinted in Davidson, D. (2001b), pp. 69-84, to which page references apply.

Davidson, D. (1997b). Seeing through Language. In J. Preston (Ed.), *Thought and Language,* (pp. 15–28). Cambridge: Cambridge University Press. Reprinted in Davidson, D. (2005), pp. 127-142, to which page references apply.

Davidson, D. (1998a). The Irreducibility of the Concept of the Self. . In M. Stamm (Ed.), *Philosophie in Synthetisher Absicht* (pp. 123–30). Stuttgart: Klett-Cotta. Reprinted in Davidson, D. (2001b), pp. 85-92, to which page references apply.

Davidson, D. (1998b). Replies. *Cr'ıtica, 30*, 97–112. Replies to articles by Barry Stroud, John McDowell, Richard Rorty, and Carlos Pereda in *Cr'ıtica, 28* (1998). Reprinted in Davidson, D. (2005), pp. 315-328, to which page references apply.

Davidson, D. (1999a). Reply to W.V. Quine. In L. E. Hahn (Ed.), *The Philosophy of Donald Davidson* (pp. 80-86). US, Illinois: Open Court.

Davidson, D. (1999b). Reply to Richard Rorty. In L. E. Hahn (Ed.), *The Philosophy of Donald Davidson* (pp. 595-600). US, Illinois: Open Court, 1999.

Davidson, D. (1999c). The Emergence of Thought. *Erkenntnis, 51*(1), 7–17. Reprinted in Davidson, D. (2001b), pp.123-134, to which page references apply.

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

Davidson, D. (2001b). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.

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

Davidson, D. (2001d). What Thought Requires. In J. Branquinho (Ed.), *The Foundations of Cognitive Science*, (pp. 121-132). Oxford: Oxford University Press. Reprinted in Davidson, D. (2004), pp. 135-150, to which page references apply.

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

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

Davies, D. (1998). How Sceptical is Kripke's 'Sceptical Solution'? *Philosophia, 26 (1)*, 119 – 140.

Dummett, M. (1959/1978). Truth. *Proceedings of the Aristotelian Society*, *59*: 141–162. Reprinted in Dummett (1978), *Truth and Other Enigmas*. Duckworth, London.

Dummett, M. (1973). *Frege: Philosophy of Language*. London: Duckworth.

Dummett, M. (1978). The Social Character of Meaning. In Dummett M., *Truth and Other Enigmas*, (pp. 420-430). Duckworth, London

Dummett, M. (1986). 'A Nice Derangement of Epitaphs': Some Comments on Davidson and Hacking. In Ernest Lepore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, (pp. 459-476). Cambridge: Blackwell.

Ebbs, G. (1997). *Rule-Following and Realism*. Cambridge: Harvard University Press.

Edwards, J. (1992). Best Opinion and Intentional States. *Philosophical Quarterly, 42 (166)*, 21 – 33.

Fennell, J. (2000). Davidson on Meaning Normativity: Public or Social. *European Journal of Philosophy, 8 (2)*, 139-154.

Gaudet, E. (2006). *Quine on Meaning*. London: Continuum International Publishing Group.

Gibson, R. (1994). Quine and Davidson: Two Naturalized Epistemologists. *Inquiry, 37 (4)*, 449-463.

Gibson, R. (2004). *The Cambridge Companion to Quine*. Cambridge: Cambridge University Press.

Glock. H. J. (2003). *Quine and Davidson on Language, Thought, and Reality*. Cambridge: Cambridge University Press.

Glüer, K. (2001). Dreams and Nightmares: Conventions, Norms, and Meaning in Davidson's Philosophy of Language. In P. Kotatko, P. Pagin, and G. Segal (Eds.), *Interpreting Davidson*, (pp. 53-74). Stanford CA: CSLI Publications.

Glüer, K. (2006). Triangulation. In E. Lepore and B. Smith (Eds.), *The Oxford Handbook of Philosophy of Language*, (pp. 1006-1019). Oxford: Oxford University Press.

Glüer, K. (2011). *Donald Davidson. A Short Introduction*. New York: Oxford University Press.

Glüer, K. (2013). Convention and Meaning. In E. Lepore and K. Ludwig (Eds.), *A Companion to Donald Davidson*, (pp. 339-360). Oxford: Wiley Blackwell.

Goodman, N. (1983). *Fact, Fiction, and Forecast*. Cambridge, US: Harvard University Press.

Hacker, P. M. S. (1997). Davidson on First-Person Authority. *Philosophical Quarterly*, 47, 285–304.

Hacking, I. (1986). The Parody of Conversation. In E. Lepore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, (pp. 447-458). Cambridge: Blackwell.

Hattiangadi, A. (2007). *Oughts and Thoughts: Rule-Following and the Normativity of Content*. Oxford: Oxford University Press.

Heal, J. (1997). Radical Interpretation. In C. Wright and B. Hale (Eds.), *Companion to Philosophy of Language*, (pp. 175-196). US, Mass.: Blackwell Publishers Inc.

Holly, W. J. (1986). On Donald Davidson's 'First Person Authority'. *Dialectica*, *40*, 153–6.

Holton, R. (1993). Intention Detecting. *Philosophical Quarterly*, *43*, 298 – 318.

Hookway, C. (1988). *Quine*. Cambridge: Polity Press.

Hylton, P. (2007). *Quine*. New York: Routledge.

Joseph, M. (2004). *Donald Davidson*. Montreal: McGill-Queen's University Press.

Kemp. G. (2006). *Quine: A Guide for the Perplexed*. London: Continuum.

Kemp. G. (2012). *Quine versus Davidson: Truth, Reference, and Meaning*. Oxford: Oxford University Press.

Kirk, R. (1986). *Translation Determined*. Oxford: Oxford University Press.

Kirkham, R. L. (2001). *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press, fifth addition.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Cambridge, US: Harvard University Press.

Kusch, M. (2006). *A Sceptical Guide to Meaning and Rules: A Defence of Kripke's Wittgenstein* .Chesham: Acumen.

Lasonen, M. & Marvan, T. (2004). Davidson's Triangulation: Content-Endowing Causes and Circularity. *International Journal of Philosophical Studies,12*, 177– 195.

Lewis, D. (1975). Languages and Language. In K. Gunderson (Ed.), *Language, Mind and Knowledge*, (pp. 3-35). Minneapolis: University of Minnesota Press.

Lynch, M. P. (2001). *The Nature of Truth: Classic and Contemporary Prospective*. Cambridge, MA: MIT Press.

Ludwig, K. (2003). *Donald Davidson*. UK, Cambridge: Cambridge University Press.

Ludwig, K. & Lepore, E. (2005). *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Oxford University Press.

Ludwig, K. & Lepore, E. (2007). *Donald Davidson's Truth-Theoretic Semantics*. Oxford: Oxford University Press.

Malcolm, N. (1972-3). Thoughtless Brutes. *Proceedings and Addresses of the American Philosophical Association, 46*, 5 – 20.

McDowell, J. (1984). Wittgenstein on following a rule. *Synthese, 58(3)*, 325–363.

McDowell, J. (1993). Meaning and intentionality in Wittgenstein's later philosophy. *Midwest studies in philosophy, 17 (1), 40-52*.

McGinn, C. (1984). *Wittgenstein on Meaning: An Interpretation and Evaluation*. Oxford: Basil Blackwell.

Miller, A. (1989). An Objection to Wright's Treatment of Intention. *Analysis 49 (4)*, 169 – 173.

Miller, A. (1997). Tacit Knowledge. In B. Hale and C. Wright (Eds.), *A Companion to the Philosophy of Language*, (pp. 146-164). Cambridge: Blackwell.

Miller, A. (2006). Meaning Scepticism. In M. Devitt and R. Hanley (Eds.), *The Blackwell Guide to the Philosophy of Language*, (pp. 91-113). USA, MA: Blackwell Publishing Ltd.

Miller, A. (2007a). *Philosophy of Language*. London: Routledge. Revised and Expanded 2nd Edition.

Miller, A. (2007b). Another Objection to Wright's Treatment of Intention. *Analysis, 67 (3)*, 257 – 263.

Miller, A. (2009). Primary Qualities, Secondary Qualities and the Truth about Intention. *Synthese, 171 (3)*, 433-442.

Miller, A. (2010). Kripke's Wittgenstein, Factualism and Meaning. In D. Whiting (Ed.), *The Later Wittgenstein on Meaning.* Basingstoke: Palgrave Macmillan.

Miller, A. (2013a). *Contemporary Metaethics: An Introduction*. Cambridge: Polity Press. Revised and Expanded 2nd Edition.

Miller, A. (2013b). The Development of Theories of Meaning: from Frege to McDowell and Beyond. In M. Beaney (Ed.), *The Oxford Handbook of the History of Analytic Philosophy*, (pp. 656-688). Oxford: Oxford University Press.

Miller, A. & Divers, J. (1999). Arithmetical Platonism: Reliability and Judgement-Dependence. *Philosophical Studies: An International Journal for Philosophy in the AnalyticTradition, 95 (3)*, 277-310

Miller A. & Wright C. (Eds.) (2002). *Rule-Following and Meaning.* Montreal: McGillQueen's University Press.

Pagin, P. (2002). Rule-following, Compositionality and the Normativity of Meaning. In D Prawitz (Ed.),*Meaning and Interpretation*, (pp. 153-181). Konferenser 55, Stockholm: Kungl. Vitterhets Historie och Antikvitetsakademien.

Pascal, E. (1997). Davidson on Meaning, Understanding and Normativity. *Cadernos de Filosofia, 2,* 9-30.

Pascal, E. (2008). Davidson on Epistemic Norms. In C. Amoretti & Nicla Vassalo (Eds.), *Davidson, Truth, Language and Interpretation*. Munchen: Ontos Verlag.

Pietroski, P. (1994). A Defense of Derangement. *Canadian Journal of Philosophy 24 (1),* 95 – 117.

Quine, W. V. (1951). Two Dogmas of Empiricism. *The Philosophical Review, 60 (1),* 20 – 43.

Quine, W. V. (1960). *Word and Object*. US, Mass.: MIT Press.

Quine, W. V. (1968). Reply to Chomsky. *Synthese, 19 (1/2),* 274 – 283.

Quine, W. V. (1969). Replies. In D. Davidson and J. Hintikka (Eds.), *Words and objections: Essays on the work of W. V. Quine*, (pp. 292 – 345). Dordrecht: D. Reidel Publishing Company.

Quine, W. V. (1970). On the Reasons for Indeterminacy of Translation. *The Journal of Philosophy, 67 (6),* 178 – 183.

Quine, W. V. (1978). Facts of the Matter. *The Southwestern Journal of Philosophy, 9 (2),* 155 – 169.

Quine, W. V. (1981). *Theories and Things*. US, Mass.: Harvard University Press.

Quine, W. V. (1987). Indeterminacy of Translation Again. *The Journal of Philosophy, 84 (1),* 5 – 10.

Quine, W. V. (1990). *Pursuit of Truth*. US, Mass.: Harvard University Press.

Quine, W. V. (1995). *From Stimulus to Science*. US, Mass.: Harvard University Press.

Ramberg, B. (1989). *Donald Davidson's Philosophy of Language: An Introduction*. New York: Basil Blackwell.

Ramberg, B. (2000). Post-ontological Philosophy of Mind: Rorty versus Davidson. In R. Brandom (Ed.), *Rorty and his Critics*, (pp. 351-369). Malden: Blackwell Publishers Inc.

Reimer, M. (2004). What Malapropisms Mean: A Reply to Donald Davidson. *Erkenntnis, 60,* 317-334.

Ricketts, T. (2011). Roots of Ontological Relativity. *American Philosophical Quarterly, 48 (3)*, 287-300.

Rorty, R. (1970). Incorrigibility as the Mark of the Mental. *Journal of Philosophy, 67,* 399 – 424.

Rorty, R. (1999). Davidson's Mental-Physical Distinction. In L. E. Hahn (Ed.), *The Philosophy of Donald Davidson* (pp. 575-594). US, Illinois: Open Court.

Ryle, G. (1949). *The Concept of Mind*. London: The Mayflower Press.

Smart, J. J. C. (1999). Correspondence, Coherence, and Realism. In Hahn, L. E. *The Philosophy of Donald Davidson,* (pp. 109-122). Illinois: Open Court.

Soames, S. (1999). The Indeterminacy of Translation and the Inscrutability of Reference. *Canadian Journal of Philosophy, 29* (3), 321-370

Soames, S. (2003). *Philosophical Analysis in the Twentieth Century*. (Volume 2: The Age of Meaning). Princeton: Princeton University Press.

Stander, P. E. (1992). Cooperative Hunting in Lions: The Role of the Individual. *Behavioral Ecology and Sociobiology, 29, (6),* 445 – 454

Stoutland, F. (1982a). Realism and Antirealism in Davidson's Philosophy of Language I. *Crítica Revista Hispano-Americana de Filosofia, 14 (41),* 13 – 53.

Stoutland, F. (1982b). Realism and Antirealism in Davidson's Philosophy of Language II. *Crítica: Revista Hispano-Americana De Filosofia, 14 (42),* 19 – 48.

Strawson, P. (1959). *Individuals*. London: Methuen.

Stroud, B. (1998). The Theory of Meaning and the Practice of Communication", *CR´ITICA, Revista Hispanoamericana de Filosof´ıa, XXX (88),* 3–28

Talmage, C. J. L. (1997). Meaning and Triangulation. *Linguistics and Philosophy, 20 (2),* 139 – 145.

Tarski, A. (1944). The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research*, 4 (3), 341 – 376.

Thöle, B. (1993). The Explanation of First Person Authority. In Ralf Stoecker (Ed.), *Reflecting Davidson: Donald Davidson responding to an International Forum of Philosophers*, (pp. 213-247). Berlin: de Gruyter.

Turner, S. (2011). Davidson's Normativity. In Jeff Malpas (Ed), *Dialogues with Davidson: On the Contemporary Significance of His Thought*, (pp. 343-370). Cambridge, MA: MIT Press.

Verheggen, C. (1997). Davidson's Second Person. *The Philosophical Quarterly,* 47 (188), 361–369.

Verheggen, C. (2006). How Social Must Language Be?. *Journal for the Theory of Social Behaviour, 36 (2),* 203 – 219.

Verheggen, C. (2007). Triangulating with Davidson. *The Philosophical Quarterly, 57 (226),* 96-103.

Verheggen, C. & Myers, R. (2016). *Donald Davidson's Triangulation Argument: A Philosophical Inquiry*. MS. Routledge.

Whiting, D. (2007). Meaning, Norms, and Use. *Philosophical Investigations 30 (2),* 179-187.

William C. (2007). *Davidson on First Person Authority and Knowledge of Meaning. NOUˆS, 41 (2),* 157–177.

Wilson, G. (1994). Kripke on Wittgenstein and Normativity. *Midwest Studies in Philosophy, 19 (1),* 366 –390.

Wilson, G. (1998). Semantic Realism and Kripke's Wittgenstein. *Philosophy and Phenomenological Research, 58, (1),* 99 – 122.

Wilson, G. (2006). Rule-Following, Meaning, and Normativity. In E. Lepore and B. C. Smith (Eds.), *The Oxford Handbook of Philosophy of Language*, (pp. 151-174). Oxford: Oxford University Press.

Wittgenstein, L. (1922/2003). *Tractatus Logico-Philosophicus*. (C. K Ogden, Trans.). Barnes & Noble Publishing, Inc.

Wittgenstein, L. (1953/2001). *Philosophical Investigations*. (G. E. M. Anscombe, Trans.) (Third Ed.). Oxford: Wiley-Blackwell.

Wright, C. (1984). Kripke's Account of the Argument against Private Language. *The Journal of Philosophy, 81 (12)*, 759 – 778.

Wright, C. (1987). On Making Up One's Mind: Wittgenstein on Intention. In P. Weingartner and G. Schurz (Eds.), *Logic, Philosophy of Science and Epistemology-Proceedings of the Xltb International Wittgenstein Symposium, Kirchberg*, (pp. 391-404). Vienna: Holder-Pichler-Temsky. Reprinted in C. Wright (2001a), pp. 116-142, to which page references apply.

Wright, C. (1989). Critical Notice of Colin McGinn's *Wittgenstein on Meaning*. *Mind, 98*, 289 – 305. Reprinted in Miller A. and Wright C. (2002), pp. 108-128, to which page references apply.

Wright, C. (1992). *Truth and Objectivity*. Cambridge, US: Harvard University Press.

Wright, C. (2001a). *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*. Cambridge, US: Harvard University Press.

Wright, C. (2001b). The Problem of Self-Knowledge (I). In C. Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*, (pp. 319-344). Cambridge, US: Harvard University Press.

Wright, C. (2001c). The Problem of Self-Knowledge (II). In C. Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*, (pp. 345-374). Cambridge, US: Harvard University Press.

Wright, C. (2001d). Postscripts: Study Note on Wittgenstein on the Nature of Philosophy and Its Proper Method. In C. Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*, (pp. 433-443). Cambridge, US: Harvard University Press.

Wright, C. (2001e). Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics. In C. Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*, (pp. 170-214). Cambridge, US: Harvard University Press.

Wright, C. (2002). Meaning and Intention as Judgement Dependent. In Miller A. and Wright C. (Eds.), *Rule-Following and Meaning*, (pp. 129-140), Montreal: McGill-Queen's University Press. The paper is taken from Crispin Wright (1989), "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics", in A. George (Ed.), *Reflections on Chomsky* (pp. 246-54). Oxford: Blackwell.

Zalabardo, J. (1989). Rules, Communities and Judgements. *Crítica: Revista Hispanoamericana de Filosofía, 21(63)*, 33–58.