Mariann Solberg

Belief, Truth, and the Enigma of Error

Thesis submitted for the dr.art.-degree
Department of Philosophy
Faculty of Social Science
The University of Tromsø
February 2006

Contents

Acknowledgements

I would like to express my gratitude to my advisor Jan Harald Alnes who has guided my work since midway in the period of the scholarship. All articles except article 2 have been written or rewritten under his supervision. His enduring demands for clarity and his uncompromising line of avoiding handy short cuts of criticism to the advantage of always taking on charitable readings have been very important for my work on the thesis. Jan Harald's competent and thorough response to my writing has always been instructive, and the related hour-long discussions have been great fun and invaluable for pointing out the direction of the work. I would like to thank Johan Arnt Myrstad, who was my advisor for the two first years of the scholarship. Arnt gave generously of his time and he inspired me to believe in my own thinking.

I thank The Norwegian Research Council and The University of Tromsø for rendering this thesis possible by their financial support. My thanks go also to Nordlandsforskning in Bodø for letting me borrow an office in the final stages of my work. Thanks to family and friends for encouragement, and patience with me while I have been working on this thesis. The Department of Philosophy at the University of Tromsø have been my professional environment since the years of undergraduate study, except one year that I spent as visiting scholar at The Department of Philosophy at The University of California at Berkeley. I thank faculty, staff and students at both institutions for good companionship, and The Norwegian Research Council for generous financial support of our stay in Berkeley.

While I was in Berkeley Donald Davidson took time for discussions, and for reading and commenting upon my written material on his philosophy. Fortunately, I had misunderstood quite a lot. In this way I learned a good deal about what he had *never said*, what he *never could have said* and what he anyhow had *never meant*. Eventually, I managed to write something about what he *had* said *and* meant, but by then I was of course also far into the project of defending different theses of his, rather than critically discussing them from the outside. It was a great experience. It is sad that I due to the unexpected passing of Donald on August 30. 2003 will not have the opportunity of receiving his comments upon this thesis. I am grateful that I had the opportunity to work with him.

One of my first philosophy teachers, Harald Johannessen, deserves thanks for trying to teach me the essentials of Gottlob Frege's philosophy by co-reading quite a few articles during weekly seminars for quite a long time. My pace and ways of reading philosophy came to be more or less settled by that experience. Many friends, teachers and colleagues have since taken part in shaping the thoughts and words of this thesis by reading and commenting upon drafts, by suggestions, by criticisms and questions related to presentations, by discussions of the problems involved, and by advice on composition. I am indebted to Roar Anfinsen, Rani Anjum, Maria Baghramian, Eivind Balsvik, Solveig Bøe, Steinar Bøyum, Elizabeth Camp, Eddie Cushman, Eyolfur Emilsson, Kjersti Fjørtoft, Olav Gjelsvik, Kathrin Glüer, Ånund Haga, Jon Hellesnes, Mikael Janvid, Hildur Kalman, Gayle Kenny, Arpy Khatchirian, Jonathan Knowles, Elin Kolsvik, Tarjei Mandt Larsen, Frode Sirnes Larsen, Trygve Lavik, Erik Lundestad, Jacob Meløe, Kevin Mulligan, Petter Nafstad, Stephen Neal, Elina Nurmi, Peter Pagin, Bjørn Ramberg, Richard Rorty, John Richard Sageng, Merethe Sollund, Øyvind Stokke, Barry Stroud, and Folke Tersman.

I thank my husband Arne for two decades of discussions on evolutionary epistemology. My elder son Martin took on the task of proofreading the articles. It has been fun to see his reactions to the research that I've been on to "for as long as he can remember", and I am thankful for his elaborate work on my thoughts and words. My younger sons, Jonas and Eivind made me rethink my beliefs about the nature of language, thought and intentionality more than once. But, most of all, I thank my family for making my life good.

I would like to express my appreciation to the editors and publishers concerned for the permission to reprint articles. These include Unipub for the permission to reprint "Davidson and the Role of Error" from *Non-Conceptual Aspects of Experience*, Unipub, Oslo 2003, pp 119-134, edited by Hallvard Fossheim, John Richard Sageng and Tarjei Mandt Larsen, here appearing as article 2, and Universitetsforlaget for the permission to reprint "Velvillighetsprinsippets epistemiske status" from *Norsk filosofisk tidsskrift*, nr 1-2, Universitetsforlaget 2004, pp. 61-76, here translated into English with the title "The status of the Principle of Charity", appearing as article 3.

At its best, doing philosophy means shaking the ground upon which to understand. The moments of "mind quake" makes it all worth the while.

Tromsø, february 2006

Marianne Solberg

Introduction

1. Davidson's overall project

Over the years, Donald Davidson's project has come to be centered around the question of the difference between dispositions to react on the one hand, and the ability to conceptualize and think, on the other. Or we could say that he has focused on the difference between having a disposition versus having a concept, and between having a reaction versus having a norm. According to Richard Rorty there need not be a difference. He holds the view that all there is to normativity is regularity of behavior. Why would Davidson, then, consider these differences so important?

The debate on intentionality and explanation of human action between logical empiricism, represented by Carl Hempel, on the one hand, and philosophers inspired by the late Wittgenstein, represented by Elizabeth Anscombe, on the other, seems to have been formative for Davidson's project. The former were taken to have the project of reducing human intentionality to physics, and thus of leaving intentional aspects out of the explanation of human behavior. The latter were taken to have the project of rescuing human intentionality at the expense of humans' material and causal conditions, considering them unimportant for the understanding of human behavior. I'm sure that each of these descriptions is rather unfair. Nonetheless, the point here is that Davidson took on the project of uniting the two presumed contradictory views. In his theory of action, he is attempting to do this by joining reasons and causes in the explanation of action, as expressed in the slogan "the reason for an action is its cause". This means that for Davidson, material monism is to be *settled* by his philosophical project, and it could therefore not be presupposed.

The aims of his anomalous monism thesis lie in showing that there is one reality, but it is also vital for Davidson to point out that we can reach it in two ways. We can reach it either by way of intentional concepts or by way of causal concepts. If regularity of behavior is all there is, as Rorty has suggested, the interpretational research program, initiated by Davidson, will have lost its subject matter. That would of course be sad, but except for that, why couldn't we settle for regularity of behavior only? Why insist that there is an autonomous area for intentionality and normativity? *My* main reason for resisting the strategy of settling for regularity of behavior is that we no longer would be able to give an account of what it *is* to understand intentional and normative creatures. If we erase human interests, reasons, orientations, values, assessments, meanings and beliefs from the board, there will be no point in interpreting creatures in order to understand them. "Interpretation" of regularity of behavior, without assessment, would, if it were possible, have no point. It would give us what we do not want and need, and in a sense it would stand out as a category mistake. It would not give us an understanding of the actions and values and beliefs of individual speakers, it would not give us an understanding of what intentionality is, and it would not give us an understanding of what normativity is.

The explanation of intentionality is the project that unites Davidson's philosophical reflections on language, mind, action and epistemology. According to my reading, his overriding question is: How can we explain intentionality and normativity in a physical world? His unified theory of action, mind and language is given in the form of a theory of radical interpretation. In these articles, however, I say almost nothing about Davidson's *unified* theory of thought, meaning, belief and action. When I present his theory of radical interpretation, I refer mainly to articles on theory of interpretation, such as "Truth and Meaning" and "Radical Interpretation", where the theory was still a theory of meaning and belief, or with a collective term, theory of language. The reason for this is clearly not that the perspectives of later articles, such as "A Unified Theory of Thought Meaning and Action" or "A new Basis for Decision Theory", in which Davidson sketches how to unite decision theory with the earlier work on radical interpretation, are uninteresting or unimportant. However, to come to grips with the particular features of Davidson's epistemology that I seek to clarify I have assumed that this particular extension of the theory is inessential. For the theory of interpretation, the most prominent change is that the radical interpreter (who was previously looking for the speaker's hold-true-attitudes) now has to search for the speaker's preferences among sentences. This is then a change from the attitude of "holding true" to the attitude of "preferring true". This implies that the radical interpreter should note the speaker's preference of the truth of a particular sentence over another. The final result of the extended theory is that in addition to developed views on what the speaker mean and believe, the interpreter will have views on what the speaker wants and what the speaker is doing in addition to developed views on the speaker's meanings and beliefs. Where I find that the extension of the theory might have made a difference, I mention it explicitly. In all other cases the assumption should remain that the results that I come to are indifferent to the restricted and extended version of the theory.

2. Davidson's epistemological project

I tentatively read Davidson's epistemological project as an attempt at warding off skepticism about knowledge of the external world, about other minds and about the contents of our own minds. According to Davidson there are three forms of propositional knowledge. There is subjective knowledge, which is knowledge of the contents of our own minds, or what we more often call first person knowledge. Then there is intersubjective knowledge, which is knowledge of the contents of other minds. Intersubjective knowledge is either second person knowledge, i.e. what another knows about the contents of my mind, or third person knowledge; what I know about the contents of another mind. Finally, there is objective knowledge, which is knowledge of the shared environment, or "knowledge of the rest of the world of nature".[1] These three varieties of knowledge differ from each other in various respects, but are, according to Davidson, all objective "in the sense that their truth is independent of their being believed to be true" and also "objective in the sense that it could for the most part be expressed by concepts which have a place in a publically shared scheme of things". According to Davidson, none of the three varieties of knowledge can be dealt with in isolation, and none of them can be reduced to the others.[2] They are mutually irreducible. This means that Davidson hampers the skeptic by refusing to let him take a stand in only one (or two) of these forms of knowledge. However, there are no knockdown arguments to be found against the skeptic in Davidson's thinking, and perhaps one cannot even claim that Davidson really *answer* the skeptic. Answers to epistemological questions are often indirectly given, and some epistemological questions are simply dismissed. My contention is not that Davidson was first and foremost an epistemologist, but rather that he has given interesting answers to certain epistemological worries, and quieted others.

---

[1] See the fist page of the Introduction, page xiii of Donald Davidson: *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.
[2] See Davidson's "Three Varieties of Knowledge", p. 208 in his *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

3. My project

These articles deal with different problems within Donald Davidson's philosophy on belief, truth and error. In the hands of Davidson the methodology of radical interpretation is a device to explain how we achieve belief, objective truth and thus knowledge. Through the thought experiment of radical interpretation, Davidson has been able to shed light on some of the conditions of knowledge, and of the nature of human thought in general. The articles of this collection are not studies in epistemology in the sense of clearing out what knowledge is or what it takes to have justified beliefs. Still, the problems discussed will be highly relevant for such questions.

According to Davidson, the source of human knowledge is to be found in the interaction between at least two interpreting minds and a world. He calls this model "triangulation", and a key point is that other linguistic minds are inevitable for propositional knowledge. A major aim in my work is to analyze certain factors of the epistemic rationale of the model of triangulation. I recover some of the epistemic problems solved by triangulation and point out some of the models more problematic characteristics, such as the phenomenological insensitivity towards non-conceptual experience.

One of the conditions for knowledge within Davidson's theory that I find particularly interesting is the principle of charity. In Davidson's thinking this principle has always been crucial. I analyze the principle's function within his theory and take a closer look at the epistemological function of the theory of radical interpretation itself. Highlighting the function of radical interpretation in an epistemological setting generates a new perspective on this core principle. The received reading of the epistemic status of the principle of charity has long been that it is an apriori and constitutive principle. I challenge this view, and find that the epistemic status of the principle of charity is a useful "acid test" for placing Davidson's epistemological enterprise in the philosophical terrain in the stretch between naturalizing and transcendental tendencies.

The principle of charity and the model of triangulation are two of the main philosophical devices focused in the present interpretation of Davidson's epistemology. Both of these devices originate in the philosophy of Willard van Orman Quine. The first article below therefore introduces the debate between Quine and Davidson, in particular on epistemological questions, while the next two articles deal explicitly with triangulation and charity. Together the three articles give an account of the basic methodology of interpretation cum epistemology.

The next three articles discuss important factors for an evaluation of Davidson's contributions to epistemology: In the fourth article I discuss a specific problem related to triangulation, viz. the problem about normativity and how to explain the existence of error in the triangular situation. Unless we are able to distinguish between correct and incorrect belief, and correct and incorrect interpretation, there can be no talk of objective truth and knowledge. Whereas Davidson has given an explanation of the conditions of correct understanding and correct interpretation, accounting for error is a problem within the theory. Indeed, the problem of explaining how error is possible has over time come to seem close to an enigma. Davidson has taken it that if we can explain error, we can explain intentionality. An explanation of intentionality would also be an explanation of the phenomenon of normativity. A solution to the problem of error would therefore be of great interest and have rich consequences. Also the fifth article deals with error, and in specific the question of how a creature can obtain the idea that error is possible, and how we could be in position to know that another creature has an awareness of the possibility of error. The sixth article discusses the role of second-order beliefs in Davidsonian epistemology. Beliefs about beliefs is an interesting case for testing how far Davidson's theory can be pushed in the direction of naturalization.

The Synthesis brings together the overarching methodology of interpretation, through triangulation and charity, and the more specific problems of error and second-order belief. The big issue in the Synthesis is however the discussion of naturalism and naturalized epistemology in relation to the results of my investigations in the articles. Throughout my work, a reading of Davidson's enterprise as genuinely gradualist has come to the fore. There are few, if any, on-or-off positions that can be taken (or anyhow, consistently defended) within his theory. The Synthesis also discusses Davidson view on truth, and objective truth, in relation his otherwise fallibilist and gradualist views.

4. An Internal and an External Perspective on Intentionality

I distinguish between an internal and an external perspective on intentionality.[3] In an internal approach intentionality is given; one is explanatorily on the inside. An internal approach could for instance take the form of laying out relations between core concepts of intentionality, such as belief, second order belief, desire, and wish, truth, knowledge. In an external approach the non-intentional is given, and the task becomes to explain the emergence of the intentional based on the non-intentional. In triangulation an explanation of intentionality is approached from an external perspective. Triangulation gives a description of how intersubjective knowledge contributes to objective knowledge, and as such it deals first and foremost with how we come to have objective knowledge of the world.[4] According to triangulation, we come to have objective knowledge because of the interaction between two speaker-interpreters, and because they have a common world which each of them react to. When speakers assess each other's linguistic and other reactions to a common feature of the world, and when they share these assessments with each other, a condition for intentionality is in place. A question that I raise is whether triangulation can explain the transition from non-intentional to intentional.

An external approach is primarily described in article 2, 4 and 5, where I discuss triangulation and the problem of error. Article 6 takes an internal approach, and describes how some of our intentional concepts are interrelated.

5. About each of the articles

The articles 2, 4, and 5 all address the problem of error in one way or another. The order in which they stand indicates a radicalization of the treatment, as well as a narrowing of the problem. The chronological order in which they were written would however be 2, 5 and then 4. Article 1 has not been publically presented before, and has been in the making since 2001, and up to the present. Article 2 was written in 2000, during my stay in Berkeley. It was written as a paper for the International Philosophy-Symposium "Non-Conceptual Aspects of Experience" arranged by the Nordland Academy of Arts and Sciences in association with a committee of Norwegian research fellows in philosophy in

---

[3] See e.g. article 4.

[4] Intentional creatures and the contents of their mind are also included in the concept of 'world'.

Melbu in July of 2000. It was published in an anthology in 2003, and I have made only a few minor phrasing changes and technicalities that have little or no effects on its content. Article 3 was written in 2003 as an invited essay to a special edition on Donald Davidson of *Norsk filosofisk tidsskrift*, and was presented at the Dr. art. / Research-seminar at the Department of Philosophy, The University of Tromsø in March 2004. The first draft for article 4 was made during 2004, and it was finished in 2005. It has not been published before. An early version of article 5 was written in 2001 and presented under the title "Error and Triangulation" in the International Workshop "Davidson and Triangulation", arranged by Professor Bjørn Ramberg at the University of Oslo in March the same year. Article 6 has been in the making since 2000, and it was finished in 2005. Some of the material for the article was presented in my paper "Intensjonalitet og naturalisme", ("Intentionality and Naturalism") which I read in the dr. art. / Research-seminar at the Department of Philosophy at The University of Tromsø, November 2000. The Synthesis was written in 2005 and finished in 2006, but some of the material is taken from a paper titled "Davidson, Naturalism and Constitutivity", presented at an International Philosophy Seminar "Contemporary Issues in Metaphilosophy" at the University of Tromsø in December 2001.

1. Radical Translation and Radical Interpretation[1]

It has been commonplace to say that Davidson's radical interpretation is an adoption of Quine's radical translation, with a Tarskian truth-theory added in order to make it into a compositional theory of meaning. Davidson's radical interpretation is clearly built upon Quine's radical translation, and, according to Davidson, one of the differences is marked by "a greater emphasis on the explicitly semantical [in radical interpretation]".[2] My aim is not to dispute this understanding. However, I argue that, in the assumed insignificant differences between radical translation and radical interpretation, we find reasons for the epistemologically significant differences between Quine and Davidson on a macro-level. Quantification structure is one of the most perspicuous "details" in this respect. I argue that the choice of ways of quantification is the background for the debate on whether we acquire our beliefs by sharing and building upon proximal *or* distal stimuli. I hold that the importance of these "insignificant details" is due to the degrees of indeterminacy in the two approaches and to the degree of elbowroom for the skeptic. Furthermore, these details affect the question of *the basis of* intersubjectivity. While Quine either cannot guarantee intersubjectivity (in epistemology) or base it in a pre-established harmony (in semantics and language-learning), Davidson bases intersubjectivity partly on dispositions and partly on the employment and exchange of language, described as the process of triangulation. In section 1 through 4, I give an account of characteristics of radical translation and radical interpretation, where I pinpoint the most relevant differences. In section 5, I thoroughly discuss the question of quantification. In section 6, I present Quine's issuant "discomfort" with his own solution as concerns proximal stimuli, as it originally was presented in his *Word and Object*. My concern is to show that (and how) Davidson's epistemology grows out of his work on semantics and intentionality, and I consider Quine's radical translation a necessary backdrop for understanding what kind of project radical interpretation is. I argue that when Davidson expanded Quine's theory of radical translation into his own theory of radical interpretation, the reason was that Davidson wanted to give a theory of meaning.

---

[1] This article is a product of five years of discussions with Jan Harald Alnes on Quine, Davidson and the relation between them. This does not mean that he agrees with the conclusions here drawn. However, I have to a large extent benefited from Jan Harald's insistence on charitable readings (especially of Quine) when dogmatic categorizations came easy.

[2] See Donald Davidson: "Radical Interpretation", p. 126 in his *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

This expansion, and the steps required to give a compositional theory of meaning, however, made Davidson grow out of and divert from Quine's epistemology.

1. Radical Translation

The "radicalism" in Quine's radical translation is that the language to be translated is the language of a hitherto unknown people. This means that there are no given similarities, no previous translations or information about the people and their language. The situation of Quinean translation, described in chapter two of his *Word and Object*, is that of a field-linguist going out into the jungle aiming to prepare a dictionary (or, rather, a manual) for translations between the language spoken in the jungle and English. The language of a "hitherto untouched people"[3] is chosen because it avoids any assumptions about the ontology and meanings of their language. One of the advantages of this situation is that it puts us in a position to get a picture of linguistic meanings solely based on people's linguistic behavior and their interaction with their surroundings. Systematic correlations of linguistic behavior and changes in the immediate surroundings are thus what we face in radical translation. The advantage of a behavioral study, according to a Quinean view, is that it does not have to deal with entities (or notions) such as meanings, beliefs, desires, intentions and the like. In such a study we can do better, we can take an extensional stand instead and concentrate on "...men's dispositions to respond overtly to socially observable stimulations"[4], and forget about meanings, at the least in the form of reified entities or in the form of "museum pieces". In this way, methodological considerations determine the ontology, and if a serious and modest methodology can make do without intentions and meanings in the form of reified entities, there is no need for such entities.[5] This is expressed in the Quinean slogan "No entity without identity". In order to meaningfully speak of an entity, we have to be able to individuate it, and not the least, we have to be able to re-identify it. A consequence hereof is that if we have no way of identifying and re-identifying the presumed entity in question, there simply cannot be an entity. In line with this view, the motivation for radical translation is that Quine wants to strip away metaphysical notions of meaning from the study of meaning and reference, in order to

[3] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 28.
[4] Ibid, ix.
[5] Davidson continues the battle against reification, but he will not as easily as Quine let go of the classical notions of meaning, belief, intentionality and normativity.

clarify "how much of language can be made sense of in terms of its stimulus conditions..."[6] This means that empirical acquaintance with a specific spoken language has to be our basis for the theory. Quine wants to show that there is every reason to be suspicious about language-transcendent meaning. The purpose of the thought experiment of radical translation is to clarify the evidential basis for linguistic meaning, and one of the central conclusions of the investigation is that "the enterprise of translation is found to be involved in a certain systematic indeterminacy..."[7]

Seen from an epistemological point of view, the situation of radical translation opens the possibility of studying the relations between input (stimulus-conditions) and output (conceptual scheme), or as Quine puts it: "how surface irritations generate, through language, one's knowledge of the world".[8] In radical translation we do not avail ourselves with theories and preconceptions of which we have no entitlement, and the situation will in this sense be an ideal basis for a serious and unprejudiced study of human intentionality, be it linguistic meanings or acquirement of beliefs and knowledge. Radical translation is Quine's "method" of studying semantics as well as epistemology. We should note that Quine seems to recommend a focus on both "socially observable stimulations", and "surface irritations". "Socially observable stimulations" would initially be ordinary middle-sized objects, since objects as these are the stimulations that we would relate to first in a situation of radical translation and also in language learning. On the other hand, it all seems to be about "surface irritations", and the surfaces spoken of are the sensory surfaces of individual human bodies. For Quine, as could be read in his *Word and Object* (1960), both the proximal (surface irritation) and the distal stimulations (objects in particular) are of interest.

The confined aim of the radical translator - to prepare a manual that enables translation between the jungle-language and English - is not identical to the confined aim of Davidson's radical interpreter, who aims to prepare a truth-theory for the speaker, or, in Davidson's words, to prepare "...a Tarski-style characterization of truth for the speaker's

[6] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 26.
[7] Ibid, p. ix. This "systematic indeterminacy" is called "the thesis of the indeterminacy of translation", and it has consequences for possible meanings as well as for references.
[8] Ibid, p. 26

language, and a theory of his beliefs."[9] This means that the aim for the radical interpreter is to confirm a theory of truth for the language of the speaker under conditions that ensure that the theory can be employed in an interpretation of the sentences of the speaker. A radical interpreter will try to interpret the speaker based on what the speaker holds to be true, and the interpreter has to be able to recognize the attitude of 'holding true' or, in other words, the by Davidson assumed identical attitude; the attitude of believing. By means of the recognition of this attitude, the radical interpreter develops a theory of the speaker's beliefs and at the same time, the meanings of the speaker's sentences. We shall soon see how this is supposed to be accomplished.

But first, why doesn't Davidson just say: "What an interpreter needs to know to be able to interpret the speaker is the meanings of all the meaningful expressions of the speaker"? In line with Quine, Davidson holds that "Appeal to meanings leaves us stranded further than we started from the non-linguistic goings-on that must supply the evidential base for interpretation..."[10] So we cannot rely on some kind of linguistic concept, such as meaning or synonymy; our evidence has to be accumulated without previous knowledge and essential use of such concepts. This is most clearly an adoption of the empirical and methodologically restricted line that we find in Quine's Radical Translation. This line has the implication that rich notions of meaning are not an essential part of our observational apparatus for taking in the world. The kind of evidence we need "...must be of a sort that would be available to someone who does not already know how to interpret utterances the theory is designed to cover..."[11] This is to say that we are not to start out with any preliminary attributions of actual particular meanings, and that empirical acquaintance with a specific language is necessary. We are supposed to start out with something non-semantic and end with something semantic and with intentional attributions. It is also presupposed that all relevant facts for a theory of the speaker's meanings are already present in the material complexity. However, it is important to note that the interpreter's given information is both non-linguistic and linguistic, even if he has no initial semantic preconceptions.

---

[9] Donald Davidson: "A Coherence Theory of Truth and Knowledge" (1983) pp 307-319 in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore, Blackwell, Oxford 1986, p. 315.
[10] Donald Davidson: "Radical Interpretation" (1973) pp 125-139 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 126-7.
[11] Ibid, p 128.

There are different objectives, though, for each of the two experiments. In Quine's radical translation we find an overall emphasis on the input-output-relation of stimuli and knowledge and an engagement with the question of justification. In radical interpretation, Davidson focuses on finding a truth theory for the language of the speaker - thereby establishing a theory of meaning for his language - and at the same time finding a theory of his beliefs.

2. Radical Interpretation

Now, radical interpretation is not a universal method of interpretation of any utterance in any language, or of any kind of creature for that matter. The radical interpreter starts out with some regularity in behavior, some regularity in some physical being, and an aim to end up with complex semantic ascriptions, we might expect that any creature with a certain behavior would do. Jane Heal, in her article "Radical Interpretation," holds: "A theory of radical interpretation should be applicable to giant octopuses or superbeings emerging from their spaceships as well as to newly encountered human beings." [12] That Davidson's theory of interpretation is applicable to only human beings is here considered a serious lack of generality. According to Heal's view, the specific orientation towards humans is a flaw in Davidson's theory of interpretation. Heal holds that the project of radical interpretation is far from radical unless it could be applied to any being. Davidson's starting point is the attitude of holding true and on this point Heal criticizes Davidson for not taking the totality of physical facts as his base instead. She holds that this "would be a much less tendentious place to start."[13] I would say that her criticisms rest on a serious misunderstanding of Davidson's thinking. When she criticizes the radical interpreter's starting point in the recognition of the attitude of holding true, she makes certain assumptions about Davidson's ways of doing philosophy. She assumes that Davidson is seeking a starting position in radical interpretation where any creature could be radically interpreted, which would amount to a universal theory of interpretation. Her comments suggest that only then will the interpretational approach be truly radical. But, as I will argue more substantially later, Davidson's thinking is generally not a specimen of universalism and absolutism. In

[12] Jane Heal: "Radical Interpretation", in Bob Hale and Crispin Wright: *A Companion to the Philosophy of Language*, Blackwell, Oxford 1997, p. 184.
[13] Ibid, p. 185.

addition, the self-imposed restriction that is being laid upon the starting point of radical interpretation, namely to start out with something non-semantic *is* fulfilled.

The project of radical interpretation rests entirely on *belief*, and the possibility of identifying particular beliefs in the creature subject to interpretation. The notion of belief is crucial to the project of Radical Interpretation (and to Davidson's thinking in general). The importance of *belief* can hardly be exaggerated, and the role that this attitude plays is often overlooked in comments on Davidson.[14] The important point here is, however, that the creature in question has to have a whole lot of beliefs (and other attitudes) in order for us to be able to locate and identify one single belief. Davidson maintains that a belief can only be identified due to its relation to other beliefs in the same set of beliefs, and for a creature to have such a complex and interwoven set of attitudes, and for us to be able to discover it on a normal empirical basis, amounts to that the creature has a developed language. In other words, Davidson holds that we do not know quite what it takes to have such a complex set of attitudes, but fully developed linguistic abilities would do. Hence it is obvious that human beings are the kind of creature that the theory is applicable to, and the attitude of holding true (or sentences held true) is not a tendentious place to start.

Further, Davidson holds that the method of radical interpretation can give an exhaustive account of the meanings of a particular language L (which is also called an idiolect). "What a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes."[15] As we have noted, the basic features of radical interpretation are not very different from the features of radical translation, and Davidson says that radical interpretation "...adds little to Quine's program of translation, since translation of the speaker's language into one's own plus a theory of truth for one's own language adds up to a theory of truth for the speaker."[16] The issue of a theory of truth for either speaker or interpreter is not of particular interest to Quine's project, though. But Davidson adds that "...the shift to the semantic notion of truth from the syntactic notion of translation puts the formal restrictions of a theory of truth in the foreground, and

[14] See article 6 for more on this.
[15] Donald Davidson: "A Coherence Theory of Truth and Knowledge" (1983) pp 307-319 in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore, Blackwell, Oxford 1986, p. 315
[16] Ibid.

emphasizes one aspect of the close relation between truth and meaning."[17] What seems clear is that Davidson's more semantic interests is the focus of radical interpretation, while the Quinean epistemological interests are the ones first of all to be taken care of in the method of radical translation. However, this is not to say that epistemological interests are *in general* left behind in Davidson's theory.

According to Davidson, "...the point of language is communication".[18] Apart from this, he maintains it is possible to find out about the linguistic meanings of other people's sentences only by observing overtly responses to "socially observable stimulations"[19] But how do we really go from overtly responses to socially observable stimulations to understanding and smooth communication between speakers? Quine does not seek normal communication and understanding; he focuses on the relation between input and output in single individuals, and the scene *is* social, as the investigation is done from a third person perspective. But the model of radical translation is not reciprocal, and Quine is not trying to answer the question of how we go from overtly responses to smooth communication, interaction and understanding. Davidson is, however, and he starts off where Quine leaves. From a Davidsonian point of view, a problem with the experiment of radical translation is that it does not deal with understanding and communication explicitly. It does not explain how we could come to understand each other, and how it comes about that we actually manage to communicate. This interest in understanding is closely connected to the aim of providing the semantics of the language to be interpreted, or rather; it is the other way around.

In order to take care of semantic interests, the theory has to solve the problem of exposure of the semantic structure of sentences. Davidson says "A theory of translation must read some sort of structure into sentences, but there is no reason to expect that it will provide any insight into how the meanings of sentences depend on their structure"[20] A theory of interpretation will, on the other hand, have to reveal "...significant semantic structure..."[21], if the theory is to be satisfactory. We must, for instance, be able to show that (and how) the

[17] Ibid.
[18] Comment in seminar, UCB, April 31. 2000.
[19] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. ix.
[20] Donald Davidson: "Radical Interpretation" (1973) pp 125-139 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 130.
[21] Ibid.

interpretation of complex sentences depends on the interpretation of words and phrases. Quine's radical translation was not into this business. Davidson obviously had to add something to radical translation in order to fulfill the mentioned aim, and it was here that he found that he needed Tarski's theory of truth for formal languages. A theory of truth in Tarski's style is, according to Davidson, "...a structurally revealing theory of interpretation for the object language..." Such a theory "...entails, for every sentence *s* of the object language, a sentence of the form: *S* is true if and only if *p*. Instances of the form (T-sentences) are obtained by replacing '*S*' by a canonical description of *S*, and '*p*' by a translation of *S*."[22] An example of an instantiation: *Schnee ist weiß* is true if and only if snow is white. With this extension Davidson was definitely closing in on the aim of providing a theory of interpretation for particular languages with a semantically revealing structure. The theory is (but not only due to this, of course) an example of a compositional theory of meaning.

One of the important differences in the transition from radical translation to radical interpretation was that Davidson's semantic theory had the form of a theory of truth. Davidson considers truth basic in the sense that it is primitive. He holds that "Truth is a single property which attaches, or fails to attach, to utterances, while each utterance has its own interpretation; and truth is more apt to connect with fairly simple attitudes of speakers."[23] He says he wants to "extract" an account of interpretation from truth. He thinks that the attitude of holding a sentence true is a good place to start as it enables interpretation to rest on evidence that does not presuppose knowledge of meanings and detailed knowledge of belief. The holding true of a sentence is itself, according to Davidson, a belief, but it is a single attitude that is "applicable to all sentences", and Davidson argues that we may, as interpreters, be able to identify this attitude before we can further interpret a speaker. Radical interpretation does not only seek the meaning of sentences, but also psychological states, such as beliefs, and this means the pair of truth and meaning is married without parole and of equal standing. The one could not be found without the other. Psychological states are in this way also brought in.[24]

---

[22] Ibid.
[23] Ibid, p. 134.
[24] This was new with Davidson's article "Radical Interpretation" (1973). In his "Truth and Meaning" (1967) printed as pp 17-36 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, it was still all about linguistic items.

As to the matter of Quine's interest in the question of truth, Quine would say that agreement (conformity to verbal dispositions), simplicity, maximizing truth, and psychological plausibility are all "...virtues that we can seek and agree on, in devising a manual of translation..."[25] It seems that "maximizing truth" is one among several virtues, but of no special or particular concern to Quine's project, as it is to Davidson's project. Quine's solution here would not do for Davidson, because translation and a matching of sentences, holophrastically, is not what Davidson seeks to accomplish. Davidson's project is interpretation of a speaker or of his language, and he therefore seeks an understanding of the speaker, and, specifically, he wants to find a detailed compositional semantic theory for the speaker. A translation-manual, which sets up the synonymy-relations between sentences in the two languages, deals with, or, anyhow, could potentially deal with, another topic. In translation we are after the relation between two languages, while in interpretation we seek an understanding of another person, and, in a theory of interpretation, we want to focus on *what it takes* to understand another person, in the sense of understanding his particular meanings and beliefs.[26] In radical translation, we could, if we push things, say that we can know that this particular sentence translates with another particular sentence, without knowing what any of the sentences means, because the manual can be given in two, for us, unknown languages. In Radical interpretation, our own language (the interpreters language) has to be both the meta-language and the language to which we translate or interpret the sentence in the object language. Hence, it is only the object language that could be unknown to us (but it certainly does not have to be unknown). The consequence of bringing in truth theory and truth conditions is that we have to deal with the *use* of the sentences; we cannot be satisfied by a sheer correlation of sentences.

3. Stimulus meaning versus truth

There is a long way to go from the Quinean stripped behaviorist scene of radical translation to Davidson's (also behavioristically based) ascription of psychological states and attitudes, as well as sentence meanings. Both Quine and Davidson reject meanings as fixed and language-transcendent, and Quine's minimal notion of stimulus meaning is never

[25] In Lewis Edwin Hahn (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, p. 155.
[26] In Davidson's later unified theory of action, mind and language, where a theory of action is fused with the truth-theory for a person, he is also after a person's desires, evaluations and general motivation.

close to anything like it. The basis of stimulus meaning on the one hand and the basis for ascription of meaning and belief in Davidson on the other are surely different. Stimulus meaning has an empirically unproblematic objective basis, as it is the ordered pair of assenting and dissenting stimulus meaning for a sentence of a subject. The stimulus meaning for the sentence of a subject is, in short, his disposition to agree or disagree that a given sentence can be a response to a given stimulation that is present. Affirmative stimulus meaning of a sentence for a speaker is then "the class of all the stimulations that would prompt his assent."[27] The negative stimulus meaning may be similarly defined with 'assent' and 'dissent' interchanged. However, we should note that the notion of stimulus meaning is subjective in the sense that the stimuli and the stimulation of the speaker's receptors are private. Stimuli are not common. It relates directly to one single speaker, and is therefore not intersubjective. But Quine notices that somehow there must be some similarity somewhere, since our outputs and standards of stimulus meanings tend to be quite alike within languages such as English. In *Word and Object* he ascribes this to the rough uniformity of our neural receptors. This means that the similarity in output is assumed to have its cause in our input apparatus, and not in the world.[28]

Davidson's use of a recursive truth theory with quantification in collating the meanings and beliefs of a speaker enables the interpreter to go much further into the intentions of the speaker. But, it is important to note that this could not be done unless the principle of charity enabled me as an interpreter to assume that the speaker's beliefs are more or less like mine, and unless I could assume that the logical structure of his language is more or less like mine. Davidson describes his own method of radical translation and interpretation as a "...process of devising a theory of truth for an unknown native tongue..." We start by trying to find the best way of reading our own logic (which may mean "the logical

---

[27] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass., 1960, p. 32.

[28] In 1990, we hear that Davidson was unsatisfied with the solution he had taken to be Quine's solution at that time. He points out the problems that Quine later has described as a riddle. (In Davidson: "Pursuit of the Concept of Truth", p. 20, in Leonardi, P. and M. Santambrogio (eds.): *On Quine*, Cambridge University Press, Cambridge, Mass., 1995.) Davidson is taking up Quine's solution to the question of when a cause or situation is counted to be the same or relevantly similar. He finds that Quine's answer is "when the patterns of sensory stimulation are the same, nearly enough, from speaker to interpreter." Davidson points out several problematic traits of this solution. "One is the difficulty in being clear how to map one person's nerve endings onto another's. A second is that people do not always share operative sense organs. Another is that variations in what one person takes to be evidence of the truth of an observation sentence may not be what another takes to be evidence. It is also a question how similarity of patterns of stimulation could serve as the basis of translation and language learning when such similarities are so evidently unavailable to learner or teacher." Ibid.

structure of first-order quantification theory (plus identity)")[29] into the unknown language, and this is being done in one instance, not by taking the constants one by one. Further, his description goes: "The evidence here is classes of sentences always held true or always held false by almost everyone almost all of the time (potential logical truths) and patterns of inference. The first step identifies predicates, singular terms, quantifiers, connectives, and identity; in theory, it settles matters of logical form. The second step concentrates on sentences with indexicals; those sentences sometimes held true and sometimes false according to discoverable changes in the world. This step in conjunction with the first limits the possibilities for interpreting individual predicates. The last step deals with the remaining sentences, those on which there is no uniform agreement, or who's held truth value does not depend systematically on changes in the environment."[30]

In this extract from Davidson's text we get a glimpse into the process of radical interpretation; we see how the interpreter starts off on his task, using the structure of his home language (which is also his own system of beliefs) to throw himself, so to speak, into the process of interpretation. At the outset the interpreter's only 'tools' are the familiar logical structure of his own language and the speaker's reactions to the common environment. If successful, the result of the process will be a theory of truth for the other language, by which we will have an interpretation of each of the sentences of the speaker's language, and in which it is displayed how the meanings of parts of the sentences contributes to the meanings of the sentences as a whole. But we cannot go directly from truth to understanding and we need to add that we are assigning truth to the sentences of the speaker when we find it plausible that he is right. By this assignment, we also get a theory of what the speaker means by his sentences, as long as we are willing to concede that knowing the truth conditions of a sentence is vital for knowing its meaning. An assignment of truth is according to this view really an assignment of a particular belief to the speaker, and so we are also ascribing a particular propositional attitude to the speaker. This is done on basis of intersubjectively accessible objects and events. Davidson's notion of meaning, however, is still neither fixed nor language-transcendent. Meanings of utterances vary with speaker, interpreter, and the time of interpretation.

[29] In Donald Davidson: "Radical Interpretation" (1973) pp 125-139 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 136.
[30] Ibid.

In the sense that the translator does not have to go into the sentences and their semantic structure, Quine's notion of stimulus meaning demands little of the translator. Neither does the translator have to go into the logical structure as concerns quantification over objects; the truth functional constants will do. But just as stimulus meaning demands little, it also gives little when it comes to an understanding of the speaker. A truth theory for the language of the speaker, on the other hand, demands more involvement from the interpreter's side, and he has to use more of his own belief system and logic in order to reach his aims. But then it also gives more in terms of understanding, and the result is a theory of the meanings and beliefs of the speaker.

4. Meaning, belief and indeterminacy

There is the process of finding out what the speaker means and there is the process of finding out what the speaker believes, and they seem to coincide. Could we separate the two processes here? On the use of the principle of charity in radical interpretation 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."[31] So it seems that the assignment of truth is basic in this regard, as "holding true" is a belief-attitude, and therefore makes it possible for us to "read off" the meanings of the speaker's sentences. We assume that the speaker is mostly right, unless it is in some way implausible, and this enables us to work on determining the meaning of the sentence.[32] But this is not to say that it is in fact possible to separate meaning from belief as two different «entities» in the method of radical interpretation. Rather, Davidson holds that "Meaning and belief play interlocking and complementary roles in the interpretation of speech."[33] This suggests that we only analytically, and not as a matter of fact, can separate the two. As an interpreter, I will not always be able to find out whether I disagree with the speaker or if he only uses words differently from me. In real interpretation, especially with speakers we know well, we can sometimes go quite a long way towards finding this out by discussing things further. In radical interpretation it is in principle impossible to get rid of the indeterminacy due to the thesis of the indeterminacy of interpretation. An interesting

[31] Ibid, p. 137.
[32] See my article 3 for more on this subject.
[33] In Donald Davidson: "Radical Interpretation" (1973) pp 125-139 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 141.

question here is, of course, whether or not this is a possibility in "ordinary" or "real" interpretation. My suggestion is that, as interpreters, we reach a point that we consider "close enough" or "sufficiently close" to a determined meaning and a determined belief of the speaker; that is, close enough to find it useless to proceed with the process of interpretation. This means that this particular point will vary from interpreter to interpreter, and that it will vary dependent on the situation as well as the speaker subject to interpretation. We could speak of this in terms of "diminishing returns", a concept from economy, which means that the utility, for the interpreter, of asking further questions to the speaker will diminish gradually, so that the value of the first questions posed, and answers given, will be of more value than the later ones. In this sense "sufficient" is neither a matter of an on-or-of-value (a "yes, it is" or a "no, it isn't), nor something that could be listed in an exact number of criteria, but a value that will vary dependent on the interpreter, the situation and the speaker.

Therefore, part of this problem relates to a general problem of indeterminacy of translation and interpretation. The point is stated in a general form in Davidson's thesis of the indeterminacy of interpretation, which states that if there is one successful interpretation, there will be alternative successful interpretations. In this scenario we have (as a minimum) two different theories of interpretation, both of which can explain the speaker's sentences in relation to the circumstances of their uttering. This has the consequence that one cannot rule out divergent interpretations, or more precisely, divergent theories. This situation is also characterized a situation of "underdeterminacy" of theory by the empirical data. The given empirical data does not then give a clue to which theory to choose. But there is also another possible scenario of indeterminacy in which I operate with one and the same theory of interpretation. In a first case within this second scenario I adjust *my theory* when I find a discrepancy, assuming that my previous beliefs were wrong, or that I have misplaced words and phrases, on account of my system. In a second case I make adjustments in *the meanings ascribed to his sentences*, on the assumption that I had previously misinterpreted his beliefs, or that I had misinterpreted his use of words or phrases, on account of his system. In yet a third case I make adjustments in *the beliefs ascribed to the speaker*, finding that he, in some sense or another, is irrational, or alternatively, is deliberately trying to deceive me, on account of either his system or my system. Either way, the indeterminacy in radical interpretation goes all the way down, and as I have suggested above, this could also be the case in "real" interpretation.[34] The indeterminacy implies that we as interpreters are in a situation where we face an equation with loads of unknown variables. We can thus never rest assured that our theory is "the correct theory of a speaker's meanings and beliefs". We can never proclaim that we, once and for all, know the contents of linguistically uttered thoughts. But the thesis of the indeterminacy of interpretation is not as radical in form in Davidson's theory as in Quine's, because of the more extensive use of our own familiar logical structure in radical interpretation, and because of the social founding of the evidential base in the form of objects and events. Davidson's version of the principle of charity is hence also applied more extensively, given more weight in the interpretational process, and is more substantial than was Quine's.[35]

5. Quantification structure

The scene that Davidson paints in his description of radical interpretation's basic elements is an account that shares many of the characteristic features of Quine's account of radical translation. In "Radical Interpretation" Davidson mentions in a footnote three features that he considers different in the two accounts: "...the semantic constraint in my method forces quantification structure on the language to be interpreted, which probably does not leave room for indeterminacy of logical form; the notion of stimulus meaning plays no role in my method, but its place is taken by reference to the objective features of the world which alters in conjunction with changes in attitudes towards the truth of sentences; the principle of charity, which Quine emphasizes only in connection with the identification of the (pure) sentential connectives, I apply across the board."

To begin with the first difference mentioned; Davidson's interpreter will apply quantification-theory with identity, and not only the logical connectives, in his interpretational procedure. Quine's reason for sticking to a more modest involvement of logical structure is that "Of what we think of as logic, the truth-functional part is the only part the recognition of which, in a foreign language, we seem to be able to pin down to

[34] The indeterminacy is stated in a weaker form in the thesis of the inscrutability of reference, which says that the assignment of reference to the constituents of a sentence may vary, while the truth conditions (and thus the truth value of the sentence) may be constant.
[35] I describe the differences more in detail in article 3 below.

behavioral criteria."[36] He says that "The categorials [or, to put it into our previously employed vocabulary; the sentences with quantifiers] depend for their truth on the objects, however external and however inferential, of which the component terms are true; and what those objects are is not uniquely determined by stimulus meanings."[37] For instance, we are unable to uniquely determine whether we are dealing with rabbits or rabbit stages, and as long as we admit to this, it seems reasonable to withdraw from the employment of quantification-theory. It should be kept in mind that Quine's premise of employing subjective stimulus meaning is the basis for the withdrawal. So, we can easily see the connection between the first and the second out of Davidson's mentioned differences; the radical interpreter *can* force quantificational structure on the speaker's language, because he also lets go of stimulus meanings. Instead of proximal stimulus meanings, the distal objects and the speaker's reaction to them are the stimuli that the interpreter builds upon.[38] This is an important move, and a move that has been vital for the development of Davidson's theory. The third difference mentioned is that the principle of charity is more extensively applied in radical interpretation than in radical translation. This last difference is, in a way, no more than a consequence of the enforcement of quantification structure on the language to be interpreted. But also this move has far reaching consequences.

It is, however, possible to approach the discussion over quantificational structure the other way around. Since objects and events are vital in Davidson's philosophy, there has to be some quantificational structure in the interpretational process. According to Davidson, objects and events in our common world are required for any interpretational process ever to start. This is also the reason why he doesn't want, or need, stimulus meaning. Surely, it is not *necessary* for Davidson to have stimulus meaning, and as long as patterns of stimulation is not what he relies on; it is not *possible* to have it. Quine is right that on purely behavioral criteria, no objects of sentences can be directly ascribed. In radical interpretation we therefore have to use much more of our own logic and linguistic "machinery" to get started compared to what was available in radical translation. For radical interpretation it is vital to operate with quantification over objects since we do ascribe objects of sentences. However, despite this possibility to turn it all the other way

[36] In Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 61.
[37] Ibid.
[38] However, in order for the radical interpreter to be able to build upon the distal objects and the speaker's reactions to them, he must be able to certify to himself that he and the speaker focus on the same object. Davidson's notion of triangulation was partly designed to account for the determination of common object.

around, my contention is that Davidson's paramount objective was to give a compositional theory of meaning. Hence, in order to make the semantic structure of sentences visible, he had to make sure that the radical interpreter was able to quantify over objects and events.[39] My claim is that *this* is how it came about that object and events became, and had to become, ontological entities in Davidson's theory.

The difference between Quine and Davidson concerning what to pick as causes for governing the translational or interpretational process, is obviously a difference of huge epistemological significance.[40] In Quine's view, we depend on patterns of sensory stimulation, while Davidson holds that we depend on the publicly available objects and events that the sentences are interpreted as being about. However, when into questions of meaning and language learning, Quine take it that we do well to deal with intersubjectively available stimulations in the form of objects. The significance of the introduction of quantificational structure, which seems to be such a neutral and philosophically insignificant move, even if vital for giving a compositional theory of meaning, is that Davidson introduced objects and events as the cause of beliefs. In radical interpretation, objects and events cause beliefs whether we see it from a semantic perspective or from an epistemological perspective.

I uphold that the driving force behind the introduction of quantification structure is the possibility for construing a compositional theory of meaning. But this interest of Davidson's was, or came, in the course of time, to be, linked to other and more paramount objectives. "Understanding" is a key word for a comprehension of Davidson's concerns. Human understanding of the world (its people included) is the "wonder" he seeks to explain, and the fact that we can grasp an objective world is due to our ability to share this world with others. Without language, sharing couldn't come about. He says that "...not only is it the case that the aim of conversation is "shared understanding"; we must also acknowledge that without sharing there is no understanding."[41]Hence, according to

[39] This is the same as saying that he had to introduce quantification in order to give a compositional theory of meaning, a theory of meaning that can display both syntactic and semantic structure.
[40] Though Davidson does not explicitly mention this difference in his "Radical Interpretation".
[41] In "Gadamer and Plato's *Philebus*", pp 421-32 in L. Hahn (ed): *The Philosophy of Hans-Georg Gadamer*, Open Court, Chicago and La Salle 1997, p. 431.

Davidson, our shared relation to the world is the source of objectivity of thought. Davidson describes the sharing of the world as the process of triangulation.[42]

This is not to say that triangulation in radical interpretation is sufficient for reaching an understanding of another human being, but it means that triangulation is a necessary, even tough "thin", condition for understanding. Without a shared world, there would be no understanding of other human beings. There could surely be contact and some form of communication. But understanding of what some other person tries to say to us, when this is presumed to be something that can be objectively true or false, is hardly thinkable without going through some third «part». This third part is some aspect of the world, something we are genetically disposed to discriminate as *one* aspect, i.e. some object or some event, or perhaps a situation. So, my contention is that for Davidson, it soon became obvious that the sharing of the world, and thereby, the intersubjectively accessible parts of the world, had to be accounted for in the theory of interpretation. Objects and events are dispositionally and socially constituted, and are in this sense already common and under public control. However, it seems to me that Davidson only later realized the epistemological importance of his own introduction of quantification in radical interpretation. Anyhow, I have not been able to find reasons for suggesting that the driving force behind this development were epistemological at the outset.

This far I have examined the major differences and similarities between Quine's radical translation and Davidson's radical interpretation over a somewhat long time-span. We started with Quine's *Word and Object* dating from 1960, and continued with Davidson's "Radical Interpretation" from 1973 and further up to the mid-eighties. Now we will take a look at the most recent developments in the relation between Quine and Davidson.

6. Quine's discomfort

Around 1994-1996 Quine changed his view on the importance of taking nerve receptors into consideration for the theory of meaning. According to Davidson, this was not due to

[42] See article 2 for a general account of triangulation.

his criticisms of Quine on this point, but because of Darwin.[43] Quine had come to realize that our capacity for categorization in relatively similar ways is a result of natural selection. We are disposed to categorize the world into objects and events as "a gift of nature".[44] According to Davidson, Quine had "wanted to explain what makes it possible for people to understand each other, given the apparent chaos of unconceptualized experience." Davidson holds that the reason why Quine had wanted to take nerve receptors into consideration was that Quine had thought "it took a lot of training for people to come to focus on the same aspects of the world, and he had thought the only way to explain this was by starting with sensory stimulations." According to Quine himself, he had always wanted to analyze reification.[45]

In "Autobiography (continued)" Quine tells about the 1986 Stanford colloquium of him, Davidson, Dreben and Føllesdal: "We agreed with the opening pages of *Word and Object* [1960] that what people share and build upon are not sensation, but the external scene itself. But in *Word and Object* I made a problem of this, and lamely conjectured an approximate intersubjective homology of neural receptors and of networks and hence of stimulations. But the conjecture was untenable, as I had recognized by 1974; [] hence my discomfort. Davidson's line was to settle for the external scene as ultimate, and this is indeed the reasonable line if one is concerned with translation or interpretation – hence semantics – and not with neurology or epistemology."[46] Here we can see that Quine maintains it is Davidson's semantic interests that led him to consider the distal stimulus basic, and that Quine for this reason accepts this line. Quine continues: "I came away cleansed of any lingering thoughts of shared stimulus meanings, but unrelieved of my discomfort over our shared reference *in distans*. Eight more years [i.e. around 1994-1995] were to pass before I saw the matter whole and clear, with its causal explanation in innate perceptual similarity and natural selection."[47]

---

[43] Davidson, Donald: "Quines epistemologier", in *Norsk filosofisk tidsskrift*, nr 1-2, 2004, translated by Eivind Balsvik from the English version; "Quine's epistemologies", p. 15.
[44] Ibid.
[45] Willard Van Orman Quine: "Reply to John Woods", in Lewis Edwin Hahn (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998, p. 727.
[46] Willard Van Orman Quine: "Autobiography (continued)", in Lewis Edwin Hahn (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998, p. 737.
[47] In Willard Van Orman Quine: *From Stimulus to Science*, Harvard University Press, Cambridge, Mass. 1995, Quine give an account of this point. "Global stimuli are private: each is a temporally ordered set of

But Davidson holds the reason for his insistence that Quine thoroughly should have taken the distal view is that only in the distal view is there "an indissoluble connection between words and what they are about – the objects and events in the world."[48] *This* is the reason why Davidson would take an epistemological interest in the distal stimuli, and this is why he had urged Quine to take the distal view not only in the theory of meaning but also in epistemology. According to Davidson's view, it is exactly when there is such an indissoluble connection that the skeptic cannot find room for maneuvering. So, from Davidson's perspective, the reason for sticking with the distal stimulus was far from restricted to his semantic interests, though to me it is pretty clear that semantic interests was his initial motivation for introducing objects and events.

In 1998, as Quine focuses on preestablished harmony yet again, this is, according to him, due to "a widespread failure to sense the problem that it solves." [49] He goes on: "Two normal observers of a scene naturally see the same thing (we say), up to differences of perspective, and cameras similarly situated agree correspondingly. Our subject matter is distal. [] We can go on from there, one supposes, skipping any talk of light rays, sensory receptors, and neural processing. If the cameras can do it, why shouldn't we?"

He continues by explaining that unlike the cameras, people have not been "designed and manufactured" alike, we vary widely in the number and arrangement of our sensory receptors and in the topology of our nervous systems. "How can our shared distant subject matter activate us so harmoniously through such disparate intervening mechanisms?" Quine asks. Further, he notes that the camera analogy confers false comfort, and that there is no parallel assurance regarding the observers. "Their verbal reports agree, but their words were likewise learned on the basis of similarly shared observations of distant events,

some one individual's receptors. Their perceptual similarity, in part innate and in part molded by experience, is private as well. Whence this coordination of behavior across the tribe? It requires that if two individuals jointly witness one scene, and subsequently jointly witness another scene, and the one witness's global stimulations on the two occasions qualify for him as perceptually similar, usually the other witness's stimulations will so qualify for the other witness. The same is required of the signal. Recurrence of it must activate, in each individual, global stimuli that are perceptually similar for that individual. So we see a preestablished harmony of perceptual similarity standards. If two scenes trigger perceptually similar global stimuli in one witness, they are apt to do likewise in another. This public harmony of private standards of perceptual similarity is accounted for by natural selection." Ibid, p. 20f.

[48] Ibid, p. 16

[49] In Quine's "Reply to Roger F. Gibson, Jr." in Lewis Edwin Hahn (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998, p. 685.

simply pushing the problem back." So Quine's point is that similarity in verbal behavior does not explain what he wants to explain. Similarity in observation can neither explain "his" problem. What is, then, the problem that is solved by the assumption of preestablished harmony? He is seeking an understanding of similarity, or the "rough intersubjective harmony of private standards of similarity". The reason is that communication has its roots in this harmony. Quine mentions language learning, in which it is important that the child achieves the required response from his mother for affirming the similarity that the child has a natural tendency to perceive. But it is the disposition, or what Quine calls the "shared structure of subjective similarity standards", that is being accounted for by natural selection. [50] Perceptual similarity standards are therefore following Quine "in part innate and are in preestablished harmony".[51]

In his "Response to George",[52] Quine talks about "the gap between the privacy of our neural intake and the publicity of our testimony" as a riddle. Or, rather, this had been a thirty-year riddle, which was now (in 1995/2000) "solved by preestablished harmony". Hence, the gap had disappeared. Subjective stimulus meanings that are private (as opposed to the earlier envisaged *shared* stimulus meanings) are presumably now, by means of preestablished harmony, in accordance with the "distal meeting of minds".

The causes of belief are still, according to Quine's epistemology, the irritations on the sensory surfaces of the speaker. This counts as evidence for the belief. On the other hand, Quine also, as we have seen, operates with "socially observable stimulations", in addition to "effects at our sensory surfaces". This means that he is also talking about objects. Objects are those to which "words apply first and foremost."[53] In semantic matters Quine is content to operate with objects, while, as an epistemologist, his interests are in stimulations at the sensory surfaces. And epistemological concerns have always been the underlying motivation for Quine's "speculations on radical translation".[54] This means that

[50] Ibid.
[51] In Willard Van Orman Quine: "I, you, and it: an epistemological triangle" in Orenstein, Alex and Petr Kotatko (Eds.): *Knowledge, Language and Logic*, Kluwer Academic Publishers, Dordrecht / Boston / London 2000, p. 2.
[52] In Willard Van Orman Quine: "Response to George" in Orenstein, Alex and Petr Kotatko (Eds.): *Knowledge, Language and Logic*, Kluwer Academic Publishers, Dordrecht / Boston / London 2000, p. 409.
[53] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 1.
[54] In Willard Van Orman Quine: "Three Indeterminacies" in R. Barrett and R. Gibson (eds.): *Perspectives on Quine*, Oxford University Press, Oxford 1990, p. 3.

the reliance on "effects on our sensory surfaces" still has a priority over objects in Quine's investigation. The reason for this is that he takes the proximal stimuli to provide evidence for theory. Should the proximal stimuli be left behind we would no longer be able to study the relation between meager input and torrential output. From a Quinean point of view, the study of this relation is what epistemology is all about, and therefore the study of this relation is of the highest interest. We should keep in mind that, in Quine's view, objects are posits, and posits can never count as evidence in a theory about the world. According to Quine: "The subject's reification of rabbits and the like is for me decidedly a part of the plot not to be passed over as part of the setting."[55]

Davidson accepts the Quinean behavior-based methodology, but, unlike Quine, he does not accept empiricism. The difference lies in the *epistemological significance* attributed to the point that all our knowledge of the world stems from our senses. If this means that our knowledge of the world apart from ourselves depends on the functioning of our senses, then Davidson would agree. If this means that our knowledge of the world can be, and has to be, *justified* by an appeal to our sensation of the world, Davidson would not agree. Our beliefs and our knowledge are propositionally structured, but, following Davidson, the world is not. Only something with propositional structure can justify beliefs. Quine's empiricism had, according to Davidson, the form of "insistence on locating the stimulus relevant to determining both meaning and knowledge at the neural input".[56] As we have seen, however, Quine changed his view about which stimuli are relevant to the determination of meaning, but not about which stimuli are relevant to determination of knowledge. In Quine's view, the neural input remained the relevant stimulus for determination of knowledge, and the reason why is that only in this way can the world function as *evidence* for theory.

But we could ask, it is a characteristically empiricist thesis to hold that the neural input is the relevant stimulus, and not characteristically empiricist to maintain that the object is the relevant stimulus? If one assumes, as Quine, that objects are products of subjectivity, one must consider the neural input (and not the object) primary, given that one wants to be an

[55] Ibid.
[56] In an unpublished manuscript of Donald Davidson: "Quine's Epistemologies" translated into Norwegian by Eivind Balsvik as "Quines epistemologier", in *Norsk filosofisk tidsskrift*, nr 1-2, pp 9-20, Universitetsforlaget 2004, p. 8.

empiricist and not an idealist. So the question of the *basis* for the truism that different people find parts of the world (objects, situations, causes, situations) similar is obviously important for such categorizations into "isms". Quine had assumed that it was the "approximate intersubjective homology of neural receptors and of networks and *hence of stimulations*" [57] (my emphasis) that could explain why people seem to have roughly similar private standards of similarity. This secured his position as an empiricist, while also explaining the similarity. But the assumption of similarity on the side of stimulations (shared stimulus meaning) was otherwise (also according to Quine), as we have seen, a very problematic solution to the problem. Now, what about his new solution of "preestablished harmony", which is so surprisingly metaphysical for a presumed empiricist and naturalist, like Quine? The idea of preestablished harmony is even an archetypically rationalist thesis. Quine's use of the term of 'preestablished harmony' could be read as a way of pointing out the thesis' lack of philosophical substance.[58] Either way, in Quine's version it means no more than that we tend to perceive similarities in a similar way across the human tribe.

In accepting the object as the relevant stimulus for determination of meaning, and also that it is the relevant stimulus for language learning, Quine suggests (probably ironically) that we can accept this solution as non-subjective due to the presence of parallelism, a preestablished harmony among human standards of perceptual similarity.[59] According to Quine's most recent expressed views, perceptual similarity is partly innate and partly in preestablished harmony, and he argues that natural selection can account for both traits.

7. Concluding remarks

Quine and Davidson agree that our biological capacity to discriminate is crucial for the possibility of linguistic communication and they agree that a shared environment is necessary for any learning of a language. But where Quine would claim that we cannot

[57] Willard Van Orman Quine: "Autobiography (continued)", in Lewis Edwin Hahn (Ed.): The Philosophy of W. V. Quine, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998, p. 737.
[58] This has been suggested to me by Jan Harald Alnes, and he assumes it to be something of an irony from Quine, and that this is Quine's way of playing a trick on the rationalist.
[59] Willard Van Orman Quine: "I, You and It: An Epistemological Triangle" in Alex Orenstein and Petr Kotatko (eds.): *Knowledge, Language and Logic*, pp 1-6, Kluwer Academic Publishers, Dordrecht 2000, p. 2.

possibly know for certain that what we react to in the external world is the same, Davidson holds that "...triangulation helps sharpen, by locating, the source of the relevant reactions, for it is a shared source." This means that, in the simplest and most basic cases, we observe an object and the other person's reactions to the object, as well as our own reactions to the object. This is, following Davidson, possible only when we can talk about what we perceive. When we are competent speakers we have already got the concept of objective truth and independent reality, and with it, an awareness of the possibility of error. Instead of Quine's "irritations of our sensory surface", Davidson hold that objects and events in the physical world are what causally affect us. Object and events are socially established entities, and they cause our beliefs; however, according to Davidson, they are still the entities they are due to human dispositions for categorization and due to human interests. Quine's method of radical translation does not presuppose triangulation and partly common establishment of objects and events. For language learning and semantic purposes he holds that we perceive similarities similarly, and the basis for perceiving the same things as objects are parallelism. For epistemological purposes this is however not satisfactory for Quine. It is *the relation* between input and output itself that is to be the object of epistemological investigation.

The dispute over what is to be the data for interpreted empirical sentences is the most prominent "purely" epistemological dispute between Quine and Davidson. Davidson's reasons for adopting Quine's method as well as his focus on the public nature of language, is that he insists on the externalist basis of meaning and other intentional phenomena. Davidson still holds that Quine's epistemological basis in surface irritations is a basis in something private, and something pre-conceptualized.[60]

On the face of it, the difference between Quine's radical translation and Davidson's radical interpretation is that Quine is doing epistemology and Davidson is doing semantics. But this philosophically neutral way of describing the difference between them also serves to cover up the significantly distinctive projects and outlooks that have developed over time in Quine and Davidson. We have seen that the reason why Quine is an empiricist, while

[60] See e.g. Donald Davidson: «Reply to Dagfinn Føllesdal» in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Davidson is not, is found in the (in terms of epistemology) innocent "detail" of operating with or without quantificational structure in the process of translation or interpretation.

Unlike Quine, Davidson is eager to meet the skeptical challenge. But the extent to which this eagerness has *governed* his different detailed moves in the development of his theory of radical interpretation is harder, if not impossible, to pinpoint. In Davidson's own description of the ways of development and change in radical interpretation, it looks as though different local problems and challenges did lead the way, and as though the question of skepticism were non-prominent. As I see the matter, the overall goal of giving a compositional semantic theory led the way. I don't think that Davidson's moves which came to narrow the space for the skeptic and decrease indeterminacy were calculated consequences. Anyhow, it came to be increasingly more important for Davidson to give an account of objective truth and objective knowledge, and the epistemological consequences of his introduction of quantification theory was a good start in this direction.

The result of Davidson's choice of the distal stimulus has paved the way for an approach that is thoroughly intersubjective, and thus social, in the sense that the distal stimulus is picked out in coordination between two speakers, in a triangular interpretational process. The distal stimulus does then, according to this analysis, lead to idealism. The social determination of the stimulus is, however, based in human nature in the form of human dispositions for discrimination and in human interests. Intersubjectivity of knowledge is therefore, following Davidson, both part of the set-up (via dispositions) and something that has to be followed up by social interaction in order to be fulfilled. Dispositions and innateness cannot alone see to it that we are intersubjectively coordinated in our linguistic dealings with the world. It is important to notice that the beliefs about the world that in this way are based upon social and perceptual externalism cannot be *reduced to* dispositions. Davidson's approach is social, and it is a non-reductive version of naturalism.

2. Davidson and the Role of Error[1]

1. Davidson's externalisms

The questions that we are supposed to approach in this seminar are: "...in what ways ... if at all, can the non-conceptual aspects of experience be said to form a basis for knowledge? What are the non-conceptual aspects of experience, and in what way are these aspects relevant to conceptual knowledge?"

Davidson doesn't say much about experience in his writings, and he says even less about non-conceptual experience. This obviously doesn't mean that he denies that we ever experience something. Neither would he deny that some of our experiences are non-conceptual, as he would certainly agree that we have sensations, like pains and itches, there are phenomena like daydreaming etc. But he would deny that the non-conceptual as such are relevant to conceptual knowledge. He will resist ways of talking about "experience" that makes it sound or function as though it is an object, or as something that lies between the world and us. Davidson's theory of knowledge is an externalist theory, which in this case means that our thoughts have intersubjectively available events and objects as their causes, and therefore he is indeed concerned with the world and our interaction with the world. But, I will hold that it is not possible to address the mentioned questions *directly*, if it is to be done from a truly Davidsonian perspective.

From a Davidsonian standpoint, our *experience* cannot play a causal role in relation to our propositional attitudes. He would say that objects and events, or the whole causal chain is what causes our attitudes. Even so, he takes perceptions to be the basis of our knowledge about the world.[2] This might sound like an outright self-contradiction, but it is really more a question of vocabulary. The *contents* of our beliefs about the world are according to Davidson directly caused by perceived objects and events in the world. But, our beliefs come to be what they are partly *also* because of where they are located in relation to other

[1] I would like to thank Donald Davidson for very helpful comments on an early draft of this paper. I am also thankful for comments and advice from Arnt Myrstad on later drafts.

[2] As do the tradition of empiricists from Locke to Quine, from which Davidson likes to dissociate himself because of their reliance on epistemic intermediaries.

beliefs in the network of beliefs. According to Davidson, beliefs are thus causally and logically related to each other, and they have the objects, properties, situations and events of the world as their causes.

For Davidson, the theory of knowledge about the world takes the form of a theory of interpretation. The interpretational situation that Davidson considers basic is a situation of "triangulation". He holds that knowledge about the world (or belief, intentionality, normativity, conceptualization, understanding; notions that in this connection are intersubstitutable for Davidson) cannot arise as a phenomenon, unless there is more than one creature present. Hence, as a condition for beliefs about the world, there has to be the world and at least two creatures.[3] The world that he is talking about is "...a world of objects, properties and events that the creatures can discriminate in perception."[4] The two individuals in the triangle switches between the roles of interpreter and interpretee, and what they have in common is the world, or rather; the objects, events and situations occurring in it. The point of introducing more than one creature, is, according to Davidson, *adding* "...something basic to the situation with one creature, for with the possibility that their actions may diverge we have introduced the gap needed to make sense of the notion of error."[5] Hence, with two persons present, there is an opening for error to occur. This connection between the second person and the notion of error will be our focus of the next part of this article.

But before we go ahead to the role of error, let's first try to exemplify the triangular situation by thinking my friend Janet and me into a specific situation of triangulation. Let's say that I play the interpreter-role, while Janet is playing the role as the one to be interpreted. Now, we said that perceived objects and events are the basis for our beliefs about the world, and that our beliefs can be identified partly on the background of where they are located in relation to other beliefs in a network of beliefs. This would mean that a certain response in the form of a belief to (a perception of) a certain event that my friend and I have in common, will be identified by me as the belief it is, partly on the basis of its

[3] I will hereafter say 'person' or 'individual' instead of 'creature', even if that really is to imply a little too much for the stage of my explanation being.
[4] In Donald Davidson: "Externalisms", unpublished manuscript, p. 4.
[5] Ibid, p. 2.

location in my friend's network of beliefs. For instance, if I am walking in the middle of a road with my friend, and she suddenly says: "car!" I will then interpret her exclamation on the background of other beliefs that I know inhibits her, e.g. the belief that cars are vehicles with four wheels and an engine, that makes them capable of running on a road, and the belief that such a vehicle with a certain velocity probably will bring damage to anything it bumps into, and the belief that a human body will be harmed by meeting a car with a certain velocity, etc., etc., and I will of course take into account the actual car that we both perceive. On the basis of all of these assumptions, I interpret her as telling me that a car is approaching us, and we will probably both hurry ourselves onto the side of the road. In this way, I interpret her as trying to warn me about the car. In a situation like this, there is a possibility of a misunderstanding between us, maybe I'm not realizing that what my friend said was meant to be a warning, or I might think that she said something else because of all the noise around us, or I might not realize that the vehicle coming toward us is what she would call a 'car', as I would have called it a 'truck' etc., etc. Ascription of a belief to the other person is thus being done both on the basis of the perceived event; the vehicle that comes driving down the road, which causes her belief, *and* that particular belief's location in her network of beliefs. In the basic triangle there are causal connections between the world and the two persons respective, and there are causal connections between the persons. We perceive the relevant aspects of the world, and we perceive each other's reactions to the world. Davidson holds that "these two sets of connections are themselves correlated", for each person associates for instance cars with car-reactions of the other person, "...except of course in the relatively rare case of error."[6]

As we can see, Davidson would insist that whatever justification we could have for attributing propositional attitudes – and other mental content – to others, must be public and intersubjectively available in the outset. But this point seems to be based on recognition of methodological restrictions rather than ontological or metaphysical commitments. This might make a difference in the present context. What we can consider in radical interpretation is what we all can observe, but *this doesn't mean that what we can hear or see is all there is,* in Davidson's opinion.

---

[6] Ibid, p. 4.

But how do we know whether even a simple utterance, such as the utterance of 'car' in that situation really is a correct way of using the language? Maybe it would be wrong to classify the vehicle as a car, and so the right thing to say would be 'truck'? Someone might, even more radically, insist that saying 'bar' is the only correct solution, or even 'bruck'.[7] And, on what basis could we say that they were wrong? According to Davidson, we have to go into the situations where we learned the language in question (here; English) for the first time, or rather; when we heard the sound of someone uttering the sounds necessary for pronouncing 'car' (or 'truck' or whatever), for the first time applying it to a specific object in the world. The most common way of teaching somebody to apply a specific sound to specific objects for the first time, is by pointing out the specific objects which the sound are meant to refer to. Ostensive learning is therefore of utmost importance in language acquisition, according to Davidson.[8] Further, he would hold that when we take this normal ostensive process of learning a language into account, we have an answer to the question of how we can know whether we have misapplied the concept of 'car' or not in specific cases. My belief that this particular object is a car has, first of all, the correctness guaranteed by the process of which I learned to apply that word in the first cases. Furthermore, its correctness is verified by the confirmations (and rejections) that others have given to my usage of the word, and by my observations of how other adult language users have applied that particular word. Through our interpretations of the other speakers, and other speakers' interpretations of us, we have the possibility of being corrected by others, and we have the possibility that our beliefs are being refined by our interaction with others. Thus, the publicly available objects and events cause our beliefs in the first place, and the correction of wrong beliefs is going on all the time. The corrections comes by way of further perceptual experiences, by the fact that our beliefs are being confronted with other and contradictory beliefs that we have ourselves, and by contradictory beliefs that others have. This makes it generally not very likely that most of our beliefs about the world are wrong.

---

[7] Anyone familiar with Kripke's *Wittgenstein on Rules and Private Language* will recognize this as the voice of Kripke's meaning skepticist. (Saul Kripke: *Wittgenstein on Rules and Private Language*, Oxford University Press, Oxford 1982.)

[8] He holds that Kripke partly misses this point in his book, because of Kripke's use of only mathematical examples. (p. 2 in Davidson: "Externalisms", unpublished manuscript.) Mathematical functions like addition isn't typically what you learn when you first learn a language.

But why do we pick out the same things to be objects and events that other creatures do in the first place? How come that a car would count as an object in our (and most other peoples') language? According to Davidson we are disposed – *genetically* – to discriminate, but we also *learn* to discriminate more fine grained by and by.[9] Without our learned discriminative powers we're just not in a position to be thinkers and believers. But we always have to think about it all as *starting from* common discrimination that we are genetically inclined to, combined with the learned discrimination or categorization that has its starting point where an experienced language user ostensively teaches us how to hook words on to the world. In the beginning, it is likely that the teacher will have to switch between ascribing propositional attitudes on a behavioral basis and overinterpreting the child, putting too much into it.[10] This will go back and forth, and in the end, when the learner has developed a full-fledged language, there will no longer be a need for overinterpreting.

For Davidson, perception is propositional and it is an open question whether non-propositional seeing is possible for an adult language-user. (x sees *that* y, versus x sees y) He would not acknowledge a separate level of concepts, as he says that: "...there is in fact no distinction between having a concept and having thoughts with propositional content, since one cannot have the concept of mama unless one can believe someone is (or is not) mama..." Further, he says that he will "...stress the connection between concepts and thoughts only to make the point that concept-formation is not a way station between mere dispositions, no matter how complex or learned, and judgment."[11] This means in effect that children's (and, of course, other people's) one-word sentences should be seen as propositions, rather then as concepts, if these first understandable utterances are to be seen as more than mere reactions.

What is then necessary for distinguishing between, on the one hand, a pure disposition to react to stimuli in a certain way, and on the other hand the ability to conceptualize and

---

[9] Bob Brandom is instructively calling "being genetically disposed to discriminate in" "reliable differential responsive dispositions". (p. 900, footnote in his "Knowledge and the Social Articulation of the space of Reasons" in *Philosophy and Phenomenological Research*, Vol. LV, No. 4, 1995)
[10] Comment made by Davidson in seminar, Berkeley spring 2000.
[11] In Donald Davidson: "Seeing through language" (1997) pp. 15-27 in *Philosophy: the journal of the Royal Institute of Philosophy*, Cambridge University Press, Cambridge 1997, p. 25.

think (what Wittgenstein is talking about as "following a rule")? Two creatures can respond in a certain way to certain stimuli and we can then say that there is a certain focus of their responses; there is a common cause to be found. This common cause will be an indication of the content of the thought that's present.[12] So one creature will see relevant similarity between its own response and the other creature's response, and so this second person's perspective that is present in the triangulation seems to be the *source* of objectivity. Davidson says that this is acknowledging that "...the only legitimate source of objectivity is intersubjectivity."[13]

Is there really anything more to add here? Have we now said what *is* to be said about Davidson and "non-conceptual aspects of experience and their relevance to conceptual knowledge "? Surely not, but by this first approach I have tried to give a general picture and a quick overview of his position regarding the relevance of non-conceptual experience to conceptual knowledge. What I have said implies that *the non-conceptual* is hardly relevant to conceptual knowledge at all in a Davidsonian theory of knowledge. *Experience*, understood as something that lies in-between the world and our beliefs, is simply non-existent in his theory. It is rather that the objects and events in the world *cause* our beliefs (and other attitudes) directly, and objects and events thereby also take part in the *content-formation* of our beliefs. Thus, the world, in the form of objects, properties, situations and events "impinges" *directly* on our belief-systems. According to Davidson, there is nothing such as the world as "given in our experience".

2. Making Sense of Error

I will now go on to concentrate on the notion of error and the role that error or mistake, and a subject's acknowledgment of the possibility of error and mistake, plays in Davidson's theory of knowledge. I will try to show that this is deeply connected to his rejection of skepticism about the world. Awareness of the possibility of error is necessary as a criterion for ascription of propositional attitudes to others, and according to Davidson, it is also a

[12] In Donald Davidson: "Externalisms", unpublished manuscript, p. 9.
[13] Ibid.

sufficient condition. I will briefly discuss whether disagreement or the possibility of alternative interpretations is another criterion for such ascription to take into account.

The concept of error is connected to the concept of truth, and to the fact that truth and belief are distinct. Truth is the way in which things are, while belief is the way in which we take things to be, and if there were no difference, there were no error in our sense. Hence, we could say that error can only emerge where there are norms or where there are rules for correctness. Mistake is then a sign of a normative, holistic system that inhibits a person. When a mistake is possible, and the person knows that it is possible, *we* know that a person might be a language user. The concept of truth is therefore central to our use of language, and for our thinking in general, according to Davidson. But this will only be so at a certain "level" of language use; it will not be so for a creature that only copies what it has "learned" by others. Thus, for truth to be important, the person will have to have reached a stage where he or she realizes that he or she could be wrong about something. The person must be aware of the possibility of a mistake for the concept of truth to have an application. Furthermore, Davidson holds that "...where error has no point, there is not a concept of thought." It seems as though thought and truth have the same conditions in some respects; they both require the possibility of error. The same holds for beliefs (or the concept of belief); there is no use in characterizing something as a belief unless one in principle could be wrong.

When a two-year-old child says "light" (or rather; something that his parents or guardians would interpret as such), pointing to the lamp, he doesn't really understand what he is saying unless he understands that this sound can be successfully applied to only a certain limited selection of objects in the room. Hence, if he has no idea that this sound does not go as well with his brother and the milk bottle as with the lamp, he does not understand the sentence he is uttering.

This seems to me to be a reasonable account of how we as adult language users usually would understand such a situation. What we actually would do if he uttered "light" when pointing at his brother, would be to correct him, saying maybe "No, this is *Eric*", pointing at his brother, and then: "*This* is light", pointing at the lamp or the candle or whatever.

Now, it might take some time to get it, that an error is possible, but when the child does get it, he might indicate this by saying for instance: "light" in a questioning voice, looking very carefully at his parents' or guardians' faces for reactions. At this point, we may begin to apply a concept of belief to him, saying that he *believes* that the lamp, or whatever he might be pointing at, *is* "light". For us, as full-fledged users of language, it is also very clear that it would make sense to tell whether it is true or false that what is pointed at is "light", but this could also be possible for us at an even earlier stage. But as far as *my* intuitions are concerned, I would neither say that a two-year old, nor that a younger child yet understands the concept of truth and falsity, or what *we* think of as truth and falsity. If a child says "light" when pointing to his brother, I think that we would rather say that the child had not yet quite learned how to apply the concepts and expressions of the language. We would say *this*, rather than holding that he was saying something that was false, or not true. When the child is in a position to play with words confidently, or "deceive" others with his words; that is, when it is obvious that the child is in charge of the word-to-world connections and tries them out in a way so as to, for instance, deliberately provide a misapplication, then we can start questioning whether he knows the difference between truth and falsity. Later on, he would be able to provide reasonable disagreements with his parents or guardians, and he would be able to argue about which descriptions of a common environment are the better ones.

An important point here is that it is not possible to make sense of error in the beginning of language-acquisition or before the individuals in question have internalized a norm. When there is only dispositions to react to something perceived that is present, we wouldn't even call deviations by the name of 'error'. (By disposition I mean something like a stabile tendency to respond in a certain way to certain stimuli, something that has no normative force.) Error can in this context only make sense once one has mastered the language, and error is therefore a sign of intentionality. But how can we identify a mistake as a mistake? How can we really distinguish between having a concept and having a disposition?

This is the same as asking: How do *you* know that *they* know that they can be wrong? Expression of uncertainty or if they ask you? But do they really think, or do they have propositional content, unless they are able to produce a disagreement? I am not sure that awareness of mistake is the sole prerequisite for us to acknowledge others as thinkers/language-users/believers etc.; we might also require that they could produce interpretations that can compete with our interpretations of the world. In other words, we might require that they are able to disagree with us. At least, it is obvious that the situation of triangulation requires this kind of symmetrical relationship between the persons present. Without "propositional communication", we wouldn't be able to share our beliefs about the world with the other(s). Or, as Davidson puts it himself: "...without the exchange of propositional contents, there is no way they can take advantage of their ability to triangulate their shared world." This means that, even if the creatures cannot give their own interpretations (which by definition is something recognized as fallible), there could still be a basis for ascription of thoughts to them, according to him, but they wouldn't be able to share their thoughts with us. So, Davidson holds that error is sufficient for ascription of thought, but it is unclear to me how this view of his combines with his description of the basic situation of triangulation.

Either way, according to Davidson, we still lack an account of how error can get into the picture, or what he calls "an analysis of the concept of error". He holds that if we had an analysis of the concept of error, we would have an analysis of the concept of objectivity or of thought itself, but the very fact that we haven't managed to give such an analysis, is a reflection of the fact that "...intentional phenomena cannot be reduced to something simpler or different."[14] Thus, it seems that Davidson is pessimistic when it comes to the possibilities of explaining intentionality.

I would agree that the triangulation-metaphor seems to lack a reference to the systematic or systematizing effect that are added by holistic considerations. There is the world and the other person, but where is the normative system? The pressure from within, so to speak? The pressure from within our own belief system and system of concepts? Does the other person represent that? The other person instantiates society, but is he also representing the normative holistic from within the system of beliefs? Or is it maybe the I-person – the interpreter that represents the normative system? The pressure from holistic considerations, or what we would consider normativity, is there in a way, as something immanent in our

[14] Ibid, p. 11.

interpretation of the other person, in the pressure from what we have already interpreted to be his other beliefs. In this way, it seems as though normativity itself *emerges* in the triangular situation. But Davidson is explicit in stating that triangulation is only one step towards explaining the difference between reaction and thought. He does not think that triangulation is a way of explaining thought, but rather that triangulation provides illuminating necessary conditions for thought.[15] The metaphor of triangulation is therefore, according to him, lying out the necessary basics for thought and language to be present, but it doesn't give an explanation of thought or for the difference between reaction and thought. The triangle where the three participants are the other person, the world and I, cannot explain *normativity* either. Holistic considerations are not brought in, and intentionality is not explained. On the other hand, it seems to me that it is impossible to explain the triangular situation without *presupposing* intention. Neither is error, and awareness of error, possible without presupposing intention. This may point to an extended role for the metaphor of triangulation.

3. McDowell's criticism of Davidson in *Mind and World*

If my presentation of Davidson's theory of knowledge is reasonable, it seems as though McDowell is systematically misrepresenting Davidson's views on the relation between thought and the world in his *Mind and World.* McDowell says that "Davidson's picture depicts our empirical thinking as engaged in with no rational constraint, but only causal influence, from outside."[16] That this comes out as a misrepresentation of Davidson's view is due to the fact that McDowell omits the theory of interpretation, which deals with precisely "rational constraints". These rational constraints on empirical thinking could be said to come from "the inside" in a way, since they come from the other beliefs in the system. But, and this is essential: the constraints apply especially and exactly to our "outside", to the world that we are dealing with. And maybe the most important point to stress here is that the rational constraints on our empirical thinking are always there as constraints coming from "the others". Thus, in triangulation the other person is interpreting my utterances on the background of what is going on in our common and publicly available "environment". This means that his or her interpretation of me, my utterances and my

[15] This does not imply that the conditions are eternal, even if he looks upon them as necessary. These conditions are conditions that are necessary *now*, that is, given our present biological constitution.

belief-system are contributing to the process of determining the contents of my beliefs. I am able to occupy the same role, the role of the interpreter towards the other person, thereby taking part in the process of determining the contents of his or her belief. This means that our empirical thinking is engaged with rational constraints, the constraints that the first ostensive learning process supplied, and then later the constraints from "society".

McDowell could seem be right in some ways here; Davidson's coherentism may seem to threaten "...to disconnect thought from reality."[17] But then it has to be shown that his insistence of the world's causal role towards our beliefs are impossible, or without effects or implications. If so, the only contact between thought and world is through truth, or through interpretation of others, and his theory would indeed be "one legged" in a sense. The reason for saying that it is *not* is that the basis of interpretation of others (in radical interpretation at least) is the utterances that they produce as reactions to a publicly available world. The causal impact of the world is already put into "the space of reasons", in Davidson's system, but this is done through the eyes of the interpreter. This rational relation between word and world is one that he has acknowledged for at least the past 30 years. He holds that our interpretation of the empirical beliefs of others is constrained by what *we* hold to be rationally sound beliefs. So, it seems that Davidson in no way is denying that our interaction with the world is epistemologically significant, quite on the contrary. So, McDowell's characteristic of Davidson, that "...Davidson's picture is that we cannot get outside our beliefs." is somewhat strange. In a way, we could say that we have never been anything but "outside" our beliefs because our beliefs have their source in what we perceive, and our thinking is intrinsically social.

If we grant McDowell the benefit of doubt as to whether he misrepresents Davidson or not, it is important to notice that Davidson's theory of interpretation has been developed and changed over the years. It has developed from its source in the situation of the Quinean field linguist doing radical translation, to radical interpretation, and over into the situation of triangulation. I would say that this has been a development from a more static behaviorist model of stimuli and response, towards a dialectic process of mutual

[16] John McDowell: *Mind and World*, Harvard University Press, Cambridge, Mass. 1994, p. 14.
[17] Ibid, p. 24.

interpretation between "speaker-interpreters" in the model of triangulation. This means that McDowell's comments may have been a bit more reasonable in 1994 than they are now.

McDowell says: "We find ourselves always already engaging with the world in conceptual activity within such a dynamic system. Any understanding of this condition that it makes sense to hope for must be from within the system. It cannot be a matter of picturing the system's adjustments to the world from sideways on: that is, with the system circumscribed within a boundary, and the world outside it. That is exactly the shape our picture must not take."[18]

I think that McDowell touches an important point here. But I also believe that this point can be accommodated by the notion of triangulation. In the situation of triangulation, the world is a fully participating, always and already present part. But it is also true that the world is only making impacts on us (and our "concepts") causally, not rationally. If we hold that concepts are bound up with the world *rationally*, our notion of a "world" would be a world of mere appearance, or in other words, a mere phenomenal world, and not a world that could actively cause anything (such as beliefs) in us. This could open up for the possibility of a system of beliefs completely out of touch with the world. If we wish to explain how objectivity of empirical thought is possible, the solution that Davidson has pointed out seems to be a better one: That we "catch" the causally active objects and events into a rational system of beliefs by interpreting others. Thus, we connect world and thought by ascribing beliefs to others, and others interpret us by ascribing propositional attitudes to us, and all of this is done based on publicly available objects and events. These interpretations take the form of hypotheses (guesses about what the other sayings mean), which in turn are being corrected or confirmed and refined as we communicate with one another, and as time passes. This makes what McDowell calls "a sideways-on understanding of our thinking" unnecessary.[19] We mutually come to have an understanding of the beliefs that we have, and we have most of them in common.

The perhaps most pressing problem with McDowell's solution is that he relies on a representational picture of the world-to-word relationship. His solution seems to require

[18] Ibid, p. 34.
[19] Ibid, p. 35.

that for a belief to be about the world, we have to be able to pick out the part of the world that justifies this particular belief. This would, in turn, require a manageable correspondence between parts of the language and parts of the world. Having witnessed many attempts at such theories for the past hundred years of analytic philosophy, it is hard to see how it could be achieved in a satisfactory manner.

4. Concluding remarks

What are the strengths and weaknesses of Davidson's theory as we have seen it unfold? As I see it, the main achievement of the Davidsonian theory lies in the non-representational, yet objectivity-oriented epistemology. His theory of knowledge gives an account of what we already know; namely, that our beliefs about the world are objective. Specifically, our beliefs are objective in the sense of intersubjectivity. He avoids skepticism by rejecting epistemic intermediaries, and also by taking the objects of a belief to be the causes of that belief.[20] As regards the non-representational features of his theory, he holds that there are no such things as representations of reality in our heads. He says that "...we ought to question the popular assumption that sentences, or their spoken tokens, or sentence-like entities or configurations in our brains, can properly be called 'representations', since there is nothing for them to represent."[21] Here, he is contesting correspondence theories of truth, specifically, the ontology of facts that would have to follow such a theory. The following excerpt is a good summary: "Language does not mirror or represent reality, any more than our senses present us with no more than appearances".[22] This is where I would say Davidson might be parting from Rorty's *less* objectivity-oriented, but *non*-representational approach and from McDowell's objectivity-oriented, but representational approach.

[20] At page 317 in his "A Coherence Theory of Truth and Knowledge", he puts it this way: "What stands in the way of global skepticism of the senses is, in my view, the fact that we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief" in Donald Davidson: "A Coherence Theory of Truth and Knowledge" (1983) pp 307-319 in E. LePore (ed): *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford 1986.

[21] Donald Davidson's manuscript "Epistemology and Truth" (1988), p. 12. (The article has later been published in Donald Davidson: *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001, pp 177-191, and the quote is on p. 184.)

[22] Donald Davidson: "Seeing through language" (1997) pp. 15-27 in *Philosophy: the Journal of the Royal Institute of Philosophy*, Cambridge University Press, Cambridge 1997, p. 18.

The main weakness of Davidson's theory regarding the problems addressed by this workshop may be that his claim that perception plays a casual role in the word-to-world relationship is not very specific. What does the statement really mean, and what does it amount to? He would probably say that it means that whatever we perceive is the cause of our empirical beliefs. We perceive the world and we perceive the reactions of the others to the world. This, combined with our considerations for the system of beliefs as a whole, forms our beliefs. But it doesn't seem that we can go much further into the question of what it amounts to, to say that objects and events cause our beliefs. Davidson doesn't appear concerned with explaining *how* our empirical knowledge can be *empirical*, or *how* our beliefs *can* have «content». Such questions can probably not be answered by his theory if something more than a story about language-acquisition and perception is asked for.[23]

Another problem that might be raised is that his theory is not very phenomenologically sensitive. Only propositional differences are taken into account. He says: "Language is the organ of propositional perception. Seeing sights and hearing sounds does not require thought with propositional content; perceiving how things are does, and this ability develops along with language."[24] Hence, Davidson is *not* letting the non-conceptual play any role when it comes to conceptual or propositional knowledge, and he agrees with McDowell by holding that "...reasons have to be geared conceptually to what they are reasons for. The relation of epistemic support requires that both relata have propositional content, and entities like sensation and sense-data have no propositional content."[25] In effect, this means that we cannot possess knowledge, properly speaking, of the non-conceptual. At least, this is the conclusion that Davidson seems to draw from the insight that only something propositional can be a reason for something propositional. As most of us will probably have noticed, McDowell draws another conclusion from the same insight, namely that the world itself is propositional. According to McDowell, the only way of

[23] Given that we acknowledge his criticism of the distinction between conceptual scheme and content, the reason why he cannot or would not even try to answer such questions, should stand out as obvious. (In Davidson "On the Very Idea of a Conceptual Scheme" (1974) pp 183-198 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.) Another alternative would be to account for the ideal conditions under which specific beliefs are caused, and he isn't exactly doing that, but the most ordinary situations, with the most ordinary objects that are middle-sized, *do* play a particular role.
[24] Donald Davidson: "Seeing through language" (1997) pp. 15-27 in *Philosophy: the journal of the Royal Institute of Philosophy*, Cambridge University Press, Cambridge 1997, p. 22
[25] Ibid.

securing the objectivity of our knowledge of the outer world is to consider the *justifications* for holding our beliefs as *propositional*. As I have argued, Davidson is in no need of this sort of contra-intuitive commitment. Given the way that we have all learned to connect the perceived world to the word in the first place, and given the way that a specific perceptual belief fits with the rest of our beliefs, we have the possible and required foundation and justification for the objectivity of our beliefs. In the question of objective knowledge McDowell has failed to appreciate the importance of the linguistic community.

Finally, we could ask: "What is the function of Davidson's move to allow us to conceptualize the world through other speakers; that he lets the world's importunity on us go through another person?" The whole point of this is to stress the fact that thought and language is both social phenomena. Following Davidson, there is no language without a community, and he goes further in this direction than many others by saying that thought does not exist without a community. Literally, society is a condition of both thought and language. This makes for an anti-individualist epistemology, and this is where Davidson anchors his rejection of skepticism about the world and skepticism about other minds. It also enables Davidson to tell the skeptic to "get lost", and it seems to me that McDowell is unable to do anything of the sort, due to his individualistically oriented epistemology. Bob Brandom's comments on McDowell shows that he, in line with Davidson, is serious about this Wittgensteinian legacy. Brandom phrases this in post-Sellarsian language: "...I have indicated how knowledge can be construed as standing in the space of reasons."[26] Further: "To individualize the space of reason is to interiorize it. To ignore the social articulation of standing in the space of reasons is to leave out what makes it possible to understand such standings as answerable for their correctness to how things actually are."[27] Thus, normativity and the objectivity of our norms are connected to the world itself in triangulation.

Davidson's two externalisms are his perceptual *and* his social externalism, and according to himself, the interconnected triangle "...fills the gap in social and perceptual

[26] In Robert B. Brandom: "Knowledge and the Social Articulation of the space of Reasons" in *Philosophy and Phenomenological Research*, Vol. LV, No. 4, 1995, p. 907.
[27] Ibid.

externalism..."[28] According to Davidson, triangulation does not necessarily explain objectivity of thought, but he grants that some necessary and interesting conditions for objectivity of thought are laid down. The possibility of error is our indication of the existence of such objectivity.

[28] In Donald Davidson: "Externalisms", unpublished manuscript, p. 5.

3. The Status of the Principle of Charity

The principle of charity has always occupied a key role in Davidson's theory of interpretation. Roughly, the principle states that in order to understand another being, one has to assume that the other being is in accordance with himself, and that he speaks the truth. In this way, the principle describes a condition for understanding other linguistic beings. Davidson's theory is not yet a theorization of actual everyday communication. His theory operates on an assumed underlying level, and the theory takes a thought experiment as its vantage point. The thought experiment is named "radical interpretation", and describes an interpretational situation in which an interpreter is trying to understand another being with an unknown culture and an unknown language, that is, unknown to the interpreter. Radical interpretation involves a basic position, in which no linguistic behavior is ascribed a specific meaning in advance. The key to understanding of the other being is to be found in a connection between truth and meaning. An assumption is that the interpreter is able to recognize the speaker's attitude of "holding true", and the interpreter reaches the meaning of the utterances on the basis of the truth conditions of the utterances. It is initially in this imagined situation that the principle plays its role as a possibility condition for understanding. In this case, the principle is a requirement for successful communication and even for the very possibility of getting communication started. Davidson anticipates that the principle is inevitable.

In this article I discuss ascription of epistemic status to the principle of charity.[1] I give an account of Davidson's version of the principle, partly by comparing it to Quine's version, and, in particular, I stress the function that the principle of charity had in the first place in Davidson's theory.

As radical interpretation is an imagined case, and in one sense, an idealized situation, I will analyze it further to get a hold of the possibilities and restrictions that the situation prescribes to the principle. I will involve Quine and Davidson's understanding of classical epistemic distinctions, and consider how this understanding influences ascription of epistemic status to the principle. I argue that our present epistemic categories are unfit for

[1] The epistemological turn that Davidson's philosophy has taken with the introduction of the concept of triangulation has actualized the question of the epistemic status of the principle.

the principle of charity; yet, at the same time, that an understanding of the function of the principle is epistemically enlightening nonetheless. Moreover, I show that the epistemic potential of the principle in Davidson's theory is elucidated by the introduction of the concept of triangulation.

1. Davidson's principle of charity

The basis for Davidson's principle is to be found in Quine's *Word and Object*, where Quine introduced the principle as a rule of "radical translation". Quine's principle states that obviously false and absurd utterances from the person you are trying to understand are more likely to be due to poor translation than that he or she actually holds the view that your translation attributes to him. He says: "The maxim of translation ... is that assertions startlingly false on the face of them are likely to turn on hidden differences of language. ...The common sense behind the maxim is that one's interlocutor's silliness, beyond a certain point, is less likely than bad translation—or, in the domestic case, linguistic divergence."[2] If we hear the other person say something that seems founded on a very unreasonable belief, we must, according to Quine, assume that there could be something in the other person's language that we have misunderstood. In a related note he continues: "Cf. Wilson's principle of charity: "We select as designatum that individual which will make the largest possible number of...statements true" (). [3] According to Wilson, the point is to choose as starting point the individual [the translation] that makes the largest possible number of utterances true[4], and we must assume that Quine counts this as a point in favor of his own characteristic of the principle (that utterances which appear untrue are likely to originate in our deficient understanding of the language).[5] If we are to believe that our informant truly means what he says, according to our manual, this may compromise many established translations and even make obsolete many of the reasonable beliefs we had

[2] In Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 59.
[3] Ibid.
[4] Wilson's example anticipates that we have five possible "versions" or alternatives to deal with, and when four out of these alternatives points towards the same "individual", while one alternative points towards another individual, we will assume that the individual pointed to by the four concurrent versions is the right one. In Wilson's example we thus deal with a predominantly quantitative assessment. See page 532 in N. L. Wilson: "Substances without Substrata", in *The Review of Metaphysics*, Vol. 12, nr 4, 1959, p 521-539.
[5] We can add that when we run into absurd utterances of the informant, these utterances can be absurd in at the least two different ways; they are absurd because they are so in the light of our own beliefs, or they are absurd on the basis of the beliefs we until now have ascribed the informant, which in its turn *also* is based upon our system of beliefs.

previously considered true beliefs of the informant. Therefore, we will not pick the translation of the utterance that makes a large number of the other utterances false. Instead, we will assume that we have bumped into a previously hidden difference between the language of the informant and our own language. Ten pages on in the book, Quine approaches another methodic point by saying: "...bold departure, [is] to be adopted only if its avoidance would seem to call for much more complicated analytical hypotheses. For certainly, the more absurd or exotic the beliefs imputed to a people, the more suspicious we are entitled to be of the translations; the myth of the prelogical people marks only the extreme. For translation theory, banal messages are the breath of life." [6]

Having simple analytical hypotheses is something to reach for, and ascription of ordinary, or even banal, beliefs to our informant, is according to Quine "the breath of life" for translation theory. Simplicity is Quine's only mentioned reason for putting aside the principle of charity. If we would have had to operate with a much more complicated hypothesis, the cost of employing the principle of charity could turn out too high.

Summarized, the function of the principle in Quine's theory is to refuse a translation that ascribes exotic beliefs to the informant. This means that the assumed meaning of an utterance is being weighed against the belief that must be ascribed the speaker, when we take this particular meaning to be a correct translation. This is Davidson's starting point for bringing the principle into the theory of radical interpretation. In radical translation, or, to be precise, radical interpretation, Davidson starts off with the operation of weighing meaning and belief. In order to be able to understand someone else, one has to understand which beliefs he or she has, as well as what his or her utterances mean. In "Radical Interpretation" Davidson says that it all turns on "...holding beliefs constant as far as possible while solving for meaning"[7], and the principle of charity has the function of holding on to the belief.

The problem of finding a way of weighing belief and meaning against each other is a reminder of the problem Davidson had been involved in earlier when trying as he tried to find a way of separating "utility" and "subjective probability" in experimentally oriented

[6] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 69, my addition.
[7] "Radical Interpretation" (1973), in Donald Davidson: *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 137.

decision theory.[8] The problem in question was how to find an empirical basis for identifying and separating the two entities, and how a theory could contribute in the separate measuring of subjective probability and expected utility. When an individual chooses between alternatives involving uncertain outcomes, at least two factors arise in the decision: The relative desirability of the possible outcomes, and the extent to which the different outcomes are considered probable. Generally, one would assume that a rational individual chooses, among high-risk alternatives, the alternative that maximizes expected utility.[9]

Davidson's inspiration came from the strategy that Frank Ramsey had chosen when trying to figure out the concept of subjective probability.[10] Ramsey had noticed a problem with subjective probability, or what we might refer to as "the degree of belief in beliefs", the problem that others cannot observe it directly. Unquantified preference can, however, be read of in behavior in concrete situations of choice, which can be partly observed. Behavior in situations of choice gives empirical evidence, and, based on this, Ramsey conceives that a theory of subjective probability can be given empirical content. He imagines setting up a sequence of options, which, when we have the theory, may give the information about the degree of belief in beliefs and the relative strength of desired preferences. 'The degree of belief in beliefs' and 'the relative strength of preferences' will thus be theoretical concepts that will have their meaning through their role in explaining and predicting decisions in situations of choice.

Having this in mind, Davidson construes a utility function based on a random event with a subjective probability of one half, which can be utilized to define the utility of the outcomes, and this utility function can subsequently be used in the measurement of

---

[8] First published as Donald Davidson and Patrick Suppes: "A Finitistic Axiomatization of Subjective Probability and Utility" in *Econometrica*, Vol. 24, 1956, p. 264-75. "Utility" or "expected utility" will here involve the question of to what extent one will value that a certain incident will occur (corresponds to "desire"/wish), while "subjective probability" will involve the question of how likely one estimates it to be that this incident will occur (corresponds to "belief"/view). Davidson has often called attention to this resemblance between theory of translation/interpretation and decision theory.

[9] See p. 1-3 in Donald Davidson, Patrick Suppes and Sidney Siegel: *Decision Making, an Experimental Approach*, Stanford University Press, Stanford 1957.

[10] See p. 84 in Davidson's "Reply to Jerry Fodor and Ernest LePore" in Stoecker, R. (ed.): *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers*, De Gruyter, Berlin 1993, pp 77-84. Davidson did however find the utility function later to be mentioned on his own, and only later did he discover that Ramsey had found the same function. For an example, see p. 32 of his "Intellectual Autobiography" in Lewis Edwin Hahn (Red.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

subjective probability. This means that by fixing the subjective probability at one half, one can calculate the relative utility of the different outcomes, and the function thereby construed can again be employed to measure subjective probability. Hence, it was now possible to "determine a person's subjective probabilities and relative degrees of preference simultaneously, assuming neither in advance."[11]

The similarity between theory of interpretation and decision theory is this difficult situation where we are facing two factors that are both obviously important in a situation where we have to decide what to choose, but where none of the two are more basic than the other, and where the value of both factors mutually influence each other. Thus, it is necessary to weigh up the two factors, and it is all about finding the right place (or reference mark) on the lever by using one of the factors to "measure" the other. Hence, one of the factors is given a specific value; in decision theory the utility function is given a subjective probability of one half, thereby deciding the utility, and in theory of interpretation the beliefs of the other person is given according to my beliefs and according to the other person's other beliefs. This fictitiously specific reference mark (a specific utility value/a specific belief) is thereafter employed to, respectively, measure subjective probability and to determine the meaning of an utterance. By inserting a specific hypothetical value for one of the factors, a specific value for the other factor appears.[12]

But what is really achieved by this move? In the case of radical interpretation, in which Davidson has sought a way to separate meaning and belief, the objective has been to gain

[11] Ibid.

[12] Davidson made it clear that he brought this solution along from decision theory to radical interpretation, and the solution came to be the beginning of a unified theory on language and action. Davidson originally took as starting point the assumption that the interpreter must be able to recognize when the speaker holds something true. This is being described in "Radical Interpretation" from 1973, but already in "Belief and the Basis of Meaning" (see pages 147-48) from 1974 (both in Donald Davidson: *Inquiries into Truth and Interpretation,* Oxford University Press, Oxford 1984) he notices that it would have been better if we could employ a speaker's preferences between sentences as a starting point for interpretation. That the speaker holds something true ("holding true") were in an expanded theory substituted by that the speaker prefers the truth ("preferring true") of one sentence to the truth of another sentence. See also "Towards a Unified Theory of Meaning and Action" in *Grazer Philosophische Studien*, 11, 1980. This substitution lead to the unification of theory of meaning and decision theory (or, more generally, theory of action) and thus a theory of interpretation, where all forms of human intentional action can be handled. A comprehensive account of this unified theory is given in "The Structure and Content of Truth" from 1990, in *The Journal of Philosophy*, Vol 87, p. 279-328. The extended theory is also radical, since it has the same kind of ignorance constraint, and the principle of charity is being used unlimited. In this article I nevertheless maintain the focus on linguistic behavior and the role of the principle of charity in this context, as the unified theory, as far as I can see, does not bring about a changed epistemic status for the principle of charity.

the ability to explain what is involved when an utterance means what it does, or in other words, what is the significance of saying that the utterance has a meaning. Davidson says that he initially wanted an answer to the question of what meaning is, and that the radical interpreter thought experiment was motivated by the question of what linguistic meaning is, and what it can be when we do *not* base it on a metaphysically heavy loaded concept of meaning.[13] But, according to Davidson, the question of what meaning is cannot be given an answer; we don't even know where to begin, or what would count as an answer. Instead, Davidson poses the following question: "What would it suffice an interpreter to know in order to understand the speaker of an alien language, and how could he come to know it?"[14] We are dealing with the unknown speaker and his physical environment. We don't know anything about the speaker's character; we don't know the culture in which he takes part, and we have no preconception of the meanings of words and expressions in his language. The radical interpreter enters the situation with an aim of getting to grips with the beliefs of the speaker, and the hypotheses that he forms will not be filled with a specific understanding of the meanings of the speaker's utterances. In radical interpretation, we want evidence that will *suffice* for the interpreter to be able to form a substantiated theory about the beliefs and meanings of a speaker. The basis for this is that the interpreter sometimes knows when the speaker affirms a given sentence, even if the interpreter doesn't understand the sentence. This is a decisive premise for the possibility of radical interpretation, where interpretation as the possibility of recognizing that somebody holds something true enables our entry into the mutual dependence between belief and meaning. The beliefs of the speaker, which the interpreter partly copies from his own system of beliefs, function as the key to finding the meanings of the speaker's utterances. Partly, the aim is to acquire a theory of meaning, a theory about the meanings of a certain speaker's sentences in a limited period of time. The principle of charity, which, in this case, performed the function of enabling the interpreter to hold onto (what he sees as) the speaker's reasonable beliefs, enabled the interpreter to reach the meanings in a legitimate way.[15] By assuming the rationality of the speaker when an alternative interpretation would

[13] In "Reply to Jerry Fodor and Ernest LePore", p. 83 in Stoecker, R. (ed.): *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers*, Berlin 1993 we find Davidson saying: "Like many others, I wanted answers to such questions as "What is meaning?", and became frustrated by the fatuity of the attempts at answers I found...."

[14] Ibid, p. 83.

[15] Not knowing the speaker and his language is a condition that is supposed to "prevent smuggling in into the foundations of the theory concepts too closely allied to the concept of meaning". (From the introduction to his *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. xiii) The legitimacy of

involve ascription of obviously false assertions to him, (all of course according to the interpreters own standards of rationality and truth) the interpreter has found the meaning of the speaker's utterances. This means that the use of the principle of charity involves ascription of belief and meaning based on unusually low evidential coverage. We cannot know for sure whether or not the speaker is, in fact, telling the truth, or whether or not he is in accordance with himself, and therefore, we cannot be certain that the meaning that we ascribe to the utterances is the meaning of the utterances at the time. By further re-examining the theory, we can obviously test the results at a later time, and thereby increase evidential support. But we have brought forward an initial hypothesis about the meaning of the utterance, which is basis for proceeding.[16]

Which role and status should now be attributed to the fictitious specific reference mark? It is important to stress that the beliefs that we, hypothetically, have been holding on to, are precisely temporary hypotheses about the speaker's beliefs; they are a means to reach the meanings of the utterances. For this reason, the result must be temporary as well. Suggestively, Davidson describes this as "passing theory"[17]. Yet again, through our estimation of them as such by our usage of the principle of charity, the beliefs are neither true nor consistent. We have utilized the principle in order to choose the interpretation that we consider most reasonable, and by doing this, we assume that we have found the best interpretation. There is nothing in our employment of the principle that *give* beliefs a permanent or metaphysic status of truth, and this applies to both speakers' and interpreters' beliefs.

For Davidson, as pointed out earlier, the principle of charity was initially some sort of lever for "measuring" or uncovering the meanings of the speaker's utterances in radical interpretation. In interpretation, employment of the principle results in a theory of meaning and a theory of belief in a single move. In Quine's version, the principle functions in a

---

the theory of meaning is secured by not presupposing a specific concept of meaning, and by letting the radical interpretation rest heavily on empirically occurring interaction between speaker, surroundings and an, as far as possible, unbiased interpreter.

[16] The truth conditions for the utterances of the speaker and that the speaker is holding sentences true, is yet not the sole support for the interpreter. Davidson is pointing out that the interpreter also must consider "...simplicity, hunches about the effects of social conditioning, and of course our common-sense, or scientific, knowledge of explicable error." ("On the Very Idea of a Conceptual Scheme" (1974) p. 196 in Davidson: *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.)

[17] See p. 442 in "A Nice Derangement of Epitaphs" in E. LePore: *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford 1986.

similar way as the rule of aspiration to simplicity, and it does not have a special status in that respect. Having this in mind, it is easy to see that the principle of charity has another role to play in Davidson's theory. The plain naturalist will not regard the principle any differently from Quine, and it is therefore important to take a look at how Davidson's use of the principle is different from Quine's. To Davidson, the principle is necessary for accessing the system of beliefs of the speaker, in order to make out what he means by what he says, and for finding out which beliefs he has. In the very beginning of radical interpretation, the principle cannot be put aside to the benefit of considerations of simplicity. The principle enables us to pick a first hypothesis and to choose between different alternatives of interpretation. In Davidson's version, the principle is also constitutive in the sense that it contributes to giving meaning to concepts such as belief, meaning, wish, etc. In this sense, it is internally constitutive in the theory of interpretation. This is so because we cannot make out what the meaning of the speaker's utterance or his beliefs are unless we can utilize the principle, and because talking about meaning and belief is meaningful only when we interpret a speaker in relation to his surroundings. Quine's version of the principle was never constitutive in character, and it is reasonably clear that Davidson commits himself more in his usage of the principle than did Quine. In Quine's theory, the principle is a methodological guide that, level with the demand for simplicity, regulates the process of translation.

Furthermore, it is important to notice that Davidson's focus on truth in the process of interpretation has an effect on the function of the principle of charity. For Davidson the principle is to ensure truth and consistency in a first approach to radical interpretation since, without a critical mass of shared belief, understanding cannot be attained. In Quine's theory, the principle did of course not have this function; one reason being that Quine's radical translation lacked an objective of obtained understanding.

2. Radical Interpretation: "Ideal" or "Normal"?

How are we then to understand Davidson's concept of radical interpretation? As mentioned, it is a thought experiment not intended to theorize on actual interpretation.[18] Is

[18] Quine is talking about it as a thought experiment on p. 28 in his *Pursuit of Truth*, (Harvard University Press, Cambridge, Mass. 1990) while Davidson speaks about radical interpretation as a "conceptual exercise"

it hence meant to be a description of the ideal interpretational situation? Davidson has explicitly denied that the radical interpreter is an ideal interpreter.[19] According to Davidson, the theory of radical interpretation and the radical interpretation situation is not meant to be "ideal" but "normal" (representing a normal rather than an ideal situation). Radical interpretation has recently been referred to as "ideal interpretation"; the radical interpreter called an "idealized interpreter" and the speaker said to possess the greatest extent of "cognitive and moral perfection".[20] It is clear that this might make a difference for the epistemic status of the principle of charity. The concept of ideality is not leading us in any specific direction, and, as opposed to the concept of normality, doesn't anchor the principle in (human) nature. Ideality of communication lead our thoughts towards transcendental pragmatics, and it can lead us into thinking Davidson's project into such a frame. The concept of "normal" can lead us in the direction of Darwin, in the direction of human dispositions, and into thinking that humans are linguistic beings in their biological constitution. Facing these two alternatives, it undoubtedly makes more sense to associate a tradition of evolutionary theory, or naturalism, to Davidson's philosophy.

The concept of ideality can, however, be understood in two different ways. It can be perceived as "prefigurative", hence as an exemplary, or it could on the other hand be understood as "conventionalized", "simplified", as a model or a typical form. Davidson has disavowed the former. However, he does agree that radical interpretation is a theoretical situation, and a fictitious situation. The situation can be likened to a scientific experiment, or a model, which is simplified in the sense that potentially interfering elements are left out. The situation is cleansed, made transparent and construed for a specific purpose, and is methodologically ideal. Furthermore, the situation cannot be an actual historic occasion. But the radical interpreter's set of beliefs is neither more ideal than are yours and mine, nor is the speaker made ideal or perfect by the interpreter. Some commentators have

---

(page 324-325 in his "The Structure and Content of Truth" in *The Journal of Philosophy*, 87 (6), 1990, pp 279-328) which was not "meant to throw any direct light on how in real life we come to understand each other, nor on how we master our first concepts and our first language."

[19] Personal comment, Berkeley, May 2000.

[20] See page 1 and page 10-16 in Bjørn Ramberg's "Naturalizing Idealizations: Pragmatism and the Interpretivist Strategy" (at http://folk.uio.no/bjoerntr/Natideal.html). Ramberg is making it clear that he means to follow Davidson in his explanation of "the interpretative strategy" as he is laying it out in the article, but that he is not claiming to do Davidson-exegesis. He is putting up something like an interpretational machine, IDA, which is very similar to the radical interpreter, but he points out that he is aware of that this is not Davidson's radical interpreter. I therefore neither argue against Ramberg's interpretation as an interpretation of Davidson.

apprehended Davidson's principle of charity as a principle of speaker idealization. There can be a considerable difference between saying that a first hypothesis is that which is constituted by use of the principle of charity, as I have emphasized, and saying that "what people actually believe and mean is constituted by what [] it is ideally rational for them to believe and mean." as Simon Evnine[21] explains Davidson. I will maintain that this is not about the "ideally rational", but about what a normal, fallible interpreter deems "normal".

A preliminary hypothesis about meaning and belief is what is constituted by the application of the principle of charity in the first place. Evnine's phrasing in the mentioned quote ("what people *actually* believe and mean") might lead us to think that other people/other creatures ontologically or metaphysically constitute the beliefs and meanings occurring in speakers. Either way, this is not the case. The perspective of the radical interpreter is methodological and epistemological in the sense that it points out the possibilities and limitations for ascription of meaning from a third-person perspective. This third-person perspective, which is a fictitious case, determines the possible theories of meaning and belief for a speaker, or maybe even the best possible theory of meaning and interpretation. But an interpreter cannot alone *determine* what a speaker means and believes; this is settled in a larger context, where the speaker's general natural dispositions and his causal and social learning history form part of the picture.

Moreover, Davidson has emphasized that both the interpreter and the speaker are fallible subjects that make mistakes, are wrong and also sometimes contradict themselves, even if mistake and inconsistency cannot be the main rule, because a consistent set of analytical hypotheses could not then be generated. The radical interpretation situation must therefore be intended to consist of a normal interpreter and a normal speaker, where the limits are set by the way in which natural language works and by the psychology of normal, linguistically equipped humans (for instance, normal humans are fallible and not omniscient), but it is a situation in which our overview is better and we are more in control of what is going on then we are in actual interpretation. We make use of the situation to find out and clarify the requirements of understanding another speaker. We want to illuminate the underlying presuppositions for catching someone else's beliefs and the meaning of his utterances.

[21] Simon Evnine: *Donald Davidson*, Stanford University Press, Stanford 1991, p. 178.

Hence, Davidson's radical interpreter cannot be an ideal interpreter, even if his competence is a model. Just as little is the speaker in radical interpretation an ideal speaker (Likewise, the speaker in radical interpretation is just as less an ideal speaker), and just as little (less) is the principle an ideal that interpreters of all kinds should strive for.[22] If we assume that radical interpretation is theoretical, but has a normal interpreter and speaker, we have to assume that the principle of charity applies to normal interpreters and speakers, even if it does not necessarily apply in the same form in actual interpretational situations. The principle is not designed for ideal interpreters and ideal speakers, but is involved in a theoretic model. If, on the other hand, the radical interpretation thought experiment in its entirety is perceived as an ideal in the prefigurative sense, it will not necessarily apply to normal interpreters and speakers. This interpretation would undermine the scope of the principle and would weaken the value of transmitting it to real-life speech and thought.

Davidson has gradually turned radical interpretation from a one-way interpretational situation, in which an interpreter is trying to understand a speaker based on the speaker's utterances as related to the common physical environment, into a situation in which mutual interpretation exists between the two speakers, though still related to their joint surroundings. In triangulation, the new thing is that the speaker and the interpreter have the same status. Both parties to triangulation are speakers as well as interpreters. Moreover, using the triangulation metaphor, Davidson emphasizes that our beliefs are partly caused by casual influence exerted upon us by a common world; in other words, through casual effects exerted on us by empirically occurring objects and events. Furthermore, ostension and the special role of beliefs related to the genesis of sets of belief have been accentuated. We have seen a shift from focusing on internal constraints in interpretation to, instead, emphasizing causality. I will return to the effects of this change as regards the principle of charity.

[22] Davidson is utmost clear that the principle is not an ideal that we can choose to follow if we wish; it is according to him not possible to choose not to follow the principle. See e.g. page 221 in his "Mental Events" (1970) in Davidson: *Essays on Actions and Events*, Oxford University Press, Oxford 1980, where he says "Crediting people with a large degree of consistency cannot be counted mere charity: it is unavoidable if we are to be in a position to accuse them meaningfully of error and some degree irrationality."

3. Apriori? Open to empirical revision? Transcendental?

Is the principle of charity empirical or apriori? The Kantian constitutive apriori seems to be a possible apriori category for the principle. In this case, the apriori is what makes experience possible; in other words, it is the condition for experience without being dependent on experience. The Kantian apriori involves that there are elements of knowledge independent of experience yet closely connected to experience. It is characterized by being necessary and non-revisable. Kant's conception is that the apriori involves knowledge, and that which is knowledge must be true, and therefore non-revisable. Can the principle of charity have an apriori status in this sense? The principle functions as a condition for interpretation and is in principle conceived into the special situation that constitutes radical interpretation. The validity of the principle is presupposed in radical interpretation, and without this presupposition no understanding would come to exist. Trough the use of the principle we can understand the other person. Hence, the principle is actively involved in forming understanding in radical interpretation; according to Davidson the principle is inevitable for such formation. Consequently, it must be referred to as constitutive pursuant to normal usage, and it is constitutive in a thought experiment. Davidson indicated as early as in 1970 that the principle of charity is constitutive, without going deeper into the details than we have as yet.[23] In "Mental Events", Davidson likens the constitutive with synthetic apriori. In the article he also says that ascription of consistency and coherence is inevitable.

But could the principle really be apriori in a Kantian sense? The thought experiment precedes experience in its entirety, and so the question does not comport well with the case, at least if we consider it epistemically committing. If we were to say that the principle is apriori, it would be natural to think that there must be something in the situation that is not apriori. Is there anything non-apriori in the thought experiment of radical interpretation? Not everything in internally in radical interpretation is in a flux. How normal (not ideal) people function psychologically, what characterizes natural languages – these are empirical presuppositions on which radical interpretation rests. We cannot avoid these assumptions if the thought experiment is to honor the value of transmittance. The thought experiment is therefore not devoid of presuppositions; it depends on a series of

[23] In "Mental Events" (1970), p. 221, and p. 236-237 in "Philosophy as Psychology" (1974), both printed in Davidson: *Essays on Actions and Events*, Oxford University Press, Oxford 1980.

assumptions that we would generally consider anthropological. The participants in radical interpretation are not machines, and they are not ideal, but normal. The events and actions that take place are still not "normal"; it is a fictitious situation, a model that is operated in order to explore and examine certain philosophical ideas.

If we still were to ask whether the principle is apriori in a Kantian sense, we could ascertain that it does make understanding possible, and therefore, it is a condition for the kind of experience we seek when interpreting others. But does the principle itself depend on experience? Could we have been aware of the existence of these conditions if we had never had interpretational experience? It is hard to imagine that we, at first, could have established the thought experiment if we had no interpretational experience. Is then the principle necessary and non-revisable? The principle constitutes a part of a theory of interpretation of human beings. Obviously, as any other theory, this theory could be made obsolete by a new or more accurate theory. But, given this particular theory, we cannot imagine the theory without the principle since the principle takes part in giving meaning to central concepts, such as belief, meaning and truth. Thus, the principle seems to be theoretically necessary, within the theory. But is it practically necessary? Can what we say about the principle be transferred to practical interpretation [and to our general reflection on thought, language and knowledge], or is it possible that our perception of the principle's necessity is due to our interpretational practice? We are at loss for a good answer.[24]

In modern usage of the concept of apriori, the elements of necessity and impossibility of revision is not always closely tied to the concept. In a collection of articles by Paul Boghossian and Christopher Peacocke: *New Essays on the A Priori,* published in 2000, the introduction reads: "Being a priori is to be sharply distinguished from being necessary, from being true purely in virtue of meaning, and from being knowable infallibly."[25] And continues: "For all that, it may still seem that a priori propositions cannot be defeated by

[24] An indication which gives reason to talk about the principle as necessary is given by Davidson in his "On the Very Idea of a Conceptual Scheme" (1974, printed as pp 183-198 in his *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984) where he says that in order for something to be counted as a language or a conceptual a scheme, it must be interpretable. Analogical to this one might say that for somebody to be counted a speaker, it must be possible to apply the principle of charity to him or her. This implies that the principle is constitutive of being a linguistic being. This is another form of necessity than the necessity involved when we say that the principle is required for getting started on an interpretation of another being.

[25] Paul Boghossian and Christopher Peacocke (Red.): *New Essays on the A Priori*, Oxford University Press, Oxford 2000, p. 3.

wholly empirical information; that is, that they may still be experientially indefeasible. It may be natural to wonder: if something is empirically defeasible, how can it be known justificationally independently of experience?"[26] It still seems that the concept of aprioricity is characterized by that the validity of an apriori assertion can be established independently of experience. This seems to be "the smallest common denominator" in a modern usage of the concept. But, according to Boghossian and Peacocke, the apriori also seems immune to plain empirical invalidation. The interesting question is whether or not we consider the apriori empirically revisable. If the principle of charity is non-revisable, this seems to be in accordance with the view that we can know its truth validity independently of experience, but if it *is* revisable, how can we then know its validity independently of experience? In other words, it won't be easy to uphold a concept of the apriori if we say that what is apriori is also revisable. Still, claiming that we can have knowledge of something independently of experience, and yet, that this something is subject to empirically conditioned revision is not incongruent. The authors talk about this form of aprioricity as "non-conclusive a priori justification or entitlement".[27] This means that the epistemic basis is not considered sufficient for an adequate or "clean" apriori status, but they maintain that these cases should still be referred to as apriori. This concept of aprioricity is consequently substantially different from Kant's notion of the apriori, as the authors focus on justification, and as they do not assume that aprioricity involves truth.

Boghossian and Peacocke maintain that much work remains before we can fully understand these cases and before we know how they should be evaluated in relation to non-empirically-revisable cases. For our purpose, the relevant question is whether or not the principle of charity can be a case of "non-conclusive a priori entitlement". In the sense that we don't have to wait for an empirical confirmation in order to determine the principle's entitlement, we can know that the principle is valid, and that it must work independently of experience. But is the principle immune to plain empirical invalidation or not? Can we imagine a situation where our experiences of interpretation and communication could show that the principle must be revised on empirical basis? If the answer were yes, we would be able to ascribe the status of "non-conclusive" or incompletely apriori to the principle. A second option would be to call the principle quasi-

[26] Ibid, p. 4.
[27] Ibid, p. 4f.

apriori, but that would be a less precise term.) Should we decide to categorize the principle in this way, we would really only be saying that it is a principle that we can validate independently of experience, and that it is a principle that can be revised on empirical basis. The first part of the characteristic is straightforward, but as yet, we haven't discussed the question of empirically conditioned revision.

In radical interpretation, as mentioned earlier, something obviously depends upon empirical affairs, and this also concerns the principle of charity. First of all, in radical interpretation there is the premise that the language spoken by the speaker is a natural language, involving certain empirical opportunities and restrictions. We can imagine that if we were wrong about the properties of natural language, we could also be wrong about our assumptions about the behavior of the principle of charity. Moreover, when it comes to the anticipated normal psychological inventories of speakers and interpreters, there are connections between radical interpretation and the empirical. If our conception of how human beings work is psychologically incorrect and it could be shown that we are generally unable to recognize that somebody holds some sentence true, this could facilitate a revision of the principle. If interpreters in general were unable to recognize assent to utterances, it would have been impossible for an interpreter, as the principle assumes, to keep the beliefs of the speaker constant. Or, what if we as interpreters are unable to distinguish between grades of belief in our own belief, and in this way cannot distinguish between more or less central beliefs in our own system of beliefs, and therefore neither can estimate when the beliefs of the speaker are absurd? Or, what if we are merely lacking grades of belief in our own beliefs, to such an extent that we consider all of our own beliefs to be true? This last state is the supposed condition for autistic persons.[28] I don't think that these very broad assumptions we maintain concerning the properties of known natural languages and the normal psychological inventory of people could turn out to be to such an extent that empirically based revision of the principle of charity would enforce itself. Neither is it obvious what would be the consequence if anything for the principle if our suppositions about such empirical matters were exposed.

[28] Regarding the principle of charity and autistic people, Peter Pagin and Kathrin Glüer argues that people with autism still does not represent empirical counter-evidence. See pp. 44-46 of their "Meaning Theory and Autistic Speakers" in *Mind and Language*, Vol.18 No. 1 February 2003, p. 23-51.

In asking whether the principle can be revised on an empirical basis, it could be that we really meant to ask whether the beliefs we ascribe to the speaker by help of the principle really by help of the principle are by and large true and consistent. We must then assume that there is a way in which we can independently discover the beliefs of other people, and then check whether or not this way is in accordance with the result obtained through employment of the principle of charity. But what does the principle really state? It states that in order to interpret others in a situation of radical interpretation, we must *assume* that speakers more or less tell the truth, and that they are by and large in accordance with themselves. It does not state that most of our own and the other person's beliefs *are* true and consistent.[29] Moreover, it is almost impossible to imagine that an independent check could be initiated without employment of the principle.

It is generally a problem to distinguish between the empirical and the apriori in an immanent naturalism of the Davidsonian kind. There could be several reasons for this. Davidson mentions that, in connection with attempts at testing the empirical validity of decision theory, he ran into a variant of this problem: "I became discouraged with attempts to test the empirical validity of decision theory because I could not think of a well-grounded way of separating the empirical and a priori demands of rationality. Too many apparent deviations from what rationality required were explicable as in accord with theory when reinterpreted."[30] This seems to be a steadily repeating experience, that when we relate to the intentional behavior of humans, we relate to variables that are almost impossible to empirically fix in unambiguous ways.[31] The picture that we have of rationality beforehand and our empirical interaction with intentional behavior mix up in such a way that it is hard, and maybe impossible, to know where to draw fundamental

[29] Still, Davidson says that the fact that interpretation is possible is due to "the fact that we can dismiss a priori the chance of massive error." ("Thought and Talk" (1975), printed as pp 155-70 in Davidson: *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 168-9.) We can then, according to Davidson, apriori rule out the possibility of speaker's thorough mistake, and this means that most of the speaker's beliefs must be true. But, it is here not made clear whether this is intended to be a methodological, ontological or metaphysical claim. I have taken the claim to be purely methodological, as a consequence of the method of interpretation from the third person perspective, and my interpretation of Davidson is here that he does not maintain that most of the beliefs therefore actually *are* true. Or rather, I would say that the veridical nature of beliefs couldn't be based on the principle of charity alone.

[30] In Davidson: "Reply to J.J.C. Smart" in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, p. 125.

[31] This is of course built into Davidson's philosophy in a general form via his endorsement of Quine's thesis of "the indeterminacy of translation". See e.g. p. 227 in "The Inscrutability of Reference" (1979), in Davidson: *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984. The indeterminacy is increasing the further away from "observation sentences" we get. (Davidson is not employing this term.)

lines. In practical interpretation, we handle this in a straightforward manner, and for theoretical ends we can avail ourselves of a pragmatically adapted use of a division between the apriori and the empirical. My point is, however, that the distinction between the apriori and the empirical cannot carry a substantial epistemic weight in a holistic theory where such differences are regarded gradual.

If we ask ourselves whether the principle could be a transcendental condition for understanding, we must ask if the principle could appear as a result of a transcendental deduction; or in other words, if it could have its justification and legitimacy via such a deduction. Provided that one finds transcendental deductions to be possible operations, one must then ask whether the principle can be brought forward by such means. I would say that even if the principle is a *condition* for our understanding, this doesn't make the principle a *justified* principle. If a transcendental deduction is characterized by justifying the subject for deduction, the principle probably cannot be deduced transcendentally. Hence, it is possible that the principle is a quasi-transcendental condition for understanding, since it on one hand appears to be a possibility condition, but on the other hand that this still isn't enough to justify the principle? As long as we cannot give necessary and sufficient conditions for something to be characterized as quasi-transcendental, such a categorization isn't much of help. A third alternative is to say that the principle is transcendental in a weak sense, owing to the fact that the "transcendental condition" concept is to be understood in the sense of a "possibility condition". But in that case, we would be just as well off by simply employing the term "possibility condition". I cannot see any obvious reasons for saying that the principle of charity is a transcendental principle, nor will we have much use of the term "transcendental" in a philosophical mindset where it doesn't make sense to utilize the concepts of "apriori" and "constitutive" in an epistemic binding fashion.

4. The possible epistemic status of thought experiments.

What status could a principle in a thought experiment such as radical interpretation have? Could a principle in a thought experiment actually be revised? Radical interpretation is not purely theoretical; it is a thought experiment that relies on a long list of empirical assumptions. One could hardly claim to be in control of the individual factors and external conditions for this type of experiment. Empirical experiments are often used with an aim to control or manipulate the external (or boundary-) conditions, even though we cannot even in this case be certain that all prerequisites are known for us (for instance in the form of theoretical supporting hypotheses). In radical interpretation our overview of premises is hardly better than in experiments in general. Therefore, I will assume that we lack control of the premises and presumptions on which this thought experiment relies, and that the principles we postulate as conditions will therefore not have a different status than other, more or less empirically based premises. Furthermore, to think that we in thought experiments can draw fundamental lines between empirical revision and revision on basis of change in theory, as Robert Sinclair, for instance, seems to assume, can hardly fall into place in a Davidsonian frame.[32] By and large, Davidson endorses Quine's holism, and adheres to the Duhem-Quine-thesis. A main point here is that individual assertions do not have their own stock of privileged observational sentences that can uniquely verify or falsify statements within a theory.[33] According to Quine, a theory will have a shared basis of empirical and theoretical assumptions, where the single parts of the theory cannot meet "the tribunal of experience" alone, as separated from the rest of the theory.[34] Davidson operates readily with a distinction between empirical and apriori, but only on an utterly pragmatic basis.[35] The radical interpretation thought experiment is theoretical, but it relies on the properties of natural languages and on an understanding of the workings of normal, psychologically equipped people. Still, the point is that any revision of a theory will be marked by holism in such a way that a distinction between theoretic and empiric revision, according to a basic Davidsonian view, cannot be attributed a decisive epistemic importance. This indicates that, founded on the assumption that the principle is "empirically non-revisable"; it will be difficult in any case to ascribe a particular epistemic status to the principle of charity within Davidson's philosophy.

---

[32] In Robert Sinclair: "What is Radical Interpretation?" pp 161-84 in *Inquiry*, 45, 2002, p. 177.

[33] Or, in Davidson's formulation: "...the problem of error cannot be met sentence by sentence, even at the simplest level." Page 152 in "A Coherence Theory of Truth and Knowledge" (1983), in Davidson: *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

[34] Quine is giving an account of this for the first time in "Two Dogmas of Empiricism", p. 42-46 in *From a Logical Point of View*, Harvard University Press, Cambridge, Mass. 1953.

[35] See for instance Davidson's comment to Peter Pagin in "Comments on Karlovy Vary Papers", page 293 and 294 in Petr. Kotatko, Peter Pagin and Gabriel Segal (eds.): *Interpreting Davidson*, CSLI Publications, Stanford 2001.

5. Conclusion

What is, strictly speaking, the objective of the principle of charity? Within radical interpretation, the original aim was to produce an empirically based and legitimated theory of meaning, a theory of the meanings of the sentences of speaker A by the language L, within the time T. This took the form of a truth theory; a theory in which T-sentences are analytical hypotheses, the relation between the uttered sentences and the surrounding world situation is a relation of satisfaction when it is fulfilled.[36] Thus, the theory of meaning consists of preliminary analytical hypotheses that are constantly revised as we gradually reach greater insight and understanding. In connection to this, the objective is to use the principle of charity to hold on to speakers' beliefs, while we on this basis try to figure out the denotation, or meaning, of her sentences. The point of the principle is that, by supposing that the speaker is right and is in accordance with herself, we can make out what the speaker means by what she says. Still, it is important to remember that radical interpretation describes an imagined case of interpretation, and is meant to clarify the conditions of achieving understanding in a situation where the person and his language are originally unknown. The conditions of understanding that we find through the thought experiment are assumed valuable for an inquiry into the general possibility conditions for successful communication and understanding in natural language. This generalization, through the transference to natural language, makes the status of the principle an interesting question. If the principle describes a general possibility condition for communication, and, in the form of exchange of prepositional content, communication is a condition for knowledge, then a clarification of the epistemic status of the principle is obviously important.

If we were to say that the principle must be either transcendental or naturalistic, we would also be saying that there is a principal difference between the formal and the empirical or between that which is as it is because of theory and that which is as it is because of empirical factors. Or, we could say that some things are true, "come what may", and that some other things can be revised on empirical grounds, or that something is "scheme" and constant and something else is "content" and varies with the scheme. It is evident that it would be a mouthful to argue that the principle of charity, as it appears in Davidson's

[36] We imagine for instance that Maria say "it's snowing". "It's snowing" is true if and only if it snows in the vicinity of Maria when she utters it.

thinking (or in Quine's for that matter), is a transcendental or purely naturalistic principle; and this will particularly be so for those who maintain the first standpoint, but also for those who by "naturalistic" means "utterly empirically based". Davidson is not a naturalist in this way; he is not of the opinion that rationality and normativity can be unpacked as fundamentally physical or empirical. Thus, the Davidsonian version of the principle cannot be clarified as naturalized in this sense. If the ambition of this article was to assign the principle of charity epistemic status, where the alternatives were naturalistic or transcendental, or perhaps apriori but quasi-empirical, it is plain to see that the goal has not been reached. We could not find a reason to include the concept of "non-conclusive apriori entitlement". On the other hand, if we sought to shed some light on the epistemic status of the principle of charity, we may have succeeded.

I have argued that using the principle of charity (as it appears in radical interpretation) is no guarantee for the truth of the beliefs of the speaker or the interpreter. This is evidently more so if using the early formulation of the principle. Using the principle of charity in radical interpretation cannot guarantee that our beliefs are mostly true, because that which is established (or constituted) are preliminary hypotheses about meaning and belief. Still, these provisional hypotheses must be internally coherent; we cannot deal with a self-contradictory theory. This means that the principle to a very limited extent has a constitutive function for beliefs, as larger evidential resources will be available for later corrections and more advanced versions of the interpretational theory. Thus, our dependence on the principle is gradual; the principle is more imperative the weaker our general knowledge, and less necessary the more we know in advance. It is always needed in order to secure the best interpretation possible, but to a graded extent. We must therefore be able to consider the principle more or less necessary as well as more or less constitutive.

By the introduction of triangulation, the causal element is emphasized; the focus is set on that our ascription of belief and meaning to others, and other people's ascription of belief and meaning to us are in causal interaction with each other and are also both in causal interaction with objects and events in our common physical surroundings. The causal history of our beliefs, (as based on interaction with the world and with other language users), is partly determining the content of our beliefs. This means that the principle of charity in the form of coherence is not the only factor supporting the reasonableness of our theory about the beliefs and meanings of others. The principle is obviously not the single actor seeing to an intersubjective agreement between interpreter and speaker. It is not until the entry of the concept of triangulation that it becomes clear to Davidson that the principle of charity can be split into two principles, the principle of coherence and the principle of correspondence.[37] The focus on triangulation has accentuated the correspondence side of the principle. This doesn't really make a difference related to the epistemic status of the principle, as Davidson has always assumed that the radical interpreter interprets based on the jointly available objects and events, but it does clarify and underline why it is a misunderstanding to think that coherence is Davidson's sole epistemological rationale.[38]

[37] In "Three Varieties of Knowledge" (1991), in Davidson: *Subjective, Intersubjective, Objective,* Oxford University Press, Oxford 2001, p. 211.

[38] I thank Jan Harald Alnes for discussions, suggestions and constructive criticism throughout the work on this article. My thanks go also to Mikael Janvid who has read and given useful comments, and to Peter Pagin who has pointed out mistakes and misunderstandings.

4. The Explanation of Error

Davidson holds that if we had an analysis of the concept of error, we would have an analysis of the concept of objectivity or of thought itself. According to him, the very fact that we haven't managed to give such an analysis is a reflection of the fact that "...intentional phenomena cannot be reduced to something simpler or different." It might seem as though Davidson is pessimistic when it comes to the possibilities of analyzing and explaining intentionality, and thereby, the possibility of explaining error. All the same, he has repeatedly stated that he is searching for an explanation of error. This apparent pessimism combined with repeated acknowledgment of the lacking explanation of intentionality and error poses a challenge that cannot be disregarded.

In this article I go through Davidson's endeavor to analyze and explain error, and thereby thought, normativity, objectivity; in short, all that is considered characteristic for intentional creatures. In short, his attempt is to point to the biological constitution of man in addition to his own account of triangulation. This means that he proposes a two-tier explanation of the intentional, in which the essential conditions are human dispositions and language learning in a society of language users. Radical interpretation is the frame of the attempt, but this is obviously not the case for the reference to the biological dispositions of man. I argue that Davidson's hardships with the problem of explaining error stems from a discrepancy between his individual notion of language as idiolect, a notion that is intrinsic to radical interpretation, and his appeal to a society of language users. I maintain that a community based, or social, notion of language is needed. The reason for this is that when communication is smooth, an interpreter has no opportunity to check whether both the speaker and the interpreter are wrong. Unless this could be done, a notion of objective truth could not be constituted trough the process of interpretation in triangulation. The concept of objective truth is partly based on the possibility of performing a social check, but to the extent that the model of radical interpretation precludes this, the theory cannot solve the problem that it has generated. My analysis is based on a pass through of the situation that the radical interpreter finds himself in when he is dealing with the two different possible errors that he could confront.

1. An internal or an external perspective on intentionality?

Let us begin with the obvious, that we are competent intentional and norm-possessing speakers, who are in fact able to describe the world in ways that enables us to successfully interact with it. We do predicate, we do *say* something meaningful with our sentences, and we are able to communicate with other minds. Now, we could ask, why are we able to do so? We can take an internal perspective on intentionality and ask: "What are the conditions for thought, intentionality and for the concept of objective truth that we already possess?" We do have the concept of objective truth, and we therewith also have the concept of error. Hence, we're saying: "look, we *do* have propositional thought, intentionality *is* an existing phenomenon" and because of this, there is also a concept of truth and a concept of error. Having stated the obvious, we can still always ask some further questions about the background of intentionality. "How did we acquire this trait?" "What has been the condition under which intentionality has emerged?" "What, if anything, proceeds propositional thought?" This means we are attempting to explain the phenomenon starting at the inside moving outwards, and I therefore consider it an internal approach. An internal perspective on intentionality could certainly also consist of a conceptual analysis, where an account of the relations between key intentional concepts is given.[1]

On the other hand, from the point of view of external conditions, we would say that we know that *the non-propositional* exists, and from this we have to explain how, and possibly also if, the propositional could have emerged. We would then start on the outside of intentionality, or at least, we would say that the approach is an attempt at starting with the outside moving towards the inside. In this approach, the very existence of intentionality as a phenomenon in its own right is not considered obvious. In any case, in this approach we will have to explain how the non-intentional could *become* intentional, or how the intentional could "grow out of" the non-intentional. An objective world exists, and we have these dispositions to discriminate, we say, but how could norms emerge out of that? At first sight, questions as these seems impossible for Davidson to answer, since interpretation presuppose intentionality, even though radical interpretation in triangulation gives us an opportunity to break into the intentional from the non-intentional and non-normative "outside". In radical interpretation we can break into the intentional because linguistic

[1] In article 6 I show how some of the central intentional concepts are related.

behavior provides non-semantic evidence for the meanings of sentences and for the speaker's beliefs. Either way, this "external" reflection on the conditions for thought and language is something that we do in addition to "internal" or intentionality-intrinsic considerations, and I would say that an informed view on intentionality should address both perspectives.

In "Externalisms", Davidson, in the attempt to reach an explanation of intentionality, sets up three conditions for the possibility of thought about an external world.[2] The first one is normality. This condition involves that contents of thoughts are being indicated in a process where (as a minimum) two creatures are «responding in characteristic ways to distal stimuli», so that a common cause of the most frequent cases is being potentially located as what determines the contents of their thoughts. This means that there is something that statistically or characteristically can be picked out as a type of cause for a given content of a thought. We may call this the normal cause of a given thought, and this will not be a normative concept of normality, but a statistical concept of normality.

When creatures respond in *non-characteristic* ways to distal stimuli, errors or misrepresentations can occur. If we base our search for the normal cause of the content of thoughts on less frequent cases, errors will also occur. Thus, error is a second condition for thought about an external world. Divergence from normally similar reactions could occur when we have said that something (specified) is normal, and this is a normality set by the most frequent cases.

But, from where did the normality come in the first place? Where does it start? What can explain that something particular turns out to be the normal cause for thought? Davidson holds that that which is relevantly similar (the cases that together pass as "normal") are based on our pre-natal dispositions to perceptually discriminate. This means that our notion of "similarity" starts in our present stage of evolution and current biological capacity. We are born with a capacity to discriminate in perception in specific ways, and this is yet another condition.

[2] Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 9.

According to Davidson, these three factors (dispositions, normality and error) are necessary, but insufficient, conditions of thought and language. That is, he maintains that they cannot explain why intentional and normative activity (such as talking and thinking about parts of reality) can arise. How could we then give a sufficient account of conditions for thought and language?[3] In bringing our attention to language acquisition, which Davidson talks about as an "intermediate stage", he points us in the direction of society. I will hold that this could be seen as a fourth condition, and I think that Davidson also utilizes it as such. The four conditions for thought about an external world are thus disposition, normality (the common cause of the most frequent cases), error, and language acquisition in a social setting.

The last condition mentioned is that we are being raised in a society of language-users. When we as children acquire a first language, society is our teacher. Through ostensive processes, we learn to connect words and concepts to «normal» middle-sized objects and events. This is an intermediate, but crucial stage. If this stage were not in the picture, it would be very hard to see how a creature ever could go from acting on dispositions to acting on norms or following rules. According to Davidson's story we will gradually move on from dispositions to discriminations, through learning to connect to the world through ostension, and towards independent conceptualization and rule following. This last step, the transition towards being an independent rule-follower and a mature language user demands interaction with a linguistic society. But, we are obviously still in lack of sufficient conditions for thought and language. Davidson says that he is «...persuaded that the basic idea is right: only social interaction brings with it the space in which the concepts of error, and so of meaning and of thought, can be given application. A social milieu is necessary, but *not sufficient* for objective thought.»[4] We can therefore sum up this attempt at explaining intentionality by saying that these four conditions do not suffice for explaining intentionality, objective thought and hence error.[5]

[3] For Davidson, this question is the same as asking for sufficient conditions for rationality, intentionality, normativity, and meaning, and also for ascription of propositional attitudes.
[4] Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 3, my italics.
[5] See also article 5 on error and the teacher-learner-object triangle, where the teacher at one vertex represents "society". Otherwise, "society" is typically hinting at the interaction of two speaker-interpreters.

Now, what about a more clearly internal perspective? The advantage of an immanent approach is that it might be less open to meaning-skepticism and less open to skepticism about the world than an external perspective.[6] The main disadvantage may be that we are unable to give a satisfactory analysis and explanation of intentionality, in the sense of really coming on the outside of, or to the bottom of, intentional phenomena. A relevant and timely question is: "What would constitute *an explanation* (not only an account) of intentionality?" In his 1996 article "Externalisms", Davidson asks what the sufficient conditions of thought are, and then says that: "...if we could answer this question we would have an analysis of thought. It is hard to think what would satisfy us, which did not amount to a reduction of the intentional to the extensional, and this, in my opinion, is not to be expected. What further progress we can make will be in the direction of theory building within the realm of the rational, not reduction of it to something else."[7] Davidson's search for an explanation of error is thus to be a search from within the intentional and from within the theory of radical interpretation. Still, the felt need to find an explanation of error from the outside of intentionality and thereby also an explanation of intentionality seems never quite to have left Davidson. I would claim that a naturalization, in the sense of giving an explanation of intentionality on basis of the principles that the natural sciences accept would anyhow not give the wanting external explanation, even if it, contrary to Davidson's expectations, should prove possible. The reason is that naturalization in this sense would give us what we do not want, or need, an account where intentionality and human interest is out of the picture. This does not imply that I think we can be generally content leaving out an external perspective on intentionality.

A legitimate and understandable reason to raise the question of the problem of error in the first place is the quest for objectivity of thought and meaning, and the quest for objective knowledge, or rather a quest for the justification of our knowledge about the world and about other minds. A possible motive for a search for an analysis of the concept of error is therefore the search for an analysis of the concept of objectivity and objective truth. A need for an analysis of the concept of objectivity is understandable for anybody who is claiming (or denying!) that we can have – and do in fact have – objective knowledge about the

[6] I will not argue for this assumption here, but it is further discussed in article 1 and the synthesis.

world. And why would we make that claim in the first place if we were not to some extent in the sway of the skeptic and his challenge? It seems to me that unless Davidson felt some kind of pressure from the skeptic, he would not feel the pressure towards giving an explanation of error. This is a pressure from meaning skepticism, from the skeptic about objective knowledge about the world, from the skeptic about objective knowledge of our own minds and from the skeptic about objective knowledge of other minds. But again, as mentioned, the problem of error is also a problem internal to the method of radical interpretation.

Another possible motive for the urge to provide an analysis of the concept of error is found in challenges from Wittgenstein. I think about the challenge of explaining what it means to follow a rule, and also the challenge of explaining what is involved in saying that there is no such thing as a private language. Following a rule means going on in the right way. A learner who goes on as the teacher does would not sufficiently prove that he possesses normativity or has thoughts. According to Davidson one has to be in the grasp of truth conditions, and one has to have an awareness of the possibility of error to be an intentional creature. Going on in the right way also presuppose that there is a wrong way. Thus, in order to be a rule follower one has to possess an awareness of the possibility of error on one's own behalf. And, if we can explain and analyze the concept of error, Davidson would grant that we could explain what rule following (and thus normativity) is about. But, the private language problem-complex questions the very intelligibility of assuming that somebody could establish and maintain a language on his or her own. Objectivity is possible only in the presence of norms, or rather, in the presence of norm-possessing knowers, and norms can only be established in a community of knowers. Hence it follows that objectivity is only possible in a community consisting of several (or, at the very least, two) creatures that we can say are knowers. An explanation of the problem of error would in this way potentially explain both meaning and knowledge, and would give a clue to the normativity of semantics as well as a clue to the normativity of epistemology.

---

[7] Donald Davidson: "Externalisms", (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 9.

Explaining intentionality, normativity and objectivity is therefore the big picture when it comes to explaining error. Within this big picture, there is the more limited and perhaps more methodological question of how to explain the possibility of error *within* the model of triangulation in radical interpretation, or rather, when we base ourselves on the method of radical interpretation. This problem seems at first sight to be more limited since it deals with the question of how we could explain, or even only give meaning to, the possibility that both the speaker and the interpreter are wrong, when communication is smooth. We will start by looking at the possibility of error within Quinean radical translation and then move on to radical interpretation.

2. The possibility of error within Radical Translation and Radical Interpretation

In Quine's radical translation, the principle of charity says that I must assume it to be more probable that I have translated incorrectly than assuming that the speaker actually believe something that is "glaringly false" by my lights. In Quine's version, the principle of charity tells me to stick to the assumption of mistake on the side of ascribed meaning rather than mistake on the side of ascribed belief. So, when communication is not successful, and I seem to have run into an error, I try to revise my translation manual as a first approach. If I find that the revision called for will bring about a radical change in the manual, and this change makes the manual much more complicated, and maybe also gives it a decrease in explanatory power, this will be an indication that I should choose simplicity over correctness on the speaker's behalf. In effect, this means that the speaker (and the belief) will be counted wrong, given that I find nothing that can explain how the speaker acquired this false belief. I therefore eventually ascribe this false belief to the speaker, thus, rather than revising my translation manual, I make him less rational. Quine's principle says we should be careful in, and avoid to the greatest extent possible, ascribing error to the beliefs of a speaker in such a way. Ascription of error to the speaker is to be done only when we have no alternative option, or when we see it as choosing Scylla, when a radical change of translation manual is considered to be Carbides.

Davidson created the problem of error for himself by not appealing to Quine's concept of stimulus meaning as what needed to be matched for translation or interpretation.[8] In Quine's original formulation of the theory of radical translation, the concept of stimulus meaning was crucial. The matching mentioned is a matching of some factor in the verbal behavior of speaker or informant on the one hand, and some factor in the (possible) verbal behavior of the translator or interpreter's side on the other hand. There has to be some matching between the two in order for translation or interpretation to occur, and according to Quine's early versions of radical translation, stimulus meanings of informant and manual were the factors to be matched.

Now, the radical interpretation approach does not appeal to the concept of stimulus meaning; on the other hand, we seek the truth conditions of the speaker's utterance. The truth conditions for the speaker's utterance is to be matched with the truth conditions of a reasonable interpretation (by the interpreter) of the utterance; that is, a reasonable utterance by the interpreter's own light, of course, which in effect means that the utterance must have as its content a belief, or a possible belief, of the interpreter. In radical interpretation, truth conditions of utterances are to be matched but what does such a match demonstrate? Does it show that both the speaker and the interpreter are right or does it show that they are both wrong? The match of the truth conditions only tells us that the truth conditions are similar for the speaker and the interpreter. What can we appeal to in this situation? The matching of verbal behavior is in radical interpretation no longer sufficient for being content that the interpretation is a likely and fair interpretation.

In Quine's radical translation, matching verbal behavior is the final factor to be found in making sure one has reached an acceptable translation. ("Acceptable" meaning that the verbal behavior is translatable, and translatability is in its turn to be judged according to success in communication.) Originally, stimulus meanings were to be matched, and Quine considered the stimulus meaning of an utterance or a sentence to be the only meaning a sentence can have. However, according to Davidson, in 1990 Quine "decided to do away with the matching of neural patterns as the key to translation". Davidson further thinks that

[8] See p. 730 of Davidson's "Reply to Dagfinn Føllesdal" in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Quine from that time on is letting empathy be a "substitute" for sameness of stimulus meaning.[9] Quine's own presentation of the changed situation is described in his article "Three Indeterminacies"[10] and comports well with Davidson's description. Instead of "intersubjective likeness of stimulation" Quine stresses the importance of empathy. The linguist, or translator, is to project himself "into the witness's position"[11] Empathy [in the perception of another's perceptual situation] instead of matching of stimulus meanings thus works well for Quine and radical translation. One could say, as Davidson does, that Quine is "bypassing the problem of error on the principle that translation will succeed if errors match".[12] Or, one could just as well say that Quine never really had a problem of error in the sense that the problem of error has come to be a problem within radical interpretation. For Quine, if verbal behavior matches, the translation manual will function, and we will know that this is so when we have success in communication.[13] There are no further criteria to be found.

Now, the champion of radical interpretation could not settle for stimulus meanings to be matched, but could he possibly settle for empathy? If not, why is it insufficient for *the radical interpreter* to share the verbal habits of a speaker when it suffices for the radical translator? For the interpreter and speaker to share verbal habits is insufficient because "both might be wrong and both might simply be programmed or conditioned to act similarly when exposed to similar stimuli"[14] Radical interpretation aims at assigning truth conditions to a speaker's utterances. The truth conditions of the speaker's utterance are to be matched with the truth conditions of an utterance's reasonable interpretation, which, in effect, means a match with the truth conditions of a possible utterance of the interpreter himself/herself. According to Davidson, if we consider the first case, in which both the

---

[9] Ibid.

[10] Willard Van Orman Quine: "Three Indeterminacies" in R. Barrett and R. Gibson (eds.): *Perspectives on Quine*, Oxford University Press, Oxford 1990, p. 3 f.

[11] Ibid, p. 4.

[12] In Davidson's "Reply to Dagfinn Føllesdal", p. 730 in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

[13] "Success in communication" is according to Quine to be judged by "smoothness of conversation, by frequent predictability of verbal and non-verbal reactions, and by coherence and plausibility of native testimony." He also points out that "fluency of conversation and the effectiveness of negation" is what we judge manuals by. (In Willard Van Orman Quine: "Three Indeterminacies" in R. Barrett and R. Gibson (eds.): *Perspectives on Quine*, Oxford University Press, Oxford 1990, p. 4)

[14] Davidson in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, p. 730.

interpreter and the speaker are wrong, we would have nothing objective to appeal to in deciding that they were wrong. Moreover, he seems to think that the second case, in which both parties are programmed or conditioned to act similarly when exposed to similar stimuli, is another but similar version of the problem. It is not obvious to me that these two cases are similar, in the sense that they represent the same problem. The problem of the first case is that some standard of truth and falsity is needed to say that both are wrong. But if both were programmed to act similarly when exposed to similar stimuli, and given that nobody ever had a chance to find that they were only "programs", do we then really have a problem? For one, if we are dealing with creatures that could be programmed in this way, it is obvious that we are not dealing with creatures that are capable of independent judgment. Anyhow, from the premises given in radical interpretation, I cannot see the problem in this situation as these creatures would we only doubtfully be able to radically interpret. However, it might become a problem if we add the premise that radical interpretation is set up for *explaining* the normativity of intentionality-equipped actors, or that radical interpretation is supposed to show a distinction between those programmed to act similarly as others and those acting from their own reasons. I will not follow this lead further here, so let's just stick to the first case, where both interpreter and speaker are wrong.

Sharing of verbal habits between speaker and interpreter could, we have said, indicate that both are right or it might indicate that both are wrong. Interpretability, then, is apparently insufficient for the defender of radical interpretation, even if translatability were sufficient for Quine. Assignment of truth conditions to the utterances of the speaker is quite another matter than sheer matching of verbal behavior. And there is not much help in empathy, because it is no longer a problem of perception. However, the concept of truth and mastery of the concept of truth cannot *come out of* matching of verbal behavior [even if I would think that creatures that can *estimate* matching of verbal behavior also would be capable of assignment of truth conditions and mastery of the concept of truth]. With the concept of truth comes the idea of an objective world, but as Davidson points out, if we ask where a sense of objectivity comes in, perception and empathy in perception cannot answer this question.[15]

[15] Ibid.

On page 321 in "The Structure and Content of Truth", Davidson notes that "The difficulty with what we may call the distal theory of reference is that it makes error hard to explain, the crucial gap between what is believed true and what is true; since the distal theory bases truth on belief, the problem is crucial." Why is there a difference between a proximal and a distal theory in this question? One would think that, in principle, this could be the same. I'd say that it is rather Davidson's demand for objective truth that seems different. And it is highly probable that this has been a motivation to introduce the distal theory in the first place. He says that "The distal theory bases truth on belief" and it is subsequently vital that he manages to establish a notion of truth based on belief.

In predicting that, in radical interpretation, the belief of the speaker will emerge based on his reactions to our common environment, I am guided by what I hold true and by what I find coherent with the beliefs that I so far have ascribed to the speaker. The belief that I ascribe the speaker is therefore a hypothetical ascription partly modeled on my own system of beliefs. Now, assuming that the verbal behavior of the speaker and me as an interpreter correlates (which, in this case, involves matching of the ascription of the truth value of a sentence/an utterance), we can see the dimension of the problem of error. The problem is that even if our verbal behaviors cohere we could both be wrong, and as an interpreter, I will not have any objective correlatives as long as continued perception will only provide more of the same kind of "evidence" that I already have. Further perception and empathy in perceiving will not bring us closer, as long as matching of verbal behavior is the closest we can get in this dimension. How can I then know whether or not we are both wrong?
Let's follow this lead back to where we started. The concept of error comes with the idea of an objective world and the concept of truth. The problem of error is therefore another side of this sense of objectivity that is involved in interpretation. What we now have to do is to explain how error is at all possible in radical interpretation. Perception and empathy in perception could obviously never solve *this* problem. According to Davidson, the reason is that those who perceive propositionally already master a concept of error as they have an idea of an objective world. We cannot explain the possibility of error by appealing to perception, because perceivers who are language users are, according to Davidson, already propositional perceivers and thus already in mastery of the concept of error. So how can we then give an account of (and explain) the possibility of error in radical interpretation?

Again, if we turn to the *differences* between radical translation and radical interpretation, we can see that the problem of error emerges somewhere in the transition from radical translation to radical interpretation. What was the crucial change that occurred in this transition? What introduced the new problem? Where, or in what, does the difference lie? I think both Davidson and Quine have been right in focusing on the introduction of a concept of truth to explain the difference. In Davidson's words, the relevant difference between radical translation and radical interpretation is that interpretation aims at "assignment of truth conditions to an informants utterances"[16] Davidson further notices that "nothing in Quine's approach seemed to me suited to explaining how the concept of truth was to be introduced." This should not be a surprise to anyone familiar with Quine's approach. Quine's focus was simply never on truth or on explaining normativity.

Quine always emphasized that there is no more to meaning than what can be pointed out empirically in a situation of radical translation. In Quine's words: "there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable stimulations."[17] Thus, regularity in verbal behavior is what we have on the table when collating meanings and other intentional phenomena. But, we will also make adjustments towards securing internal coherence and rationality in the systems of meaning and belief that we ascribe to the speaker. This is secured by the principle of charity. So, there is regularity of verbal behavior *and* there are some internal and holistic considerations for the system as a whole to take into account when collating linguistic meanings. The weight of these considerations, as well as the weight of the empirical clues, must be measured up against the aim of a well functioning manual of translation. We know that the manual works well when we can communicate successfully with the speaker. If, in this setting, we were to think of what is involved when the verbal behavior of the speaker and the verbal behavior of the translator complies (which means that, as translator, I agree to the utterance of the informant based on what I can see that he sees), we would simply think that as long as agreement exists and communication is smooth, there is no more to ask for in the sense that nothing more is needed and nothing more is possible. The

[16] Ibid

[17] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, preface.

possibility that both of us might be wrong when in agreement therefore has no place in the setting of radical translation, and is not a problem anyhow.

What is then different in the setting of radical interpretation? Something must be different, since the possibility that both of us might be wrong poses a problem in radical interpretation and not in radical translation. First, the concept of truth has been brought in. It seems to be the shift from evidence to truth that makes such a big difference. Could there be something about the metaphysical pretensions of the concept of truth that simply does not cohere well with the otherwise empirical and naturalized conception of the radical interpretation thought experiment? Even if truth is extensional in Davidson's theory and a relation of satisfaction is all that is demanded, a T-sentence also requires a previous understanding of the concept of truth. When we say "'snow is white' is true if and only if snow is white", we must understand the concept of true in order to understand what this sentence means.[18] Hence, it is insufficient to know the truth conditions of "snow is white". This previous understanding of truth (which we rely for the theory of truth to provide a theory of meaning) is crucial. It is the attitude of "holding true" that is the interpreter's guiding star, but as we know, neither the speaker's nor the interpreter's attitude of "holding true" are more than analytical hypotheses. Secondly, the principle of charity, and thus the consideration for the system as a whole, has gained weight. There is definitely a difference between the "across-the-board-ascription" of charity that has been prescribed in radical interpretation and the minimal principle that Quine introduced. This made for a shift away from the more empirical orientation (which was predominant even in the earlier versions of radical interpretation) towards a greater emphasis on overall considerations of rationality and internal coherence.[19] Otherwise, as it were in radical translation, we basically have to

[18] Davidson does not explain the concept, as he takes it to be primitive.

[19] According to Dagfinn Føllesdal, Davidson's second thoughts about the maxim of maximizing agreement "saved" him from an even heavier leaning on "reason" to the disadvantage of "perception". Føllesdal's own example, with the big tree between the rabbit and his interlocutor as a cause of disagreement that could be explained by the interpreter, is presumed to be the reason for Davidson's change of mind. It is by Føllesdal taken to have lead to Davidson's modification "maximize agreement where you expect to find agreement". One wouldn't expect one's interlocutor to agree when you know that he doesn't see what you see. The observational constraint on interpretation was according to Føllesdal after that well taken care of by Davidson. (Dagfinn Føllesdal "Triangulation", pp 719-728, in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, p. 722 and 723) If Føllesdal is right, it seems that interpretationism as an epistemological strategy for explaining objective knowledge about the world very nearly were lost with the maxim of maximizing agreement, or at the least all but coherence-theoretical ambitions would be lost.

rely on observations of behavioral regularity when comparing meanings and beliefs in radical interpretation. But it is obvious that the crucial difference does not have to lie with *only* these two differences, a combination of them, or with either one alone.

Hoping to address the problem of explaining the possibility of error in radical interpretation in a more fruitful manner, we will take a closer look at the notion of error itself. What do we really mean by "an error" or "a mistake" in connection with radical interpretation?

3. Two kinds of error

For an interpreter, there are two possible sources for mistakes. He can be wrong on what is perceived by the two parties as the common object, or he can be wrong about words. Being wrong about words could be a mistake about where to locate the belief ascribed to the speaker into the already ascribed system of belief of the speaker, that is, its location in the truth theory for the speaker. Alternatively, the interpreter could employ a non-appropriate translation that certainly also has the result of giving the belief a wrong location in the system ("conceptual inadequacy"). Now, obviously, the speaker can also make mistakes. In a similar fashion, the speaker may make a factual mistake, make use of the wrong words, or employ inadequate concepts. Hence, there are two kinds of error: On one hand, error that involves false beliefs and, on the other hand, error of meaning. In terms of an individual, or from the speaking subject's perspective, this difference will, on one hand, involve error as saying something false and, on the other hand, error as not knowing the meaning of the words used, or applying words in a way that precludes an interpreter's understanding of the other.[20] Of course, both kinds of error are familiar to us, and we have experienced this often enough, in others and in ourselves.

[20] There could of course be many ways of making a linguistic mistake, or being wrong about words. In the radical interpretational frame we could say that it lies somewhere in a continuum from the speaker's being insensitive to or unjustified in her assumptions about how the other will take what she says, and to, in the other end of the line, how the other (the interpreter) is insensitive to or unjustified in the way she is taking what you say. This is all there could be to a linguistic mistake in the interpretational frame that we here have acceded to. And sometimes the failure will perhaps be beyond sensitive and justified speaking and interpreting, The failure could be complete, such that communication breaks down, or it could be only slightly confusing, or any place in between.

According to the radical interpretational perspective there are only two factors that determine the content of a belief. First, there is the object or event in the world that caused the belief, and second, there is the location of the belief in relation to other beliefs in the network of beliefs. So, according to this view, we would believe that if something goes wrong, the error must be in one of these two areas. We would not say that both of these errors deal with false belief, we would rather say that the first kind of error (error in the assumption of what caused the belief) is about false belief and that the second kind of error (error in where the belief is located) is error in meaning, or inadequate employment of concepts. Now, within radical translation the question is, of course, whether we can separate belief and meaning so that we can know what kind of error we are dealing with. In early phases of radical translation there is no actual possibility to separate the two, but we use the principle of charity as an *analytical* tool to find the meaning of the speaker's sentence by holding his beliefs constant. This means that the separation is hypothetical and we are producing an analytical hypothesis about the meaning of the sentence. So, in the very beginning of radical interpretation, we could not actually find out whether an error is caused by wrong belief or wrong meaning. However, the situation changes when we reach a level of smooth communication, as we can ask questions and, possibly, collect enough evidence to make a more substantiated decision on the origin of the error. As the process of interpretation proceeds and our hypotheses are either corroborated or replaced, we have more evidence to back up our claim (that the error was an error of meaning or an error of belief).

In radical interpretation, we never reach a level where we can be completely certain that *the* correct hypothesis has been found, and this will also be true in the case of determining which kind of error we are dealing with. Indeed, Davidson would hold the view that there is a principled indeterminacy in this question.[21] Since we have said that there are two main

[21] In "What Thought Requires" Davidson asks "How will the learner or interpreter discover when he is applying a different concept than the one his teacher or informant had in mind, and when one of them is misapplying the same concept? Some answers to this question will appeal to the power of consensus, but this cannot be conclusive. Of course, consensus of *use*, where use is assumed to reflect what the teacher or society *means*, is just what the learner or radical interpreter needs to recognize, but consensus of *application* does not distinguish the two varieties of error. As far as I can see, nothing in the observable behavior of teacher or learner with respect to an isolated sentence can sort this out. Further distinctions depend on relations among uttered sentences. The relation of evidential support provides powerful clues…." In this sequence Davidson is dealing with the two kinds of errors, but he focuses on the situation of teacher and learner, not the situation with two speaker-interpreters. He further says: "the large and necessary step is learning to *explain* errors. [] It

kinds of error in interpretation, it seems natural to ask: "What are the conditions under which they can emerge?"

Propositional seeing – a condition for two kinds of error

Where does error enter the picture? As long as there are only casual relations to the world (reactions to objects and events that a creature perceives), I will argue that error of the "meaning kind" cannot exist. A purely causal relation to the world, and only reactions to the objects and events that a creature perceives, seems equivalent to Russell-Hanson's notion of "seeing".[22] With only "seeing" going on, there can be no error of the kind "mistake in employment of words", but with "seeing as" (or, even more obvious, "seeing that"/propositional seeing) there is a chance of this kind of error coming in. The possibility that I can be wrong in employing words and concepts cannot arise when only seeing (or perceiving) is going on, but this is not to say that there can, in fact, be only seeing going on for us at all (if "us" refers to "adult human beings" or, rather, "human beings with a fully fledged language"). It seems to me that when we "see as" or when we "see that", as we may have little chance of avoiding, the possibility for this kind of error is always present.

This is not to say that everything that goes on in our heads is propositional, but surely, it means that all that we would properly call *thinking* is conceptual or propositional. According to Davidson's view, there is no difference between the conceptual and the propositional; indeed, having a concept means having already classified, and that *is* to form a proposition. In the instant a category is formed, some things belong to it and some things do not. In this way, the propositional is not an additional step upon the conceptual, as the conceptual is always propositionally structured in advance. Thus, propositional seeing is not a general condition for error in perception, but a condition for error of meaning. This means that propositional seeing is a condition for the possibility of *two* kinds of error.

---

is [] at this point that the distinction becomes clear between falsely thinking a bull is a cow, and simply applying the word 'cow' to both." In Davidson's *Problems of Rationality*, Oxford University Press, Oxford 2004, p. 144 and 145. It thus seems that Davidson hold that we actually *can* sort out this difference when we are full-fledged users of language, and thus that we can decide which of out of two kinds of error we are dealing with.

[22] See Norwood Russell Hanson: *Patterns of Discovery*, Cambridge University Press, London 1958, rpr. 1969, pp. 19-24.

What about the other kind of error, the kind that involves only "seeing"? If we think of the example with the dog barking up the wrong tree, it feels natural to say that he is in error. It seems to be an adequate description of the situation to say that the dog is making an empirical mistake. But how can we know this for fact? What would our evidence be for saying that the dog, upon the discovery of the cat coming down from another tree than the one he was barking up, realizes that he has been in error? If I maintain that we couldn't know what kind of mistake the dog has made, this will *not* automatically amount to a denial of any form of thinking in the animal. However, it does imply a denial of the methodological availability of the mind of a non-linguistic creature, and this is so because we cannot, in the dog, find and specify a particular set of beliefs in which to locate the (presumably) wrong belief. Hence, we cannot decide that the dog's error is empirical because this would demand another kind of notion with which to contrast it, and methodologically this is out of question. The case of the possibility of mistake in non-linguistic beings is therefore out of reach for us, and this is *my* main reason for reserving our specific notion of mistake for speakers.

Davidson could be willing to go a bit further here, and say that as long as the dog doesn't have concepts it cannot be making a mistake. Davidson explicitly denies that the notion of mistake here involved applies to speechless. He would hold that the dog has the power to discriminate, and maybe also that the dog is doing what it is programmed to do. Davidson would hold that as long as the dog, from its point of view, is unaware of its own mistake, we would not have the right reason to say that he has mistaken one tree for another. We should note that the applied concept of mistake is not just any concept of mistake, but a concept established from within intentionality. In general, there are, of course, good reasons for saying that dogs make mistakes, but it could not be justified in such a way that it could be part of a set of reasonable analytical hypotheses of a radical interpreter. One of Davidson's other reasons for limiting ascription of concepts is that we could end up in a situation in which we would have to ascribe concepts to any living organism responding to its environment.[23] And it is plain to see that this could come close to making the notion of concept meaningless. According to Davidson, "A creature that has a concept knows that

[23] See for instance Davidson's "The Problem of Objectivity" in *Problems of Rationality*, Oxford University Press, Oxford 2004, p. 8.

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."[24] Thus, we have to conclude that the concept of truth and objectivity combines with awareness of the possibility of error, and also, that the awareness of the possibility of error comes with propositional thinking. On this point, there is a potential discord between Davidson and me, and the discord would lie in the unclarified justification for the case. I will maintain that the reason for the case is epistemological; based on the methodology of radical interpretation it is a question of limited access to a dog's beliefs. An alternative stand is that we also rely on an ontological reason in drawing this conclusion.

Idiolect and two kinds of error

Eva Picardi has also been preoccupied with the notions of error and mistake within a theory of radical interpretation. She is questioning the possibility for distinguishing between the two different kinds of error in Davidson's theory of radical interpretation. "... if we forfeit all appeal to a socially shared language, the distinction between making a linguistic mistake and making a conceptual or factual mistake becomes too blurred. It is a danger of Davidson's whole approach that the notion of an error becomes an utter riddle."[25]

I agree with Picardi that it is important to uphold the distinction between the two kinds of mistake. If we think about Quine's original formulation of the principle, in which he holds that we should assume that we have misunderstood the speaker's language rather than that he is wrong about the world, and Davidson's description of the application of the principle as "... holding belief constant while solving for meaning", it soon becomes clear that in radical translation and radical interpretation we have to be able to make some sense of such distinctions as the ones Picardi mentions. The distinction between belief and meaning is crucial to the project of radical interpretation, even if we cannot separate the two until relatively late in an interpretational process. Picardi's point seems, however, to be that it is Davidson's denial of language as socially shared that threatens the notion of error. It is not clear from the context which part of Davidson's theory she hints at. But he has famously

[24] Donald Davidson: "What Thought Requires" in *Problems of Rationality*, Oxford University Press, Oxford 2004, p. 141.
[25] Eva Picardi: "Sensory Evidence and Shared Interest", pp 171-185 in Mario de Caro: *Interpretations and Causes*, Kluwer Academic Publishers, Dordrecht 1999, p. 182.

denied that there is such a thing as language, which then is identified as "a clearly defined shared structure which language-users acquire and then apply to cases".[26] In this context, it is first of all a denial of conventions as governing sentence meanings, and the trouble with Picardi's point is that it is by no means obvious that language has to be involved with conventions in order to be socially shared. Even so, Picardi seems to hold that it is Davidson's focus on the idiolect that endangers or blurs the distinction between making a linguistic mistake and making a factual mistake. And I think she is right about that, even though I disagree that Davidson's anti-conventionalism is to be blamed for the predicament.

4. The Social Check for Error

The thesis of the indeterminacy of translation says there are many *correct* ways of describing a specific part of reality. But how can we know if, or when, one of the ways of describing reality is *wrong*? Some ways are wrong ways of describing things; that is to say, some descriptions or beliefs are false. As we have seen, there could be at the least two sources for an incorrect belief. First, that one thinks one has seen something that one possibly couldn't have seen, thus that one has a perception that is not in concordance with the ways of objects or events themselves. The only possible check here is a social check, since we have to check our percepts, or rather perceptual beliefs against the percepts/perceptual beliefs of others. Even if one would not admit of something like "the ways of objects and events themselves" ("the given"), one would have to admit that some beliefs or descriptions of reality could be outright false. The other possible source is wrong words, i.e. that one gives a wrong description because one has chosen words to express one's perceptions that do not cover the perceived, or one has chosen words in such a way that the interpreter could not be expected to understand. One then has perceived things correctly, but employed words in an unacceptable or inadequate way. Again, only a form of social check can be performed to say this much. Both sources of error could lead to impossible descriptions and false sentences, but only one of these two possibilities involves false beliefs about the world on the part of the speaker.

[26] Donald Davidson: "A Nice Derangement of Epitaphs", pp. 433-46 in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford 1986, p. 446.

Now that we have established that a social check is the relevant check for error, no matter which sort of error, it is time to deal with the details of a social check. How is the check to be conducted within the theory of radical interpretation? In Eva Picardi's "Sensory Evidence and Shared Interests", as recently mentioned the question is how a social check at all could be performed when we operate within a theory where the notion of language is confined to idiolects.[27] Could the radical interpreter ever determine whether or not the interpretee employs words in an unacceptable way? In fact, it seems that the radical interpreter could only check this against earlier occasions of description of parts of the world from the mouth of this particular speaker, at a given time-sequence, as we have to assume that the only available concept of language is individual, e.g. idiolects. Each act of radical interpretation deals with only one speaker for a limited period of time. If the speaker seems to have made an error and the interpreter's analytical hypotheses seems in some sense invalid, the reason for this has to be found in the speaker's previous utterances, in specific circumstances, or the reason could not at all be found. The radical interpreter can only consider previous utterances and previous behavior in general. Thus, he can check the coherence with previous utterances and correlate the circumstances for the present utterance with those of previous utterances. This means that the interpreter is (again) checking the relationship between what he can perceive, which he can see that the speaker also perceives, and the utterance that the speaker gave as a response to the common perception. A perception check will not provide anything new. On the other hand, the interpreter would often have a so-called "prior theory" about how to deal with the utterances of the speaker, and this prior theory could be based upon earlier experiences with members of the same community or even with the same speaker.[28] The prior theory of the interpreter *could* be seen as a rudimentary theory about the language of the society.

[27] Davidson has tried to calm similar worries by saying that "the intention of the speaker to be interpreted in a certain way provides the 'norm'; the speaker falls short of his intention if he fails to speak in such a way as to be understood as he intended." (Donald Davidson: "The Second Person", in his *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001, p. 116.) Thus it seems that the norm for the speaker, that by which he knows how to go on, in order to go on according to the rule, lies in himself. This means that the speaker, and not (only) the interpreter represents society, in the sense that the speaker holds the norm to such an extent that he knows what he have to do in order to be understood by the interpreter. But, as we will see, this will explain the problem of error only up to the point of understanding and smooth communication between speaker and interpreter. And this is Quine's criterion for success in radical translation, but Davidson cannot resign to smooth communication due to his demand for objective truth. (The worry were mouthed in the question: "But haven't we, by eliminating the condition that the speaker must go on as the interpreter (or others) would, at the same time inadvertently destroyed all chance of characterizing linguistic error?" Ibid.)

However, based on this, the interpretation would not typically be radical; and even so, the point is that the interpretation of the speaker is expressed in the interpreter's "passing theory". It seems the interpreter has no way out. Therefore, the common or societal aspect of language use between the radical interpreter and the speaker who communicates smoothly, is so to speak, fastened only in one end. It seems the social confinements have had the chance to manifest themselves only in the emergence of the speaker and interpreter's individual languages (i.e. in ostensive language learning).

It seems as if though intersubjectivity and the social are important only for language *acquisition.* This is in discord with Davidson's view on the matter. According to his general view, corrections and refinements of beliefs and linguistic expression of beliefs is a continuous process. But, given that the concept of language is individual, and therefore, that a speaker's language can only be checked against itself, how can a social check be performed? The radical interpreter can only perform such a check based on the communication between the speaker and himself on a given occasion. Hence, what he can check, and which could be relevant here, is only whether or not he and the speaker in a setting with "occasion-sentences" seems to be choosing the same objects or events in a manner that establishes regularity of verbal behavior. In this way, he can check over time whether the correlations between verbal behavior and stimuli are the same for the speaker, and whether, over a given time, these correlated sets are the same for the two parties. Otherwise, disrupted communication is the only way to notice that something is wrong. But this means that the interpreter cannot know that anything is wrong when communication is *not disrupted.* This is what I would call "the hard problem"; *(how) could an interpreter come to know whether or not both interpretee and interpreter are wrong, as long as communication is smooth?*

Hence, in the sense that the radical interpreter cannot access a "society", or third party, the radical interpretation situation does not seem to allow social checks. The interpretationist framework assumes that the two-creature methodological frame is already intersubjective and, thus, social. But from the radical interpreter's point of view, social checks cannot

---

[28] Donald Davidson: "A Nice Derangement of Epitaphs", pp. 433-46 in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford 1986, p. 442.

actually be performed. There is nowhere to turn for further corrections or corroborations. That established; why do we then need a social check at all when communication is smooth? Why wouldn't we just think that as long as communication is smooth nothing is wrong?

Again, it seems to be the demand for objective truth that forces us into this situation. In Quine's theory of radical translation, this situation never occurs, and it seems to me that one of the reasons is that objective truth was never a basic concept in his approach.[29] As long as communication is smooth, translation manuals can be set up in satisfactory ways, and something more is never demanded. The notion of objective truth seems to be out of reach for radical interpreter and speaker. What could then be the solution to our problem? To stop yearning for objective truth? To change our language concept from individual to social to enable understanding and explain that two communicating speakers can both be mistaken about the world and about the employment of words? If we choose the latter, we will, in effect, have dismissed radical interpretation as an approach in its entirety. The methodology of radical interpretation is built on interpretation of a speaker (and his language) within a given period of time. I cannot see a way to modify this fundamental condition within the framework of radical interpretation. If we decide that we should stop yearning for objective truth, it is hard to see how we could still find the approach of radical interpretation relevant for our understanding of human intentionality, thought and language, and not the least; how radical interpretation can be important for an investigation into the conditions of objective knowledge. A notion of objective knowledge without a notion of objective truth is not easy to cash out.

My ambition is not to give a solution to the problem, but our interests in the truth about the world other minds and ourselves cannot be dismissed easily. On the other hand, if we dismiss the idiolect as the interpreted language (and this would involve a dismissal of radical interpretation), it is difficult to see any alternative methodology, if we still would

[29] Another reason is of course that in Quine's radical translation the situation between linguist and informant is never one of equality. There is no mutual interpretation going on, as in radical interpretation, where both parties could take each other's reactions to a common environment as a possible source for correction of one's own system.

like to endorse interpretationism in epistemology.[30] It seems that we are facing a dilemma, and I suggest that this dilemma is the reason for Davidson's problem with explaining error.

5. The partial solution to the problem of error - Triangulation

In my analysis, I have assumed that Davidson, among other aims, seeks an understanding of the transition from the pre- or non-propositional to the propositional. It seems to me that he wants to understand how the transition can take place, but does not think that we can account for this in detail. He also does not think that we can reduce the propositional to the non-propositional. Still, he has been constantly searching for an explanation of "the origin of, or basis for, the concept of error." [31] In his intellectual autobiography, he says that he eventually found at least a partial solution to the problem. He thinks that reading Kripke's "Wittgenstein on Rules and Private Language" may have given rise to the solution, and that one aspect of the problem anyway is stated in that book. "What Kripke saw was that a person cannot make a mistake without an external check, and the inanimate world cannot provide such a check."[32] Based on this, Davidson formed his picture of triangulation, in which two people share visual access to an object or an event, and in which each of them can correlate the other person's reaction to the same object or event to what he or she senses. This will *not*, according to Davidson, give an explanation of error, but triangulation "makes room for it, for there will be cases where one person's reaction to a shared stimulus is not what the other person has come to correlate with that stimulus."[33] This means that an external check could only be made by another person who is also a language user, and thus, that the concept of error relies on a society of language users. Therefore, Davidson's conclusion is that error can only be ascribed and made sense of in interpersonal situations. In triangulation, we find no more and no less than necessary conditions for a concept of error.

---

[30] In article 5 I discuss a possible solution for the dilemma, (instead of giving up radical interpretation) where we seem to be able to uphold our interest in truth as well as the methodology of radical interpretation. I think of the suggestion of introducing a third speaker-interpreter.

[31] Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, p. 66.

[32] Ibid.

[33] Ibid.

Given that we accept propositional thought as language dependent, what more could be said, within Davidson's framework, about the origin of the concept of error? In the theory of radical interpretation, the concerned language is the idiolect. It is hard to see how we could give an account of language as social based on the idiolect and the model of radical interpretation. We can give a story along the lines of disposition plus ostensive language learning in triangulation, but if more is desired in terms of an external explanation of language, I fear that a reduction of intentionality and normativity would be imminent. The reason why we, for the time being, have to confine ourselves to this partial solution, seems to be, on one hand, the mismatch between the idiolect as the language for investigation, and, on the other hand, the necessity of a community-based concept of language as fundamental for a concept of objective truth and a concept of knowledge. [34]

[34] I thank Jan Harald Alnes for discussions and suggestions on the problem of explaining error, and for constructive criticisms of several drafts of this article. I thank Frode Sirnes Larsen for inspiration and discussions on Davidson and error, and for comments on this article.

5. Error and the Teacher-learner-object Triangle

1. The problem

This article concerns the issue of the notion of error in one of the situations described as triangulation in Donald Davidson's writings. The situation in focus typically involves a teacher, a learner and an ordinary middle-sized object. I am approaching the problem of justifying a difference between a learner that *is* aware of the possibility of error and a learner that is *not* aware of the possibility of error. I discuss two suggestions intended to solve some of the problems involved. Davidson has taken up the first of the two suggestions[1], which is that the teacher tripod is to be understood as involving at least two creatures in discussion. The second suggestion is mine, and it is that in order to ascribe thoughts to a creature, the possibility of communicated disagreement is to be counted as a criterion.

In the question of how a difference can *emerge* between being and not being aware of the possibility that one can make mistakes in predication about the world, Davidson points us in the direction of society. He calls our attention to language acquisition, which he talks about as an "intermediate stage" between being and not being in mastery of the distinction between truth and falsity. He emphasizes the fact that when we as children acquire a first language, "society" is our teacher. Through ostensive processes, we learn to connect words and concepts to «normal» objects and events. This situation (and stage) is intermediate, because there is no possibility for learner error when he still cannot fully grasp truth-conditions. This is not to say that error presupposes a grasp of truth-conditions, but if the learner is *aware of* the possibility that something he says can be wrong, he *is* in the grasp of truth-conditions. The awareness of error, then, is rather an indication of such competence than the grasp of truth conditions being a presupposition. Further, considering the stage intermediate does not reduce its importance, on the contrary; if this stage and such a situation are not in the picture, it is very hard to see how a creature can go from acting from dispositions to acting from norms or following rules. Hence, when we become language users, we move on from dispositions to categorizing the world in specific ways, through learning to discriminate through ostension, and towards conceptualization and

[1] The occasion was his seminar "Truth and Predication" at UC Berkeley in the spring term of 2000.

rule-following. But the transition demands interaction with a linguistic society. This kind of interaction is the subject matter for the triangle with the teacher and the learner.

In this article, I examine the teacher-vertex extensively in order to evaluate the task(s) that is incumbent on it. My line of thought suggests that this vertex seems more static in Davidson's theory than necessary. The teacher sometimes symbolizes the third vertex (in which the speaker and the world is located at the other two vertexes), and, sometimes it is symbolized by the second person, and sometimes by the whole linguistic society. By securing a dynamic and dialogic third vertex, we can further develop the picture of language provided by Davidson's theory.

2. An adequate account of error

Davidson thinks that we lack an account of error, or what he calls "an analysis of the concept of error". In relation to the teacher-learner triangle, he says that: "In the potential gap between how the learner goes on and how the teacher goes on there is room to introduce the idea of error, not only in the mind of the teacher, for we are assuming it is there, but also in the mind of the learner. But how that room gets filled I have not said."[2] We therefore seek an account of error, and, specifically, an "awareness of the possibility of error". The problem we are approaching is; how come that the learner ever gets the idea that error is possible? A counter-question here is why the child couldn't get the idea of error from his mother's corrections of him.[3] I would say that the child would then get an idea that he made a mistake, even if he would not yet himself be able to figure out the norm that his mother was acting from in her corrections. Either way, this would not give the child the distinct idea of error we are concerned with as this kind of response could also give a non-language using creature a feeling that something didn't quite work out well. A dog will not develop *a concept of* error, no matter how often it is being corrected. To turn it on its head, the problem that we are approaching is: How could *the concept of* objective

[2] Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 11.
[3] Bjørn Ramberg posed this question on the occasion of my presentation of an earlier version of this article in a Workshop on Davidson and Triangulation in Oslo in 2001.

truth ever dawn on the learner? And, in terms of behavioral confirmation, what would be sufficient evidence for that the learner possesses awareness of the possibility of error?[4]

3. The third person

In the triangular situation with a learner and a teacher, the first of the two suggestions to be discussed is to expand the triangle, or, rather, one of the vertexes in the triangle. The motivation is that this should enable us to see how a learner could get the idea that error is possible, i.e. how a learner could get hold of the distinction between belief and objective truth. A fourth element, a third participant, a teacher number two, in other words, another competent speaker-interpreter, now enters our little society.[5] His or her presence will signify the presence of a "debating society", a society in which different opinions compete.[6] The third participant will then be able to question the first teacher's authority, and this behavior can eventually become the model for the learner. The idea of the possibility of error, mistake or frustration of expectations *could* then dawn on the learner. She or he would learn something more than just how to relate words and world, she or he would also have a model-couple, so as to be able to learn how to be a proper participant in linguistic interaction. Competent speakers in discussion provide a learner with very different and, indeed, more dynamic surroundings than can be provided by an environment with a single teacher.

This "picture" of what it is like to master a language is the picture that Davidson assumes in his general view on language. The pure ostensive process with adequate conditioning that is dominant in the triangle with a learner and only one teacher is not meant to cover the ground, but is construed for making specific points explicit and for testing out specific

[4] A penultimate problem here is that we need to answer these questions on an empirical (the factuality and psychology of language learning) as well as a theoretical level (the concept of error in the thought experiment of radical interpretation and triangulation).

[5] Davidson discussed the idea of a third person in his seminar "Truth and Predication". It was partly inspired by experiments, whereby psychologists were trying to teach a parrot to talk, and the method was called "the rival-model method". He advised me to try and push his idea further.

[6] Davidson gives a formulation of a scenario with a debating society in "Replies", in Critica 90, 1998, when he says "Seeing rationality in others is a matter of recognizing our own norms of rationality in their speech and behavior. These norms include the norms of logical consistency, of action in reasonable accord with essential or basic interests, and the acceptance of views that are sensible in the light of evidence. *These various norms can suggest conflicting ways of interpreting an agent (for example, there are different things an agent may mean by what she says), and there may be no clear grounds for preferring one of these ways to others.* Hence, balancing the claims of competing norms in interpretation introduces a form of indeterminacy not found in the indeterminacy that abounds in physical measurement." (p. 102, my emphasis.)

ideas. Davidson's general picture of language is a picture in which uncertainties are at home, where fallibility is more obvious, and language is more dynamic and less reified. The picture of competent speakers in discussion is simply a more adequate picture of what it is like to master a language, while the teacher-learner-object triangle is an intermediate stage. Hence, my suggestion introduces fallibility also at the intermediate stage, and I would like to emphasize that even when a child acquires a first language, competing descriptions will be abundant. One of Davidson's motives for bringing in this triangle has been to show the opposite, that the learner does not, and cannot, ascribe error and false belief to the teacher at this stage. The introduction of another teacher may therefore seem destructive in bringing this insight. Either way, the point of launching this suggestion is to investigate into eventual theoretical implications of a second teacher. I will discuss this suggestion in detail after introducing of the second suggestion.

4. Criteria for ascription of propositional attitudes.

On what grounds can we ascribe propositional attitudes to somebody, according to Davidson? On this, in "Rational Animals", Davidson says that full-blown language is required for thought. He says that "...in order to have the concept of belief one must have language"[7] He therefore admits to believe that only language can supply the condition for thought, even though he cannot demonstrate "...that language is necessary to thought."[8] Thus, he *is* saying that language is a reason for ascription of propositional attitudes, but he realizes, of course, that this doesn't help much in explaining thought. This means that thought and language have the same criteria, and we are no better off by saying that language is a criterion for ascription of propositional attitudes.

The first criterion that he mentions, and that he is serious about *as* a criterion (in "Rational Animals") is that for us to be able to "...intelligibly ascribe single beliefs to [somebody], we must be able to imagine how we would decide whether [somebody] has many other beliefs of the kind necessary for making sense of the first."[9] Thus, a first criterion for ascribing a propositional attitude is that a creature can intelligibly be ascribed it *as one*

[7] Davidson, Donald: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 478.
[8] Ibid, p. 477.
[9] Ibid, p. 475, 'somebody' is being replaced with 'a dog'/'the dog'.

attitude in a *set* of such attitudes, because propositional attitudes "...come only as a matched set."[10], and this set of attitudes is (partly)[11] what makes it possible for anyone to individuate, and therewith, identify, particular attitudes.

Next, Davidson argues that since "...propositional attitudes require a background of beliefs..."[12] he holds that "Without belief there are no other propositional attitudes..." and so he chooses to focus on beliefs when stating conditions for thoughts generally. The next criterion is therefore a criterion for having beliefs, and it says "...in order to have a belief, it is necessary to have the concept of belief."[13] Having the concept of belief could be described as having second-order-beliefs. Hence, possession of second order belief is the second criterion for ascription of a propositional attitude to a creature.

In addition to language, we have so far found two criteria for ascribing propositional attitudes to a creature, whereof the first is that a creature can intelligibly be ascribed a propositional attitude as one attitude in a set of such attitudes, i.e. that the creature *has* a set of attitudes, and the second is that the creature has the concept of belief or that he possesses so-called second-order-beliefs.

In "The Centrality of Truth", Davidson says that: "...we can only say that the child *thinks* something is red, or a ball, if it appreciates the distinction [between success and failure] for itself; the child thinks something is red or a ball only if it is in some sense aware that a mistake is possible. It is classifying things, and it may have put something in the wrong

[10] Ibid. p. 473

[11] As Davidson is hinting at in the end of "Rational Animals", where he introduces the metaphor of triangulation for the first time, and as he goes on to argue more carefully in "A Coherence Theory of Truth and Knowledge", the objects and events that cause a particular belief (or any other attitude) is the other factor that identifies this particular attitude. (Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985 and Donald Davidson: "A Coherence Theory of Truth and Knowledge" (1983) pp 307-319 in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore, Blackwell, Oxford 1986) Thus, it is the third point in the triangle (where we; another creature and I, are located as the two other points) that, together with an individuation of a belief as placed in a network of beliefs, can identify a particular attitude. "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." (P. 475 of "Rational Animals")

[12] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 478.

[13] Ibid.

slot."[14] Davidson describes this as the difference between "conditional response" and "...what Wittgenstein called 'following a rule'". He further notices that "This is where the concept of truth enters, for there is no sense in saying a disposition is in error - one cannot fail to 'follow' a disposition, but one can fail to follow a rule."[15] In effect, a conditional response will be a kind of copying, or "parroting", while if one is following a rule, one is in charge of language, and in charge of the distinction between belief and objective truth. One possesses the *concept* truth. For Davidson, being aware of the possibility of error (and thus having second-order beliefs) means that one is in possession of the concept of truth.[16]

What we have said so far is that only when we can know that a belief (or any other propositional attitude) is involved does the concept of truth have an application. This means we can know that somebody think when they *can* potentially think something that is false, *and* are themselves aware of this possibility. But we have to ask how this can be checked out in terms of behavior. So, this is my second problem. As mentioned earlier, my first problem was: How can the learner get the idea of the possibility of error or the idea of the concept of objective truth? This problem is related to the second problem: how can the difference between being and not being aware of the possibility of error be detected?

5. Disagreement and interpretation - new criteria?

So how could *we* know that *they* know that they can be wrong? Expression of uncertainty, or if they ask us? Typically, the child will look at you with big round eyes, and a questioning expression in his or her face. But could we really say that we know they have propositional content unless they are able to disagree? In the end, I believe that awareness of the possibility of mistake might be insufficient for us to acknowledge others as thinkers and believers; we would perhaps even require that they could produce accounts that compete with our accounts of the world. Thus, that they *at least* are able to disagree with

---

[14] Donald Davidson: "The Centrality of Truth" (1996) pp 105-15 in *Truth and its Nature (if any)*, ed. by J. Peregrin, Kluwer, Dordrecht 1999, p. 112.

[15] Ibid.

[16] We could always ask what kind of a concept or notion of truth that is involved when such a connection is being set up. A psychological state is set up as the basis for mastery of a concept. This could imply that the concept of truth has some kind of a naturalistic basis, and that its basis is anyhow not metaphysical. In other connections Davidson's concept of truth could seem to be metaphysical. We can anyhow note that ascription of truth and ascription of conditions, a basic operation in radical interpretation, is not necessarily presupposing a metaphysical concept of truth.

us, but it might be that we would also expect others to be able to give competing accounts. Disagreement and the ability to give competing accounts (or; the ability to independently describe) will then be further criteria for ascription of propositional attitudes to others. It seems to me that when we judge whether or not we can (with reasonableness and an adequate amount of certainty – a question of methodology) ascribe propositional attitudes to specific creatures, the possibility of disagreement, and not only awareness of the possibility of error on their own part, is what we look for.

Davidson says: "...without the exchange of propositional contents, there is no way they can take advantage of their ability to triangulate their shared world."[17] In this case, the 'exchange' of propositional content is assumed mutual. But even if the creatures cannot give their own competing interpretations (which, by definition, is recognized as fallible), there could, according to Davidson, still seem to be a basis for ascribing thoughts to them, on the indication of awareness of error alone. As I see it, the problem is that they wouldn't be able to share their thoughts with us, so how could we know? My suggestion is therefore that the ability to describe and the possibility for ascription of thought coincide in creatures. I think that this suggestion is more Davidsonian in spirit than Davidson's own suggestion.

Even in this case we could ask: What are the basic conditions for disagreement? Are there such conditions, and if so, are they different from the conditions of error? What does it mean to understand disagreement? What are our criteria for being sure that others disagree with us? The answers here may come easy, but also lack explanatory power; when others produce a competing description of the world we know that they disagree.

Is disagreement another criterion, or is it an assumption of "error"? My claim is that if I know that I can be wrong, then I can know that if I say something wrong, you would disagree. And further, if I know that I can disagree with you, this has to do with the fact that you can be wrong. But the possibility of another kind of error (unfortunate use of words or conceptual inadequacy) also exists. Hence, it seems that error and disagreement are connected, in the sense that disagreement depends on the idea of error, but not the other

[17] Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 13.

way around. If an idea of error or mistake did not exist (meaning we had no distinction between truth and falsity), there would not be much sense in an idea of disagreement.

Anyway, this second suggestion is closely related to the suggestion of introducing a fourth "participant" to the triangle. The motivation for this suggestion is that we need to explain how a learner can get his idea of error. Without an idea of the possibility of error, the learner cannot yet be ascribed propositional attitudes (even if more than awareness of error might be needed). As long as the attitude of "holding true" is of such crucial importance in interpretation, and as this again depends on a distinction between belief and truth (so-called "second-order beliefs"), it is hardly an exaggeration to say that it *is* important to be able to explain from where a learner can get the idea of the possibility of error. And I maintain that if a learner only has *one* teacher to rely on, he or she will, in general, hold the teacher infallible; and hence, the learner might understand that something he or she said was mistaken without getting a specific grip on what is going on, and without getting hold of a distinction between belief and objective truth. With two potentially disagreeing teachers present, the idea of the possibility of error will have its obvious source. Another advantage is that our theory of thought and knowledge will be empirically trustworthier.

The disadvantages of both suggestions are numerous. A fourth element makes our theory more complicated; we are bringing an additional participant into *one* of the situations we call triangulation – will we have to rename the whole ting? This objection can be easily met by assuming that one of the vertexes of the triangle represents society, and not specifically one single person. This assumption coheres with Davidson's many statements about the role of the second person, and so the second person doesn't literally have to be *one* person. Indeed, it seems that in this situation it *could not* be one person, and this might be the case also in other triangular situations. If our aim is to provide the necessary and sufficient conditions for thought and language, another serious implication is whether this suggestion supplies more than we need. Might we risk giving more than what we minimally need when it comes to our theoretical use of the teacher-learner-triangle? There is a question as to whether we could end up with an unnecessary complicated theory.

But the demand on ascribing attitudes only to creatures that are potentially able to disagree with us seems to solve some problems. We had the problem that we didn't know how we could ascertain whether the other creature possessed awareness of the possibility of error.[18] If we set the criterion to be the possibility of disagreement instead, it is a stronger demand, but easier to know whether or not it is fulfilled; the other creature has to be able to produce a disagreement. We avoid the uncertainty about whether the learner is only parroting or if she or he is in charge of the norms. That also Davidson is concerned with this question of (what he calls) "observable behavioral fallout", is rather obvious from his generally Quinean methodological ideal. A disagreement would undoubtedly be classified as observable behavioral fallout. In "Externalisms" he talks about observable behavioral fallouts and criteria for knowing whether a creature is merely aping or not, and mentions that: "...one such fallout is willingness to change one's mind (and therefore what one says), not because one differs from others, but because the change fits in with further observations."[19] This is a description of an advanced and mature language user, and if we can observe and understand another creature's change of mind, this indicates that there is a propositional content in this mind, and thus, we can assume that he is in charge of the concept of error. This criterion or fallout is then certainly not as minimal as the criterion of communication. If we ask why we need to know whether or not the other creature has an awareness of the possibility of error in the first place, one of Davidson's reasons is that "It takes two points of view to give a location to the cause of a thought, and thus to define its content".[20] Hence, determination of content is one of the reasons for the need to decide whether or not the other creature has an awareness of error.

Another, and perhaps even more underlying motive for establishing a criterion, is to be found in the insight that, according to Davidson, there is no point in saying that neither another creature's thought *nor one's own thoughts* have propositional content, unless we can communicate. "Until a base line has been established by communication with someone else, there is no point in saying one's own thoughts or words have a propositional content."[21] We also need to have knowledge of the propositional contents of our own minds in order to ascribe thoughts to others. There is, then, a reciprocal dependency

[18] This problem arises in the context of the methodology of radical interpretation, where we are supposed to start out from something non-semantic and end with something semantic.
[19] In Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 9.
[20] Donald Davidson: "Three Varieties of Knowledge", in Davidson: *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001, p. 212f.
[21] Ibid, p. 213.

between knowledge of our own minds and knowledge of the minds of others. We depend on other minds that are propositionally structured, and Davidson would grant that it is important that we can know that the other creature has an awareness of the possibility of error. But, as I see it, the problem is that we would be able to *communicate* with another creature, and he or she could still potentially lack an awareness of the possibility of error. Thus, the criterion of communication that Davidson has suggested might be too weak to satisfy our needs.

6. Concluding remarks.

Firstly, my suggestion has been to expand our triangle with a fourth "element" in the case of the triangular situation of language-acquisition. This fourth element will be another competent speaker-interpreter, a second teacher that symbolizes the society of "debating" participants. This extension is found non-threatening to the metaphor of a triangle and triangulation as such, as "the second person" is intended to symbolize society. If we decide to go for an "extended triangle", so that, in the case of language acquisition, the tripod consists of more than one person, what about the other triangular situations, especially the situation of symmetrical linguistic competence? The reliance on "the second person" symbolizes the dialectic feature of thought, and creatures in discussion and a debating society are already present in the metaphor of triangulation. The triangle of symmetric linguistic competence, as I have suggested, is already "extended" in the sense that we extended the triangle with a teacher and a learner.

The related, and second, suggestion is that the possibility of disagreement is to be counted as a criterion for ascription of thoughts to a creature. This will make it possible for us to effectively discern parroting from rule following and we will have a situation that can be easily handled when it comes to deciding whether or not a creature is a rule-follower.

Both suggestions point in the direction of a more thorough dialectical or dialogical understanding of thought and language. They point in a direction where we can see that when we communicate, we both exercise *and* form language and thought. In my opinion, this is not a contrast to Davidson's views; rather, it is a question of taking one more step in the direction Davidson was heading. What we still have *not* obtained, of course, is an account that can *explain* normativity from the outside; we are still looking at it from the perspective of the radical interpretation participants.[22]

[22] I am grateful for comments upon early versions of this article given by Johan Arnt Myrstad and Frode Sirnes Larsen. I thank the participants of the international workshop "Davidson and Triangulation", arranged by Professor Bjørn Ramberg at the University of Oslo in 2001, for comments and suggestions on an earlier version of this article. I am in particular indebted to Eivind Balsvik, Kathrin Glüer, Peter Pagin, Bjørn Ramberg, John Richard Sageng, and Robert Sinclair. I also thank Jan Harald Alnes for comments on later versions.

6. Ascription of Belief and Second Order Belief

Davidson holds that having beliefs about beliefs is a necessary condition for having the concept of belief at all. In this view, the existence of second order beliefs, or generally, higher order thoughts, is a condition for the existence of first order beliefs. In effect, this means that in order to know that a belief of mine is a belief, I must be aware of the fact that my belief might be false. In this way, the conceptual difference between 'belief' and 'truth' is something that we can get from a reflection on the concept of belief itself. Davidson even claims to have drawn a concept of *objective* truth partly from such reflections, and objective truth is therefore being constituted partly from within the sphere of intentionality. This means that by showing the internal relations between intentional concepts (such as belief, belief about belief, truth, and error) Davidson gives meaning to these concepts. Hence, according to Davidson, objective truth relies partly on our ability to think about our own thoughts. Our subjective judgment of the relative probabilities of our own beliefs thus partly constitutes objective truth.

From what or where could we have started when drawing the line from second-order beliefs to the concept of objective truth? Our starting point is a frame of interpretation, in which language is seen as a means of communication.[1] In this frame, language is about communicating thoughts, and understanding is the goal of communication. When we make an attempt to communicate some propositional content to someone, we want to be understood. Therefore, we employ whichever words (and other means) considered necessary to achieve understanding. In addressing us, we assume that the other party has a similar aim. When we try to understand the verbal behavior of others, we assume that they have some propositional content in the form of beliefs, wishes, and other attitudes that they want us to know about. Similarly, when we want to rationalize our own verbal (and other) activity, we ascribe propositional attitudes to ourselves. In short, ascription of propositional attitudes is acts of interpretation. Thus, "interpretationism" is an enlightening description

[1] I here consider communication to be distinct from language. Even if language is first and foremost a vehicle of communication, this is not to say that all there is to communication is language.

of this view of truth, knowledge, thought and language. [2] I will give a brief account of Davidson's interpretational frame, or "method".

Davidson's version of interpretationism is based on a thought experiment named "radical interpretation". The point of this exercise of thought is to establish what an interpreter needs to know (and how he could come to know this) to understand a speaker of an alien language. The radical interpreter deals with an unknown speaker and his environment, and the interpreter only has to be able to recognize the speaker's attitude of "holding true" (meaning "believing"). Based on this, the interpreter is assumed able to form a theory about the speaker's beliefs ("What are his beliefs?") and meanings ("What does his sentences mean?"). A central tool for the interpreter is the principle of charity, by which the interpreter presuppose that the speaker speaks the truth and also that the speaker is consistent. The principle enables the interpreter to hold "beliefs constant as far as possible while solving for meaning".[3] This means that the interpreter, based on his own system of beliefs and sparse evidence about the speaker's beliefs, goes on to form hypotheses about the speaker's meanings. These hypotheses about meaning again contribute to new hypotheses about the beliefs of the speaker, and so the process of interpretation goes back and forth.

Davidson assumes that the speaker and the interpreter have a common biological basis in the form of dispositions for categorizing the world in similar ways. The pillars that support the interpretational platform are hence these common dispositions, the shared physical environment, the interpreter's ability to recognize when the speaker holds a sentence true, and the interpreter's assumptions of the truthfulness and consistency of the speaker. In Davidson's approach, this process, in which the interpreter and the speaker meet in a common environment, is called "triangulation". The notion of triangulation hints at the

---

[2] Bill Child has employed this terms in his description of an approach to mental phenomena generally, where it is assumed that we can reach an understanding of the nature of the propositional attitudes "by reflection on the procedure for interpreting a subject's attitudes and language", p 1. in William Child: *Causality, Interpretation and the Mind*, Oxford University Press, 1994. For a key expression of the view, he refers to Davidson's claim that "What a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes", and, Child adds, "and desire". (Ibid) Child thus mentions Davidson, but also Dennett as representatives of this tradition, and holds that there is a strong Wittgensteinian strand to the view.

[3] Donald Davidson: "Radical Interpretation" in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 137.

three vertexes: the interpreter, the speaker, and the world. According to Davidson, the world consists of "objects, properties and events that the creatures can discriminate in perception."[4] There are casual connections between the world and the speaker, the world and the interpreter, the speaker and the interpreter, and the interpreter and the speaker. This means that the interpreter perceives the relevant aspects of the world and he perceives the speakers linguistic reactions to the relevant aspects of the world. When the interpreter then ascribes a belief to the speaker, he has correlated these two causal connections; he has been looking for relevant similarity between his own response to the world and the other speaker's response. Based on the interpreter's own system of beliefs, he interprets the speaker's utterance to have a specific meaning and ascribes a certain belief to the speaker.[5] This means that a belief (and belief content), as well as a meaning of a sentence, according to this version of interpretationism, is constituted in triangulation.[6] In this way, the basis for this constitution is social, or specifically, intersubjective, and as such, the process of triangulation is considered to generate hypotheses about beliefs that stem directly from the world that they are about. In this sense, belief content is objective, and objective truth, then, partly dependent on triangulation in interpretation. But this external side of the constitution of a notion of objective truth will not be the focus of this article; the subject is second order belief and the possibility of an internal constitution of a notion of objective truth on basis of this.

An interpretational frame for a study of belief and knowledge is by no means self-evident. I will, however, not discuss this frame in relation to other alternatives here. And likewise, without discussion, I assume that interpretation means ascription of attitudes. What is required for having propositional attitudes? To whom can we ascribe attitudes? I start by presenting Davidson's four criteria for ascription of attitudes, and thus for interpretation, and I discuss some of my observations of related problems. I pay particular attention to

[4] Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 4.
[5] This will typically go both ways, such that both creatures in triangulation are speaker-interpreters.
[6] The mentioned factors of triangulation: coherentism, causality through social externalism and perceptual externalism, and an evolutionary biological basis, are all seen as elements in a unified approach. I have argued in article 3 that this is Davidson's view, and it seems to me that many of his commentators have overlooked someone or other of these factors. The social and the biological have in most of the history of philosophy been seen as opposing areas, as in e.g. asking about whether there is a cultural OR a natural basis for action. Likewise, the causal and the intentional have been considered separate, as in the Kantian distinction between the realm of necessities and the realm of freedom, respectively.

Davidson's notion of belief. Next, I discuss in detail the status of second-order beliefs (which is one of the criteria) within Davidson's theory. We need to question the 'foundational' merit of second-order beliefs in this respect as the notion is essential to our understanding of intentionality, especially in a process of co-constituting a notion of objective truth from the inside; indeed, we could even claim that the notion of objective truth hinges on it. The existence of second order belief as a presupposition for having beliefs has been challenged by research on autism. It has been suggested that some autistic speakers represent a counter example to some of Davidson's claims. I argue against this suggestion and I discuss the consequences of considering, or not considering, second order beliefs a prerequisite. Finally, I briefly consider the question of a possible tension between interpretationism and naturalism in epistemology.

1. Propositional attitudes comes in sets

In "Rational Animals" Davidson's first criterion for ascription of the attitude of belief is that for us to be able to "...intelligibly ascribe single beliefs to [somebody], we must be able to imagine how we would decide whether [this somebody] has many other beliefs of the kind necessary for making sense of the first."[7] Thus, a first criterion for ascription of a propositional attitude is the criterion that a creature can intelligibly be ascribed an attitude *as one attitude* in a *set* of such attitudes. The reason for this is that the propositional attitudes according to Davidson "...come only as a matched set"[8] Such a set of attitudes is (partly) what makes it possible for anyone to individuate, and therewith, identify, a particular attitude. "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."[9] I believe that this point is fairly straightforward, at the least as a methodological principle.

As Davidson suggests towards the end of "Rational Animals", in which he mentions the metaphor of triangulation for the first time, and as he goes on to argue more carefully in "A

[7] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 475, 'somebody' is replacing 'a dog'/'the dog'

[8] Ibid. p. 473.

Coherence Theory of Truth and Knowledge", the objects and events that cause a particular belief, or any other attitude, is the other factor that plays a part in identifying this particular attitude. It is therefore the third point in a triangle (where we; another creature and I, are located at the two other points) that, together with an individuation of a belief as placed in a network of beliefs, can identify a particular attitude. This bipartite dependence on a set of attitudes (holism) and causally effective objects and events (causality) for the identification of a particular attitude is of crucial importance to Davidson's theory. None of these parts can be missing, even if we sometimes are close up to, sometimes far from being actually confronted with the whole system of attitudes as well as with the actual cause, insofar as that is possible.[10] Anyway, if we overlook either of these parts, the result will be a misreading of Davidson. I have elsewhere argued that McDowell has presented us with one such misreading in his *Mind and World*, where he sees Davidson's philosophy as coherentism at the risk of being out of touch with reality.[11]

But what is the actual argument behind this first criterion; that attitudes come only in matched sets? There are at least two different claims involved here; first, that we cannot *know* that another creature has one single particular propositional attitude at all unless he has several, and second, that the other cannot *have* one single particular propositional attitude, unless he has several. The effect is in both cases that he and we cannot *identify* a singular particular attitude in him unless he has several attitudes. This could be taken to mean that Davidson is claiming that what can be known for an interpreter is all there can *be*, while only having argumentative resources for the first, epistemic claim (that we cannot have access to a single particular attitude unless the creature has several), and not for the second, ontological claim. Holism about propositional attitudes as an epistemological presumption is not the same as holism about attitudes as an ontological presumption. I would suggest that we restrict ourselves to the epistemic claim from a third person point of view, as our argument here stretches no further. This amounts to a methodologically based restriction, and I consider this unproblematic and non-controversial.

---

[9] Ibid, p. 475.
[10] When I say, "insofar as that is possible", I think of the claim that, in many cases, it is not possible to really say that one object is the cause of a belief. Rather, several objects and events or even the whole causal chain must be said to be responsible for causing a belief.

2. Beliefs occupy a certain position among the attitudes

Next, Davidson argues that since "...propositional attitudes require a background of beliefs..."[12] he holds that "Without belief there are no other propositional attitudes..."[13] and he focuses on beliefs when stating conditions for the existence of thoughts in general. This means that the existence of beliefs in a creature is a general condition for ascription of *any* propositional attitude to the creature. Davidson thus assumes that beliefs are prior to the other attitudes in some kind of way. If we think of beliefs in relation to desires, denials, intentions, hopes, fears and wishes, we would have to ask: what is it that makes beliefs so special? Let's grapple this by asking what a belief is according to Davidson.

A belief is a sentence held true by someone who understands it. [14]
Beliefs are states in creatures with intentions, desires, wishes and sense organs.[15]
Beliefs are not internal states that only the person that has them can know about. If beliefs didn't have an external basis, we could never know, and not even make sense of, that another person held a particular belief.
Beliefs are states that are caused by internal and external perceptions[16], but these perceptions can never be evidence for the beliefs. A causal explanation can never show that or why a belief is justified.[17]
Beliefs are states that cause events inside and outside the bodies of their entertainers.[18]
Beliefs come in sets, not one by one, or else they couldn't be identified.[19] Radical incoherence in a set of beliefs is impossible because of internal dependencies among the beliefs.[20]

---

[11] In article 2 I argue that the reason for this misconstruction is Mc Dowell's insensitivity to the importance of the social constitution of knowledge, via the social location of the objects and events that causes our beliefs. See article 2 and John McDowell: *Mind and World*, Harvard University Press, Cambridge, Mass. 1994.
[12] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 478.
[13] Ibid.
[14] Donald Davidson: "A Coherence Theory of Truth and Knowledge" in *Subjective, Intersubjective, Objective*, Oxford 2001, p. 138. Or also, sentences held true are "the linguistic representatives of belief" ("The Method of Truth in Metaphysics", (1977) pp 199-214 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984, p. 201)
[15] "A Coherence Theory of Truth and Knowledge", p. 138 in *Subjective, Intersubjective, Objective*, Oxford 2001
[16] Ibid.
[17] Ibid, p. 143.
[18] Ibid, p. 138.
[19] Donald Davidson: "Rational Animals", pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 475.

A belief is given its content by one's knowledge of what is required for the belief to be true. Since beliefs and truth are related in this way, belief can serve as the human attitude that connects a theory of truth to human concerns.[21]

This list, which is meant to be relevant to our questions and not exhaustive, gives us more than a hint yet not a complete overview of what a Davidsonian belief is. The notion of belief is different from "platonist" conceptions of notions like proposition, statement, judgment, mental content or propositional content. When entertaining a platonist notion of proposition one assumes that the content of each proposition is objective and eternal, and that we display this content (or a "copy" of it) when we utter an instantiation of it in the form of a grammatical unit. The point here is that the propositional content (the content of the statement, or the mental content) is seen as fixed once and for all, and it is considered more basic than the "instantiations" of it, so that one may ask whether or not it has been correctly represented by the instantiations. Davidsonian beliefs are definitely nothing like this. Indeed, the contents of beliefs can be instantiated in a variety of sentences, but there is no common stock of fixed beliefs in which we can find the authoritative versions for comparison. As the radical interpreter has his own system of beliefs to rely on in interpretation, this is the closest we come to a store. But it is not authoritative in the sense of being universal or eternal. But, on the other hand, we *can* assume that *most* of the interpreter's beliefs are true, due to the way in which the beliefs were constituted. According to Davidson, beliefs are being constituted intersubjectively, based on the common physical surroundings. But beliefs are still hypothetical. A belief about the world is a hypothesis about the world, and the hypothetical relation to the world is a characteristic feature of Davidson's epistemological thinking, a part of his Quinean heritage. One of the otherwise most important differences between the notion of proposition of the analytic tradition and the notion of belief addressed here is that the belief is always related to a

[20] Ibid, and in "Afterthoughts", p. 155 in *Subjective, Intersubjective, Objective*, Oxford 2001. Or, as Davidson puts it in a reply to Carol Rovane in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999: "Thoughts belongs to the domain of rationality. This is because thoughts () have, as we say, propositional contents. They therefore have logical relations. We cannot entertain any arbitrary collection of thoughts; logic constrains us. This… is due to the fact that the company in part determines the contents of thoughts they keep, and enough of this company must be logically congenial to confer a particular content on a thought. This is the sense in which I have meant that rationality is constitutive of thought." (p. 480)

[21] Donald Davidson: "Epistemology and Truth" (1988) pp 177-191 in *Subjective, Intersubjective, Objective*, Oxford 2001, p. 189.

believer and his system of beliefs. The notion of belief is therefore, in a way, relational.[22] We are personally connected to the belief-contents that we entertain. Our different beliefs have a specific subjective probability for us, and our attitudes to our own beliefs can vary a lot. However, some of our perceptual beliefs about the most familiar middle-sized objects usually have less room for probability variation than many of our other beliefs. A personal relation to a belief-content is necessary for something to be a belief; a person has to both understand the belief-content and hold it true, for something to *be* a belief. It is yet not private, as it has a linguistic representative; it has a sentence as its representative. Hence, a belief is propositionally structured, and it is publicly accessible.

Davidson's characteristics of beliefs constitute a truly amazing achievement as he manages to bridge the presumed gap between "subjective minds" and "objective world". This is done in such a way that the subjectivity of mind and belief as mental content is preserved; while at the same time the objectivity of the world and the objectivity of the belief as sentence is intact. This analysis also provides a demonstration of the claim that we are directly influenced by the world;[23] there is no filter between linguistic minds and non-linguistic world.[24] Yet, according to Davidson, language is the "...main entry into the mind". [25] This suggests that belief is of crucial importance to interpretationism as it operates at the joints of subjects, language and world.

What would be different if we chose to let, say, the "propositions" or the "contents" of other people rather than their beliefs be the feature that we were trying to get a hold of when aiming at understanding? When we aim to understand other people in ordinary attribution of intentional states, we are after the person's own attitudes, that is, his attitudes as he holds them from his own perspective. The subjective relation of probability that a speaker has towards his own attitudes would be lost if we chose to focus on a Platonist notion of proposition instead of the notion of belief. And then we, as interpreters, would

[22] Jan Harald Alnes made me aware of this difference between the proposition and the belief, and I have adopted his characterization of the belief.
[23] See e.g. Davidson's "Seeing Through Language", p. 20, in *Philosophy: the Journal of the Royal Institute of Philosophy*, Cambridge 1997, "...language is not something that comes between us and reality".
[24] Besides, we should note that society and we ourselves exercises a pressure on our belief-system, through constant corrections that more or less lead to revisions. This can explain why conversation is so favorable to us. Second-order belief is important in order to be able to revise our belief-systems.
[25] Comment in seminar, UCB, April 31. 2000.

also loose what we were after. What we *are* after is to understand the other. We want to understand his or her interests, actions and orientation, and it is then inevitable that we take his or her beliefs into account, because beliefs are relational in the sense that they need a subject. A person's own assessment of the belief-contents is constitutive of what a belief is.

Now, let's take a closer look at the background for saying that beliefs are special *among the attitudes*. Davidson holds that all the other propositional attitudes depend for their "particularity on a similar world of beliefs. In order to believe the cat went up the oak tree, I must have many true beliefs about cats and trees, this cat and this tree, the place, appearance and habits of cats and trees, and so on; but the same holds if I wonder whether the cat went up the oak tree, fear that it did, hope that it did, wish that it had, or intend to make it do so. Belief – indeed, true belief – plays a central role among the propositional attitudes."[26] It is apparent that Davidson's focus is on *true* belief as particularly central. But this is still not an explanation, nor even a rationalization, of why beliefs are special *among* the attitudes. We might think that what he says in this quote is not very different from saying that truth plays a special role, and that is not unreasonable, but it's not what we are after here. We should notice that Davidson also says: "...beliefs demand other basic attitudes such as intentions, desires..."[27] and so it seems as though the dependence-relation between beliefs and the other attitudes goes both ways. Later in "Rational Animals", he repeats the point that beliefs are special among the propositional attitudes: "I think I have shown that all the propositional attitudes require a background of beliefs..."[28] But all that he has given in this article are the aforementioned quotes, and they contain the statement that there is a particular dependency, but no argument for it. And so, even though I think that Davidson is right, and even though we know that according to Davidson the ability to recognize the attitude of holding true is vital for the radical interpreter and thus that the belief in this sense is the starting point for radical interpretation, it doesn't seem that Davidson here has shown either *that* or *why* "all the propositional attitudes require a background of belief". At the least he has not shown this any more than he has shown that beliefs also require a background of other propositional attitudes.

[26] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 475.
[27] Ibid, p. 473.
[28] Ibid, p. 478.

One of the reasons that the attitude of belief is of such unique standing to Davidson is that it is the notion that takes the place of the fregean thought or conceptual content, and the place of the russellian proposition, and even the place of the Quinean (in comparison to these two, very de-metaphysicalized notion of) sentence. It is also reasonable to claim that the notions of assertion and predication have its "replacement" in Davidson's notion of belief. The reason for the focus on such a notion in the first place is deeply connected to our human interest in the world, our human interest in other humans and to our human interest in truth. The thought, proposition, statement, assertion or predication has been considered subject to truth validation or rejection. Moreover, it is assumed that a person who entertains these entities considers them true. The prevailing entity has been deemed to be about the world, and has been considered to have the potential of being either true or false. The evaluative force of the entity is considered the reason for its potential to have a given truth-value. The entity is often considered to be a judgment, and in its typical form it is an indicative sentence, a description.

Davidson relates to the tradition within analytic philosophy, and seems concerned with keeping up the dialogue, perhaps unlike Quine. Quine's break with the tradition is more radical, and he is happy to confine himself to the sentence. Davidson seems to be after the same as was Frege with his thought and Russell with his proposition. According to Davidson's theory, a belief can be true or false. When somebody entertains it in the form of an assertion, he or she represents him or herself as believing it, and hence assumes it true. If one knows the belief's truth requirements, one knows the belief's content. The belief's truth conditions are thus the conditions under which we can recognize the belief, but it is also required that the belief is a part of a network of beliefs in order for us to identify it. Within radical interpretation, identification of belief is, and has to be, accompanied by identification of meaning, and we need to solve for both at the same time.[29] Davidson's theory of meaning for a speaker is based on assent to sentences that are caused by events in the world, and the theory of meaning is in this way also a theory of belief for the speaker. This is so because when we can interpret a speaker's sentence that he, based on an intersubjectively available event holds true, then we have identified one of his beliefs.

[29] I have elsewhere given an account of the early phases of radical interpretation and the weighing of belief and meaning by use of the principle of charity, see article 3.

The notion of belief was obviously a key notion in Davidson's early theory of interpretation, and still is, even if the notions of intention, desire, wish and the other attitudes are of equal standing in his unified theory of language, action and value. After the extension of the theory of interpretation, it would be reasonable to confine to the claim that a creature that possesses one kind of propositional attitude also possesses other kinds of attitudes. However, it is obvious that *true belief* plays a very central role among the attitudes. It is true belief that connects us to the world. I also happily succumb to the claim that *perceptual belief* plays a particular role in connection to this.

3. Second order beliefs

The third criterion for interpretation is a criterion for having beliefs, and it says "...in order to have a belief, it is necessary to have the concept of belief."[30] Having second order beliefs, according to Davidson, is thus (as noted in the introduction of this article) a criterion for having the concept of a belief, which again is a criterion for being attributed the ability to think and for having language in general. Davidson's claim is that in order to be a "qualified" speaker, you have to understand what it means to have a belief, and basically, I think Davidson is right in holding that the ability to have beliefs about beliefs normally will be co-occurring with the ability to speak a language. Research on autism, however, suggests the existence of speakers who are unable to ascribe false beliefs to themselves and other people.[31]

It is a common trait for autistic individuals that they take all sentences to express literal meaning (they cannot, for instance, conceive linguistic jokes or irony), and when tested in so-called false belief tests, the majority of persons with autism fail to pass the tests.[32] The

[30] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 478.
[31] In my account of this research I base myself mainly on Kathrin Glüer and Peter Pagin's "Meaning Theory and Autistic Speakers" in *Mind and Language*, Vol. 18, No. 1, pp. 23-51, 2003. Not only do they give an impressive overview of the research, they also discuss some of the central implications for philosophy of mind and language.
[32] One of the standard tests is a change-in-location false-belief task, which is often called the Sally-Anne test. A child is presented with the puppets Sally and Anne, where Sally has a basket and Anne has a box. Sally puts a marble into her basket and she then leaves the room. Then Anne takes the marble from the basket and puts it into her own box. When Sally comes back she wants her marble. The child will now be asked where Sally think her marble is, and also where Sally will look for it. If the child then answers that Sally will think the

tests are designed to find out whether the test-person understands that a belief (of him- or herself or another person) can be false. The reason why false belief has been of interest in cognitive psychology is that they are crucial in testing so called so-called "theory of mind theories" or "mindblindness theories" of autism. It has generally been assumed that the tests will be vital in perceiving how much of the autistic person's linguistic and communicative deficits can be understood and explained by the theory, and it is assumed that false-belief tests indicate the ability to "read" the minds of others. [33] In this connection Kathrin Glüer and Peter Pagin point out that: "To understand the difference between being true and being believed to be true, one must understand that a belief can be false, and this understanding is manifested by means of the ability to ascribe beliefs one takes to be false."[34] The fact that persons with autism consistently fail the tests is considered an indication that autistic individuals lack, or have a seriously impaired, capacity for thinking about thoughts. For our particular context, this also seems to provide empirical evidence that suggests an impaired capacity to understand what it is to have a belief (and other intentional concepts) in these speakers. Hence, they do not have "the concept of belief" in Davidson's sense. Autistic individuals do have general communication problems, or rather; a problem with communication is a key diagnostic criterion for autism.[35] This involves poor interpretational abilities and the autistic individual is generally unable to see another person as an intentionally determined subject. They have problems taking the other person's point of view; particularly, when a speaker's utterances are not meant literally, they are unable to take clues from the context to interpret the utterances. Yet, some of these individuals are able to express themselves creatively; they have generative capacity, which means that they can produce and understand new sentences, they are able to communicate by means of language and will by all means qualify as speakers. Given that the description

---

marble is in her basket, or also that Sally will look in her basket, this will be the correct answer. This answer will involve ascription of a false belief to Sally and a prediction that Sally will act on the false belief. According to acknowledged tests 85 % of non-autistic 4-year-olds pass the test, while only 20 % of children with autism do. This account of the test leans heavily on Glüer and Pagin, Ibid, note 8, p. 27, and Janet W. Astington and Jennifer M. Jenkins: "A Longitudinal Study of the Relation Between Language and Theory-of-Mind Development", in *Developmental Psychology*, Vol. 35, No. 5, pp. 1311-1320, 1999.

[33] The tendency in newer literature on false belief tests and theory of mind is that there is no reason to identify a failed false belief-test with a lack of a theory of mind. One might have an otherwise well developed reading of other minds and still fail the test, for several other reasons.

[34] Kathrin Glüer and Peter Pagin "Meaning Theory and Autistic Speakers" in *Mind and Language*, Vol. 18, No. 1, pp. 23-51, 2003, p. 27.

[35] Uta Frith and Francesca Happé in "Language and communication in autistic disorders" in *Philosophical Transactions of the Royal Society*, Series B, 346, pp 97-104, London, 1994.

of this group of autistic speakers so far is correct, the question for us is whether or not they represent a counter-example to Davidson's theory. But before we examine this matter in detail, let's have a brief look at Davidson's last criterion for interpretation.

4. Language, the ultimate criterion?

Davidson argues that "in order to have the concept of belief one must have language."[36] This is thus his fourth criterion for ascription of propositional attitudes to somebody. Davidson admits that he doesn't have a waterproof argument for his claim that language is the ultimate criterion for distinguishing between creatures with thoughts and creatures without thoughts. He says that all he has is some basis for a suggestion that "...there probably can't be much thought without language"[37] and I take it that the term "thought" should (or at the least could) be understood as "propositional thought". Given the interpretational frame this could be read as the modest claim that we cannot *grasp* or have access to the thoughts of another creature unless he or she has language, or it could be read as an empirical or ontological claim. Either way, it is probably fair to assume that propositional thought demands language. What seems to be beyond doubt is that the fairly sophisticated concept of belief that Davidson is talking about actually could be found only in creatures with a full-fledged language. As I have done before, I would stick to the more modest epistemological or methodological version of interpretationism in this particular situation, and interpret Davidson's fourth criterion as saying that we cannot *have access to* the propositional thoughts of another creature unless he or she has language. Hence, I would not read his claim as a claim about the ontology of mind. But now, let's go into the question of the merits of second order belief, and the proposed empirical refutation of the third criterion for interpretation.

5. Beliefs Without Second Order Beliefs?

Glüer and Pagin argue that speakers who lack the capacity of thinking about thoughts are counter-examples to some of Davidson's claims. In particular, they argue that Davidson's

[36] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 478.
[37] Ibid, p. 477.

claim that "one cannot have beliefs at all unless one has the concept of objective truth" has its counterexample in a group of autistic speakers. In the same connection they also mention the Davidsonian claim about second order belief that "no one believes anything who does not grasp the concept of belief'.[38] When they point out that there are persons with autism who "combine a sufficient degree of linguistic ability to be regarded as language users with a seriously impaired capacity for higher-order thoughts"[39] Glüer and Pagin seems to take 'sufficient degree of linguistic ability to be regarded as language users' and Davidson's utterances about 'having beliefs' to be interchangeable terms here. And that may well be correct. We do have quite detailed information about Davidson's idea of what it is to have beliefs. And having beliefs in Davidson's sense certainly demands "a sufficient degree of linguistic ability to be regarded as (a) language user(s)". Davidson has also gone far in the direction of reversing the order by admitting that "communication in the full linguistic sense" will suffice as a basis for belief.[40] But the question is whether Glüer and Pagin's concept of belief and the concept of belief involved in empirical psychology is the same as Davidson's concept.

Davidson's concept of belief is not just any concept of belief. The concept is defined in relation to a cluster of other intentional concepts,[41] and it is also marked by its role within the methodology of radical interpretation. Therefore, this concept of belief is to some extent technical, or quasi-technical. In Davidson's sense, a belief has a propositional structure and can be stated in the form of a sentence. A belief can only be found in creatures with other identifiable propositional attitudes like intentions and desires and wishes and other beliefs, and to whom we can ascribe *an awareness of* the possibility of false belief. Beliefs in Davidson's sense are caused by perceptions and they are themselves causally effective. Beliefs can be justified only by other beliefs, not by something in the world, and the contents of beliefs are given by its truth conditions. But, and this vital: beliefs are hypothetical. I said earlier that a belief about the world is a hypothesis about the world. As such, we do not take all of our beliefs to be equally probable. We frequently

---

[38] Kathrin Glüer and Peter Pagin's "Meaning Theory and Autistic Speakers" in *Mind and Language*, Vol. 18, No. 1, pp. 23-51, 2003, note 5, p. 25.
[39] Ibid, p. 25.
[40] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 480.

assess and re-assess our attitudes towards our beliefs. It is apparent that Davidson's concept of belief is a distinctly theoretical term, and, on a number of issues, non-neutral. This is not to say that the characteristic features that we in this case associate with belief is incompatible with an innocent (what ever that may be) or non-technical notion of belief, but the features in question are certainly more specific and specified than what we would expect of a definition of the notion in a dictionary.[42]

Now, we could question the methodology of radical interpretation and refuse to accept an interpretationist notion of belief. Even if we would question the particular methodology of radical interpretation, and the investments in the concept of belief following from it, I would hold that some of the characteristic features of the concept, as Davidson presents it, cover common intuitions. In particular, I am thinking about the insight that an explanation of the notion of 'belief' that does *not* involve a contrast to the notion of 'truth' is inadequate. I will argue that anyone who reflects on the notion of belief will come to contrast it with a notion of truth. I think that this part of Davidson's notion of belief (which is one of the traits of the notion stemming from participation in the cluster of inter-defined intentional concepts) is widely supported.

On the other hand, we could choose to say, more bluntly, that any speaker who does not have second-order belief is not a full-fledged speaker, that he is not a speaker in Davidson's meaning of the word or that he hasn't got a language in the strict sense, or in Davidson's sense. On this basis we could refuse to accept the autistic speaker as a proper speaker. This would, however, come close to saying that our theory or thesis is immune to empirical counterexamples. But then it could be feared that our philosophical analysis of the phenomenon of belief only covers a concept of belief relevant to a closed circuit of intentional concepts. A conceptual analysis of the relations within such a circuit will by no means be uninteresting or irrelevant. But the empirical phenomenon of belief might then have to be left to the empirical psychologists and others to study. In that case, we would

[41] Jan Harald Alnes has persistently argued this point to me, and it took some time before I realized its impact on false-belief-tests as possible counter-examples to Davidson's thesis.

[42] Similarly is there a reason to assume that the concept of belief employed in empirical psychology will have some theoretical ballast and it is also reasonable to assume that the concept involved in the tests is marked by its role in the experiment situation. I am however not sufficiently familiar with the literature to give an account of the exact features of that concept here.

have to ask in what sense Davidson's study of knowledge, and an interpretational approach in epistemology in general, could be a specimen of naturalized epistemology. Radical Interpretation is admittedly a thought-experiment, and a conceptual exercise, but a claim to empirical immunity could make it a less interesting alternative as a frame for an understanding of human belief and knowledge.

According to Glüer and Pagin the question of whether there in fact are speakers of the kind in question is an empirical question.[43] However, they admit that it is "not straightforwardly empirical, since some conceptual care is needed in the evaluation of the evidence". I find it a bit surprising that they think of "conceptual care" only "in the evaluation of evidence" and not also in the involved key concepts themselves. In fact, they refrain from discussing the concept of belief, or the possible differences in the concepts of belief employed by the different philosophers and psychologists involved in their article. The reason might be that they want to keep the field together in this way, and that they assume the concept of belief to be the same throughout, in order to be able to show the effects of this move. But they don't explicitly say that this is what they aspire to do.

Could autistic persons become competent speakers by learning to speak in an environment with only autistic people? If this much could have been demonstrated, we would, perhaps, have a convincing counter-example to the thesis that the existence of second order belief is a criterion for ascription of propositional attitudes to somebody. If non-autistic speakers have introduced autistic persons to language (which I presume is the more common case) the autistic individual certainly will have profited from the non-autistic speaker's capacity for higher order thought. Is it likely that the autistic person could be a speaker at all if this was not the case? Research on autistic disorders does show that well-enough-qualified speakers that don't have higher order thoughts exist. However, this is not the same as showing that speakers in general do not need this capability for there to be "ordinary" linguistic activity going on. Another way of phrasing this point is to say that the deviant linguistic behavior (as shown by some autistic persons) is parasitic upon "ordinary" linguistic behavior, i.e. behavior that presupposes and includes second-order beliefs. Of

[43] Kathrin Glüer and Peter Pagin: "Meaning Theory and Autistic Speakers" in *Mind and Language*, Vol. 18, No. 1, pp. 23-51, 2003, p. 25.

course this is true also of non-autistic children up to a certain stage. The difference between some autistic persons and non-autistic kids will, however, be that non-autistic kids will develop second-order beliefs if their language learning proceeds in ordinary fashion, while it is a deviance (or hurdle) that some autistic persons never overcome.

Glüer and Pagin also raise the question of the possibility of "parasitic membership" for some members in a linguistic community.[44] They raise it in connection with Lewis' conventionalism in meaning theory. They hold that finding linguistic communities consisting only of subjects with autism is highly unlikely, and that it might be empirically correct "that a community of only persons with autism could not develop language from a pre-linguistic state."[45] However, they argue that if it could be possible that only autistic members of the linguistic community survive, and we assume that they continue to use language, the lack of counterexamples will be a "mere contingency". But I will hold that the background for, and the way in which humans has learned to speak is not a mere contingency and not irrelevant to the question. The conditions for development of language, and the question of how a person has come to be a speaker also matters in Davidson's theory. And this is not first of all a metaphysical question, but a biological and causal one. If it is true that "a community of only persons with autism could not develop language from a pre-linguistic state", then we probably cannot reconstruct such a community's ways into language and thought in the same manner as has been done for communities of non-autistic speakers by Davidson.

Davidson has pointed out three factors that are necessary conditions for language and thought.[46] These factors are dispositions, normality and error. It is a common trait for humans as biological creatures to be disposed to categorize the world and thus to discriminate in perception. That something like the normal cause for a thought has come to exist is the basis for establishing a notion of normality. When two creatures respond to objects and events, so that a common cause of the most frequent cases have been established, errors can occur. The three factors are claimed to be necessary factors;

[44] Ibid, p. 43, note 29.
[45] Ibid.
[46] In Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, page 9.

however, they are not claimed to be sufficient for development of language. In order to acquire language, Davidson assumes that we also have to be raised in a society of language users, where ostensive learning plays an important role. As individual speakers, we share a common biological background with other humans and a personal, social and causal history of language learning. In particular, we should note that interpretation of other speakers is a part of the establishment of the conditions of normality and error. If a creature were unable to evaluate a possible relevant similarity of the linguistic reactions of himself and the other person in triangulation, and that he thus was disabled to ascribe a belief to a speaker on the basis of correlated causal connections and his own system of beliefs, no normality could be established. Interpretation requires personal assessment of the belief-contents of oneself and others. Hence, it is doubtful that the conditions for language and thought that Davidson has pointed out would be an acceptable reconstruction of the conditions for language and thought for a linguistic community of only persons with autism. My point here is that language and thought cannot be disconnected from its biological and causal constitution and from the way in which we learn to use language in the first place. We cannot just take away the learning environment and the causal history of language learning for the linguistic community of autistic survivors, and then use the remains as a counter example. As long as it is highly improbable that an autistic linguistic community could develop language from a pre-linguistic state, the mere accidental existence of a survivor-group of autistic speakers, raised in a community with non-autistic speakers could not be a convincing counterexample to Davidson's theory.

6. Concluding Discussion

Davidson's thesis is that unless a person has the concept of belief (second order beliefs), he or she doesn't have a belief. This view faces the challenge that actual people can be interpreted as having beliefs – they seem to be in the mental state that the concept belief is supposed to be the concept of – without having second order beliefs. The challenge could be read as a claim that the conceptual analysis that Davidson has based his thesis upon contradicts physical reality, and is refuted for this reason.

There are several problems here: first, a *conceptual* analysis, where we show that certain concepts stand in specific relations to certain other concepts, cannot be straightforwardly refuted by *empirical* evidence, per se. The reason is that the meaning of the concepts will be settled in relation to each other, and besides, they will not necessarily have empirical counterparts in the world that they correspond to. Even so, if we assume that the intentional concepts *do* have empirical counterparts, it is still not obvious that we can have access to them in any other way but through interpretation and we will in this way have access to the concepts as intentional concepts. On the other hand, a conceptual analysis that has empirical implications can in principle be proven irrelevant for a given empirical phenomenon, as has been held to be the case here. The conceptual analysis could be shown to give a faulty description of some particular state (or relations between states) of an organism.

As Pagin and Glüer argue, one specific state-type in a person (belief) does *not* necessarily and always presuppose another state-type (second order belief) in that person. We can assume that this is correct, but it is still an open question what Davidson's claims really amount to, and thus, whether this represents a genuine counter-example. On the one hand he says things like: "I merely describe a feature of certain concepts."[47] This statement justifies an interpretation of Davidson as "merely" analyzing concepts, and seems to suggest that his analysis of belief is immune to empirical research on the area. On the other hand he says: "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."[48] That Davidson sees the concept of belief as the concept of a state of an organism could indicate that he is after more than *just* a conceptual analysis. When he talks about states of organisms, it seems clear that Davidson seeks an empirical phenomenon of belief. Based on an epistemological interpretation of his thesis, it seems that we can still keep Davidson's conceptual analysis of the relation between the concepts of belief, second order belief and truth. But, we should be careful to think that we then address the same notion of belief that is being employed in a range of empirical studies within psychology. We should be even more cautious in our thinking about the belief when it comes to the belief as an ontological entity and the belief

[47] Donald Davidson: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985, p. 473.

(or any other type of attitude) as a brain state. This would, of course, also apply to second-order beliefs. On the other hand, if we do not consider empirical issues as even potentially serious objections we could of course defend the claims about second order belief a priori. But this can hardly be tempting for a naturalist.

I have given two arguments against reading the results of false-belief tests as counter-examples to Davidson's thesis that second-order belief is required for having beliefs. The first is that Davidson's concept of belief, defined as it is from within a cluster of intentional notions, is not the same concept as the concepts employed in the studies in question. Even though I think that I have demonstrated that this argument refutes Glüer and Pagin's claim that autistic speakers constitute a counter-example to Davidson's claims[49] when it comes to an epistemic or methodological version of Davidson's claim (that second order belief is necessary for belief), an ontological version of the claim has demonstrably not been refuted. Whether or not Davidson has held an ontological version of the claim, and whether it could be defended, will not be further discussed here. The second argument against the case of autistic speakers as a counter example is that the belief-like concepts and states of autistic individuals are parasitic on the language of non-autistic individuals. Therefore, I hold that autistic speakers do not constitute a genuine counter example, since it must be assumed that autistic speakers can only acquire language in a non-autistic linguistic environment, or in an environment in which a majority of the speakers have second order beliefs. I argue that the reason for this is that interpretational abilities are a condition for thought and language.

For Radical Interpretation as a method, and an interpretational approach to language and knowledge, it is inevitable that we uphold second order beliefs as criteria for ascription of prepositional attitudes. If we must give up on second order belief as criteria for interpretation, no interpretation in the sense that is described in Davidson's approach will be possible, since we need to assess both our own beliefs and the beliefs of other speakers in order to interpret in accordance with the principle of charity. The principle of charity

---

[48] Ibid.

[49] They mention Davidson's claim that "one cannot have beliefs at all unless one has the concept of objective truth" and the claim that "no one believes anything that does not grasp the concept of belief". (In Glüer and Pagin 2003, note 5, p. 25)

presupposes that the interpreter is able to evaluate beliefs based on causal impact from the world and hypotheses about sentence meanings. The weighing of belief against meaning and meaning against belief obviously presuppose the ability of assessing beliefs. Assessing beliefs means to have beliefs about beliefs. This means that second order belief is interwoven with this version of interpretationism to such a great extent that we simply cannot do without it. Radical interpretation cannot be construed without second order belief and still make sense. Furthermore, our ability to question and assess our own beliefs is vital for the ability to see and uphold the difference between belief and truth. If we loose the distinction between truth and belief we will loose the opportunity to have an internally based constitution (or co-constitution) of the concept of objective truth.[50]

[50] Parts of the material for this article were presented in my paper "Intensjonalitet og naturalisme", ("Intentionality and Naturalism") which I read in the dr. art. / Research-seminar at the Department of Philosophy at The University of Tromsø, November 2000. I thank the audience at that occasion for critical comments and constructive suggestions. Jan Harald Alnes has commented upon several versions of this article, and I thank him for discussions, suggestions and knowledgeable advice.

Synthesis[1]

Throughout the work on the articles above, I have often run two different situations where I have had to choose between two alternative ways of interpreting Davidson's philosophy. One of these situations is the one where I had the choice of interpreting Davidson into a transcendental-philosophical vein or interpreting him as a naturalistic philosopher. The major theme of this synthesis is therefore the relation between the naturalizing tendency that we find in Davidson's philosophy and a possible transcendental ambition. Over time, I have come to see Davidson's project as naturalized rather than transcendental-philosophical, as I have found that one inevitably needs a naturalist reading to achieve a proper overall understanding of his project. The principle of charity is one factor in Davidson's theory that I had suspected would be a challenge to such a reading, but I found that it didn't. Quite to the contrary, I realized that it is impossible to ascribe the principle a status compatible with a transcendental approach. The problem of error is another factor that has turned out to be important in the question of an overall reading of Davidson as a naturalist philosopher. I have argued that the problem of error stems from an inner discrepancy in Davidson's thinking, where it has become clear that the craving for objective truth is incongruent with the model of radical interpretation and the general naturalistic orientation found in Davidson's thinking. A third momentum for the overall analysis of Davidson's epistemological position is Davidson's reliance upon second order belief. It has turned out that second order belief within Davidson's project cannot be naturalized in the sense of ascribing it a biological basis in humans. Second order belief cannot be given a material equivalent in the human mind. In this question I have reached towards one of the outermost points of Davidsonian naturalism, and found a way in which Davidson's epistemology cannot be naturalized.

The other situation of choice is that between interpreting Davidson's philosophical theses as epistemologically and methodologically based *or* ontologically based. In this case, the methodological path has, over and over again, turned out to be the most reasonable choice.

[1] I thank Jan Harald Alnes for helpful comments on drafts of this synthesizing essay. Some of the material on naturalism has been taken from a paper entitled "Davidson, Naturalism and Constitutivity", read at an International Philosophy-seminar, "Contemporary Issues in Metaphilosophy" at the University of Tromsø in December 2001. I would like to thank the audience at that seminar for criticisms and questions, and I am in particular indebted to Kjersti Fjørtoft, Mikael Janvid, Jonathan Knowles, Anita Leirfall, Bjørn Ramberg, Richard Rorty, and Folke Tersman.

The reasons for this are given in each of the situations, but I have found a general tendency: Davidson's arguments cover only epistemological or methodological readings, and only very unlikely cover ontological readings. However, I will not further go into this second situation of choice in this synthesis.

We will take a closer look at the question of naturalism, as this is an underlying factor throughout all the articles. An account of what I mean by the term "naturalism" is therefore mandatory for discussing my overall reading of Davidson. I give a short presentation of the characteristics of Quine's naturalized epistemology, as this is the form of naturalism that Davidson relates to. Quine's naturalism and Davidson's are certainly similar, but by no means identical. Therefore, I discuss in the following how Davidson differs from Quine, and what naturalism can be within Davidson's thinking. Naturalism and studies concerned with the analysis of normativity are often considered disparate. In the question of whether we can talk about "normativity" as a phenomenon in its own right in connection with human action and language, or whether mere regularity of behavior is all there is, the reductive naturalist would say the latter. I discuss why Davidson would choose the former and how a study of normativity can find its place in Davidson's version of naturalism. Discussions of naturalism will occupy the three first chapters of this synthesis. However, the full meaning of Davidsonian naturalism will only be evident when we go into details in the summary and discussion of the articles. Chapters 4 through 9 summarize each of the articles, and in addition to highlighting some of the main points from the articles my aim is to confront each of the articles with my overall reading of Davidson's naturalism. The order in which I go through the articles reflects their importance in an evaluation of my overall reading, the more decisive first. The three remaining chapters of this Synthesis deal with the questions of objective truth and knowledge within an interpretational approach. But now, let's turn to naturalism.

1. Naturalism

Quine's 1969 article "Epistemology Naturalized" has been the "locus classicus" for the many who have tried to tie philosophical theories of knowledge up to scientific reasoning, in attempts to naturalize epistemology. The article has functioned as a research program for some of the philosophers that have later characterized their own thinking as "naturalized epistemology". Some of these theories are being worked out on basis of empirical studies in psychology, which means that the theories in question are more or less based on descriptions of how we arrive at our beliefs. However, studies of how people actually reason are often related to how one ought to reason in order to reach specific aims or fulfill specific norms of reasoning. In this sense, naturalized epistemology has not necessarily been on the whole descriptive, and as such insensitive to "the normative", even though this has been a more or less constant accusation. Anyway, most "naturalized epistemologists" would admit that "...questions of how we actually arrive at our beliefs are [] relevant questions about how we ought to arrive at our beliefs."[2] So, whether to go for a "thorough" Quinean program, popularly termed "the replacement thesis", in where epistemology is to be replaced by empirical psychology, or to go for a more mixed solution, Quine's article has been the ultimate source for naturalization of epistemology.[3]

In general, naturalism can be the view that everything that exists and everything that happens are empirically accessible traits of the world, and can be explained as such. A version hereof is the view that everything must be reducible to, or be understood via, scientific laws and theories. These views are both reductionist views in the sense that intentional phenomena must be reducible to empirical entities in order to exist or in order to be understood. Hence, we are either dealing with ontological reductionism (that all that exists is the physical or that everything must be reducible to the physical) or with methodological reductionism (that everything that exists must be explained as, or understandable as, physical), and the mentioned specific version is, in addition, scientistic. We will keep these conceptions of naturalism in mind, but since it becomes obvious that Davidson disavows reductionist versions of naturalism, we will not go into further detail on this point. However, it should be clear that naturalism and normativity have been regarded as belonging to different spheres. A naturalized theory of knowledge and a theory of knowledge ensuring the normative aspect of human intentionality are traditionally

---

[2] Hilary Kornblith: "What Is Naturalistic Epistemology?" in Kornblith, Hilary (ed): *Naturalizing Epistemology*, The MIT Press, Cambridge, Mass. 1994, pp 1-14.

[3] A replacement version of naturalized epistemology implies studies of actual belief-acquisition whereof the works of Hilary Kornblith and Stephen Stitch are examples. They are among the philosophers that see themselves as carrying out a Quinean program of naturalized epistemology, albeit different from Quine's own way of doing naturalized philosophy. These empirically oriented studies are also different from Davidson's way of doing epistemology.

considered opposites. I argue that Davidson's theory of knowledge questions the very point of a clear-cut opposition.

2. Quine's notion of naturalized epistemology

In "Epistemology Naturalized" Quine's overall aim is to show that epistemology is not a "first philosophy". He is a critic of those who think that we can give our knowledge of the world a fundament outside the physical world itself. Descartes would be an obvious candidate in this respect, in the sense that the fundament for knowledge of the physical world is to be found in Res Cogitans and that it further is secured by God. Even if Quine does not mention Descartes, he argues against the idea of a philosophical fundament for the certainty of scientific knowledge, a fundament steadier than ordinary scientific method. Epistemology cannot supply a justification for our knowledge of the external world, he holds, and any justification for such knowledge is thus to be found in further empirical or scientific investigation. Quine sees epistemology as a continuation of the natural sciences, and philosophy can be distinguished from the sciences by its generality. It engages in broader and more general matters than those entertained in sciences such as mathematics and psychology.[4] Otherwise, when it comes to methodology, evidence and founding, he maintains, epistemology is of the same standing as the sciences.

Quine claims that skeptical questions arise within science. It is our success in understanding the world, and the subsequent realization that there could be a difference between the way the world appears to us, and the way the world is, which renders the skeptical question possible in the first place. Therefore, Quine upholds, we can employ the resources of science itself (in his broad sense) in order to answer the questions that science has given rise to. If we are trying to explain how empirical knowledge is possible, the classical epistemological view will be that it is illegitimate to use the resources of science in answering this question. The idea is that we would then employ the very knowledge that has been questioned, and this would therefore not be regarded as an answer to the question,

[4] This could be put even stronger, as we read Quine in his "Has Philosophy Lost Contact with People?" Quine points out that many of the greatest philosophers of the past also were scientists, and that their philosophical studies, in terms of a struggle with "concepts and the quest for a system on a grand scale" were integral to their scientific enterprise. Willard Van Orman Quine: "Has Philosophy Lost Contact with People?" pp 190-193 in *Theories and Things*, Harvard University Press, Cambridge, Mass. 1981.

or alternatively, it would be deemed a bypassing of the problem.[5] According to Quine, on the other hand, we should rather look at the question of how knowledge is possible as an empirical question. It is a question of how creatures like us could end up having correct beliefs about the world. The question is: "How do correct beliefs about the world come into being?" And it is precisely this question he addresses in his epistemology, even though his investigation is far from what we would normally deem empirical, since it, as we have seen in article 1, consists of a thought experiment. The thought experiment of radical translation is, however, informed and restricted by our knowledge of empirical matters.

What Quine thinks epistemology properly should engage in is the relation between humans and their surroundings. Epistemology should be an inquiry into how, based on sensory stimulation, people can arrive at the beliefs they do have about the world, including scientific beliefs about nerve endings etc. According to Quine the object of study is the relation between theory and evidence, or more specifically, a study of *how* evidence relates to theory. "The relation between the meager input and the torrential output is a relation that we are prompted to study for somewhat the same reasons that always prompted epistemology; namely, in order to see how evidence relates to theory, and in what ways one's theory of nature transcends any available evidence."[6] It is the relation between "surface irritation" and "knowledge of the world", or again the relation between "sensory stimulation" and "belief" that is to be studied. Epistemology as naturalized[7] is, in Quine's view, the study of the relation between sensory evidence (i.e. the stimulation of our sensory receptors) and our descriptions of the external world. Observation sentences[8] play a particular role in this study. Firstly, they are the evidence on which our theory rests. Secondly, they are located at an intersection between language and reality where language confronts reality in a sufficiently direct manner for single sentences to be learnt one by one. When learning a language, Quine holds, we go from learning simple observation

[5] Descartes, for instance, well aware of the "pitfall", tried to answer the skeptic by denying all the beliefs that he found to be logically possible to doubt, and is generally considered as thereby having managed to avoid getting into a scrape.

[6] W. V. Quine: "Epistemology naturalized", pp 69-90 in his *Ontological Relativity & Other Essays*, Columbia University Press, New York 1969, p. 82f.

[7] Naturalized as for *natural* knowledge, not *formal*, even if there according to Quine are only gradual differences here. Hume, on the other hand, whose «predicament» Quine here relates to, distinguishes sharply between the two forms of knowledge. Quine makes use of this distinction in his analysis of Russell and Carnap's achievements in his "Epistemology Naturalized".

[8] Observation sentences are situation sentences that have a stimulus meaning that is not varying when influenced by additional (collateral) information. Their stimulus meaning yields full justice to their meaning. This is so because they are initiated by an intersubjectively observable present stimulus.

sentences towards an understanding of more complex and more theoretical sentences. As the quotation above suggests, the evidence does not "cover" the theory and this means our theory, for principled reasons, will not be properly supported.

Quine's own investigation into the relation between humans and their surroundings consists in a study of language learning. The investigation is pursued as the thought experiment of radical translation. The inquiry could, perhaps, be seen as a substitute for the empirical inquiries that Quine generally advices us to carry out, but as Quine denounces strict boundaries between that which is empirical and the theoretical, he would, on his own terms, be agnostic to the question. Quine's theory of radical translation is thus his choice of investigation into epistemology, and radical translation is the basis for his developed theory of knowledge.

Quine's naturalized epistemology is concerned with the relation between philosophy and science. Holism (which in this case means that a scientific hypothesis cannot be tested on its own) is one of Quine's strongest arguments for naturalism. The reason why a scientific hypothesis cannot be tested alone is that an empirical hypothesis is tied up with the rest of our system of beliefs (and our more theoretical hypotheses) to such an extent that it is impossible to localize something like the hypothesis' own stock of privileged observation sentences to uniquely verify or falsify it. According to a Quinean view, there is a continuum from the theoretical to the empirical, and from philosophy to science.[9] Accordingly, there is no first philosophy. Hence, the thesis of gradualism is at the center of Quine's naturalized epistemology.

3. Davidson's understanding of Naturalism

Davidson's understanding of naturalism is not obvious even to faithful and observant readers of Davidson's writings. The theme is seldom discussed, and a general reason for this can be found in Davidson's reluctance to characterize philosophy in terms of "isms". When Davidson takes up the theme, it is in relation to the problem of other minds, i.e. the problem that arises when we ask how we can tell that another person has experiences and

[9] There is, according to Quine, also a continuum from ordinary belief to science. The difference between them is thus not a principled epistemic difference, but a difference in exactness and strictness of method.

thoughts anything similar to our own. Observation of the other person's behavior, verbal as well as non-verbal, is what we have to go on. According to Davidson, "The problem has not been solved, at least to most philosopher's satisfaction. What has happened instead is that the subject has undergone a sort of naturalization."[10] In a footnote he adds that: "I say a "sort" of naturalization because the word has been used so differently by different philosophers. Here I am not taking naturalization to involve showing, or trying to show, how to reduce talk of mental states to something that can be subsumed under the natural sciences; *I do take it to involve shifting from trying to justify our claims to knowledge to describing our normal ways of achieving knowledge.*"[11] According to Davidson, the problem of other minds has been transformed from being a presumed solvable philosophical problem and into being a study of normal ways of achieving knowledge.

Davidson says that where we earlier tried to answer the skeptic, we now assume that we know more or less what goes on in the minds of others. We acknowledge that we already have a good grip on what is going on, and we wonder how it came about that we got this good grip on things. What we do is that we start noting how we go about when we actually acquire or arrive at the knowledge that we have of other minds. "How are we able to find out what goes on in the minds of others?" is then our leading question, and an answer to this question could be a description of our normal ways of finding out what goes on in others' minds.[12] Thus, we have to describe our normal ways of interpreting each other. This means that an interpretational approach in epistemology, to the extent that it describes normal ways of interpretation between speakers, *is* a naturalized study of knowledge, according to the understanding of naturalism found in Davidson. But interpretationism as a version of naturalized epistemology is still not empirical psychology. I would say that interpretationism is guided by quite different ideas about the relation between science and philosophy than those that guide defenders of the so-called "replacement-thesis". Radical interpretation is an epistemological approach where the differences between ordinary belief, philosophy and science are gradual. It could be argued that those who take Quine's advice about replacing traditional epistemology with empirical psychology to be an advice about leaving philosophy altogether, haven't understood Quine well.

[10] Donald Davidson: "Interpretation: Hard in Theory, Easy in Practice", p. 31 in Mario De Caro (ed): *Interpretation and Causes. New Perspectives on Donald Davidson's Philosophy*, Kluwer, Dordrecht 1999.
[11] Ibid, p. 43, n. 1.
[12] Donald Davidson: "Interpretation: Hard in Theory, Easy in Practice", p. 31 in Mario De Caro (ed): *Interpretation and Causes. New Perspectives on Donald Davidson's Philosophy*, Kluwer, Dordrecht 1999.

On the other hand, it has been assumed that Davidson's theory of interpretation cannot be naturalized insofar as it contains at least one principle with a supposed constitutive status, i.e. the principle of charity. The presumed epistemic status of the principle of charity has, in my opinion, rightfully, been regarded incompatible with the naturalistic doctrine of gradualism.[13] If we presuppose that constitutive principles do have a principled *different* epistemic status than other kinds of principles, that we e.g. ascribe constitutive principles a status as synthetic apriori, and we simultaneously embrace gradualism, then we have an inconsistent theory. If this is the case, we have to choose between operating with principles with a fundamentally different epistemic status than the rest of our principles and theses on the one hand, and the denouncement of gradualism, on the other. Constitutive principles are normally localized in theories seen as opposed to naturalism. If it is correct that Davidson's epistemology contains principles with another epistemic status than empirical knowledge, it is not likely that his epistemology could be naturalized in a strictly Quinean sense. This means that there are serious reasons for doubting that there could be something like a naturalized epistemology containing constitutive principles.

Anyhow, Davidson is accepting a form of naturalism in the elongation of Quine's thinking. Davidson regards Quine's naturalized epistemology as first of all aiming at giving an account of *what normally happens when we acquire knowledge*. This means that epistemology should not try to give a fundament for knowledge, and neither should it in other ways take its purpose to be the delivery of a foundation for knowledge.[14] Davidson notes that the method of starting by accepting common sense or science, then going on "to ask for a description of the nature and origins of such knowledge" has been called naturalism, and he seems to have no misgivings if his way of doing philosophy is being characterized in this way.[15] The next feature of Quine's naturalized epistemology that Davidson says that he wishes to continue to pursue is Quine's *third person approach to epistemology*. I assume that Davidson is referring to Quine's "method" in epistemology, the situation described as radical translation, in which an interpreter takes the third person perspective in observing the relation between stimuli from external physical surroundings

[13] See Mikael Janvid: *Naturalism and the status of epistemology*, Preprint no. 7, Department of Philosophy, Stockholm University, Stockholm 2001.
[14] "Epistemology Externalized" (1990), p. 193 in Donald Davidson: *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.
[15] Donald Davidson: "The Problem of Objectivity", p. 5 in *Problems of Rationality*, Oxford University Press, Oxford 2004.

for the speaker, and the speaker's theory. This is Davidson's methodological starting point, but he soon advances into a more complex situation, in which the interpreter also observes his own relations of stimuli and belief-system and correlates this with the speaker's. This more complex situation presupposes a more extensive use of the principle of charity, but does not break with the original methodology in any principled sense. Davidson does not, on the other hand, accept Quine's account of the nature of knowledge, because Davidson takes Quine's view to be essentially first person oriented and Cartesian.[16] Quine's insistence on stimulation of the individual's sensory surface as the source of knowledge of an external world is an important element in his view upon knowledge. Quine takes for granted that our knowledge of the external world comes from the causal effect this world has on the nerve endings of each subject. Davidson holds, on the other hand, that objects and events in the physical world are what influences us causally. This is why Davidson considers the Quinean approach first person oriented. According to Davidson, objects and events cannot play an epistemically justifying role, whereas Quine, on the other hand, takes sensory stimuli to be evidence.[17] Hence, the reason why Davidson's approach presumably is not first person oriented is that objects and events, according to Davidson, are partly socially constituted. It is, however, something of an irony that Quine is charged with being Cartesian, given that Quine's overall aim in his project of naturalizing epistemology is to overcome foundationalism, a characteristic trait of Cartesianism.

What Davidson thus brings along from Quine's naturalism is, firstly, a denial of the claim that epistemology is a kind of first philosophy, that can form a basis for all other kinds of knowledge. This implies that instead of asking *whether* we have knowledge, epistemology should be occupied with *how we acquire* knowledge. The study of how we acquire knowledge should be carried out from a third person perspective, and this is the second aspect that Davidson brings along from Quine's naturalized epistemology, even though this second feature is not specifically associated with naturalism. Hence, I would say that with (only) these two features, Davidson's own formulation of naturalism is quite general.

[16] Donald Davidson: "Epistemology Externalized" (1990), p. 194 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.
[17] See W. V. Quine: "Epistemology Naturalized" p. 75 in *Ontological Relativity*, Columbia University Press, New York 1969.

However, it is pretty clear that Davidson has also adopted Quine's holism and fallibilism. As Quine, Davidson could be characterized as doing immanent, as opposed to transcendent (-al) philosophy, in the sense that they both deny that epistemology can be foundational. But if Davidson has accepted Quine's thesis of gradualism; that there are only gradual differences between philosophy and science, and fallibilism; that everything in principle is up for revision, then we would suppose that neither could Davidson's radical interpretation be apriori in a classical sense. Davidson does not explicitly say that philosophy and science is of the same standing, but has denied that epistemology can be a first philosophy. Epistemology does not precede science, and, according to Davidson, epistemology cannot stand above and "evaluate" scientific contributions. Both Quine and Davidson would basically reject the existence of eternal and unchangeable structures as something that knowledge could be founded upon, and when they both hold that all forms of knowledge are in principle fallible, this would also have to be so for epistemic structures. The question that I have raised is whether some of the conditions and presuppositions that are part of Davidson's philosophy deserve another kind of status than Davidson has assigned knowledge in general. We could more generally say that the question is whether Davidson could be read as a naturalist, in the sense accounted for above, through and through. Now, we go on to summarize each of the articles and the question of naturalism thus related will be our particular interest.

4. The Explanation of Error

Article 4, "The Explanation of Error", gives the most thorough account for the problem of error within this collection of articles. I start by distinguishing between an internal and an external perspective on intentionality and the problem of error. Davidson's attempt at an external explanation points out the following three conditions for thought about an external world: First, common dispositions for discrimination in perception that humans share, such that we categorize the world in similar ways, and judge similarly about similarities. Second, there is the condition that something statistically will be picked out as a type of cause for a given thought, and this is the normal cause for a thought. The third condition is that error can occur, and this happens when creatures respond to stimuli in non-characteristic ways, creating a divergence from normally similar reactions. These three conditions are necessary conditions for thought and language, but are insufficient for an explanation of these. Hence, Davidson points out a fourth condition – regarding the acquisition of language in a society of language users. According to Davidson, the acquisition of language with the help of ostensive teaching is necessary in order to explain how the intentional can be possible and how thought and language can emerge. But Davidson holds that we have still not, even with this fourth condition, given sufficient conditions for intentionality, and thus for objective thought. However, he maintains that we should not expect further progress in an attempt at an external explanation, as he holds that this would "amount to a reduction of the intentional to the extensional". I may add that such a reduction would explain the phenomenon away, in the sense that it would disappear in the form in which we are interested in it. Davidson recommends that we continue our search for an explanation of thought about an external world, and our search for an explanation of error, from within the intentional.

Explaining the intentional, explaining thought and language, explaining objective thought, and explaining error are, as we can see, assumed to be different sides of the same project. If we could give an explanation of one, we would have an explanation of all. But what is the motivation behind this felt need for an explanation? There can be many reasons for such a craving. It could be as simple as a "what is really...?" question, and such questions are indeed intrinsic to philosophy. However, I have taken it that there are two different motivations for the quest for an explanation of error. One is the more general quest for objectivity of knowledge of the world, of other minds and of our own minds, and objectivity of meaning. This motivation arises from a pressure from different kinds of skepticism towards Davidson's account of meaning and knowledge. The other motivation is internal to the method of radical interpretation and the model of triangulation, and it is a question of explaining how the interpreter can explain error in the interpretational situation. Moreover, these two motivations are connected.

When I compare the accounts of radical translation and radical interpretation on the question of explaining the possibility of error, I find that Quine never had this problem. I argue that Davidson's focus on matching of truth conditions for the utterances of speaker and interpreter "creates" the problem in the first place. For Quine, smooth communication is the only criterion needed for saying whether or not a translation manual is a success. This is due to his notion of stimulus meaning, which is an empirical concept. Matching stimulus meaning for informant and manual is enough to secure smooth communication. According to Quine, all there ever was to semantic meaning is absorbed in successful translation. Davidson, however, has the matching of truth conditions for the speaker's utterance with the truth conditions for the interpreter's analytical hypothesis of the utterance (which is built on the interpreters own set of beliefs) as the correlation to be made. Now, the truth conditions will necessarily be similar for the speaker's utterance and the interpreter's interpretation of the utterance, since this is all built upon the interpreter's reasonable beliefs in the first place. This means that there is nothing further to appeal to. We could always ask why we would need a court of appeal in this case, when we didn't need one in Quine's case. The answer is that Davidson has brought in truth. While we, in the case of radical translation, have all the evidence there is, and evidence is all we need, Davidson is not content with evidence. He wants objective truth. Matching utterances' truth conditions and sharing verbal behavior is not sufficient since, according to Davidson, both the speaker and the interpreter might be wrong. Smooth communication is simply not sufficient. The demands have been raised, and from Davidson's perspective there was a good reason for the demands to be raised. I maintain that anything less than the truth is too little to cash out Davidson's concept of knowledge.

I argue that we have to recognize two kinds of error in order to get to grips with the problem of explaining error. There is error in belief, which amounts to false belief, and there is error in meaning. The determination of what kind of error we are dealing with is in principle hypothetical, as is radical interpretation in general. I argue that propositional seeing and language is a condition for error in meaning, and that it thus is a condition for the possibility of recognizing two kinds of error. The reason for this is mainly methodological; if the creature in question is non-linguistic, its seeing is accordingly non-propositional, that is, as far as an interpreter can know. It must therefore be clear that the notion of error that we are dealing with is related to *linguistic* creatures.

Since the notion of error is so clearly a notion related to speakers, it is important to uphold the distinction between making a mistake in meaning, or a linguistic mistake and making a mistake in belief, or a factual mistake. I take up a suggestion by Eva Picardi, where the contention is that it is Davidson's focus on the idiolect (the language of an individual speaker) that endangers or blurs this distinction within Davidson's theory. This is due to the methodological situation of the radical interpreter, as the interpreter relates only to one speaker at a time. When the interpreter finds that some specific verbal behavior of the speaker diverges from the normal ways of the speaker, he tries to check whether the irregularity is due to a mistake in belief about the world for the speaker, or if it is due to some conceptual mistake in the truth theory for the speaker that the interpreter has conjectured. The interpreter's check will have to be a social check no matter what kind of error has been made. It is possible to disregard errors that are due to different localization and therewith - different perceptions. The interpreter can therefore find that regularity of behavior is established on an observational basis, so that the correlations between verbal behavior and stimuli are the same for the speaker over time, and that these correlated sets are the same for the speaker and the interpreter over some given time. The interpreter can thus check whether he and the speaker are picking out the same object in such a way that regularity in behavior has been established. But the problem is that in order to find out whether the speaker has employed the words in a "wrong" way, the interpreter can only rely on the speaker's previous utterances. The interpreter can in this way check for coherence with the speaker's previous utterances, and correlate utterance with circumstances, then and now. And there it stops. A check of the speaker's utterances against the speaker's own utterances, as the interpreter has interpreted them, is not a sufficient social check. It is not sufficient for testing whether both speaker and interpreter can be wrong.

This means that the social comes in at the point of the emergence of individual languages, and that it due to the radical interpretational model cannot come in, in significant ways, at later occasions. The reason why this is a problem, does, however, not turn up when irregularities occur. The problem turns out to be that the interpreter cannot know whether or not both the speaker and the interpreter are wrong when communication is smooth. The reason why one would need a social check when communication is smooth is Davidson's presumption that objective truth is to be established in triangulation. As long as objective truth is out of reach for the speaker-interpreters in radical interpretation, we will not be able to solve the problem of error. My contention is that the individual concept of language, together with the craving for objective truth, so far has been an obstacle in the path of an explanation of error within the radical interpretational approach. We need a community-based notion of language, but this does not blend well with the basically individual methodology of radical interpretation.

In article 4 we have seen that Davidson's version of epistemology is naturalistic in yet another sense than already noted in the previous treatment of the theme in this synthesis. Davidson's account of intentionality is based in evolutionary thinking in the sense that he considers human dispositions for discrimination one out of four conditions for thought. This means that human rationality is partly based in human nature and not in some McDowellian "logical space of reasons" or some Kantian "realm of freedom". Furthermore, the statistic notion of normality likewise points in a naturalist, rather then a transcendental (-pragmatic) direction.

Article 4 also raised the question of whether a concept of truth, and, in particular, the concept of objective truth, has pretensions that are more metaphysical than the otherwise naturalized conception of radical interpretation allows for. Objective truth is most certainly an on-or-off value; it is not gradual and fallible.[18] This question will be dealt with further in the three last chapters of this synthesis.

5. The Status of the Principle of Charity

In article 3, "The Status of the Principle of Charity", I analyze the function of the principle of charity within the theory of radical interpretation and discuss the epistemic status of the principle. An account of the genesis of the principle within Davidson's authorship is given, as I believe that such an approach will be important for a proper understanding of the principle's status. I present and discuss an account of the weighing function that the principle has towards meaning and belief, as the principle makes it possible for the interpreter to hold beliefs constant while solving for meaning. I emphasize the principle's role to enable the production of a first hypothesis of the speaker's meaning, and also the principle's role to ensure truth and consistency in a first attempt at a theory about the speaker's meanings and beliefs. The claim that the principle brings out *hypotheses* about meanings and beliefs is decisive for our reading of Davidson's project. I maintain that an

[18] See especially p. 83 and p. 92f of my article 4.

alternative claim that the principle *constitutes* meanings and beliefs is the origin of incompatible readings.

Radical interpretation is a thought experiment, and a conceptual exercise. I stress the importance of not taking radical interpretation to be ideal, in the sense of a prefigurative or exemplary situation of interpretation. The interpreter is not ideal, and neither is the speaker. They are both normal in the sense that they are fallible. Thus I would say that it would be wrong to assume that the principle of charity in radical interpretation could constitute a speaker's actual meanings and beliefs as what would be *ideally rational* for him to believe and mean. However, radical interpretation is a theoretical model, and the principle is designed for this model, not for real interpretation. I argue that if radical interpretation did not use normal fallible interpreters and speakers as models in radical interpretation, the thought experiment would have no transmittance value. Moreover, if we think of the principle of charity as a normative principle, in the sense that we think about the speaker as being *more* consistent than an ordinary speaker and that the principle suggest that we impose *ideal* rationality on to the speaker, Davidson can easily be construed as an idealist and indeed an anti-realist philosopher. Even if he hadn't cared much, due to his lack of respect for thinking in "isms", many objections based upon misunderstandings of his project would be apt to be produced.[19]

I discuss whether the principle of charity can be synthetic a priori, as Davidson once suggested. The principle is a condition for interpretation, and it is constitutive in one sense, since one inevitably needs to use the principle in the forming of a theory of meaning and belief for a speaker. According to a Kantian conception of a priori, the a priori makes experience possible, but the a priori is independent of experience. However, the entire thought experiment of radical interpretation completely antedates experience, and this generally makes it hard to evaluate whether it is apriori in Kant's sense. According to Kant, that which is a priori is supposed to be necessary and non-revisable, and the principle is arguably necessary within the theory of radical interpretation, but it is otherwise hard to tell. I find that the theory as such is revisable, but it is not easy to find a way to decide

[19] See e.g. J.J.C. Smart's article "Correspondence, Coherence, and Realism" in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, pp 109-122.

whether or not the principle is revisable. At the least, I argue, there is no *particular* experience that could be relevant for revising it.

Since the Kantian concept does not quite seem to fit our case, I consider a more modern concept of a prioricity, employed in the introduction to an anthology, *New Essays on the A Priori,* from 2000, edited by Paul Boghossian and Christopher Peacocke. In this version of a prioricity it is assumed that the truth of an a priori assertion can be known independently of experience. However, being a priori is neither to be considered necessary nor infallibly knowable. Still, the a priori seemingly cannot be invalidated by "wholly empirical information". Transferred to the case of the principle of charity; we can know that the principle of charity must be valid in radical interpretation, and this comes close to saying that we can know its truth independently of experience. But if the principle were found to be revisable, it would be hard to uphold a claim that the principle comes within the scope of a concept of a prioricity. However, Boghossian and Peacocke hold that precisely these two traits are hallmarks of a specific form of a prioricity. They maintain that the case for "non-conclusive a priori entitlement" is a case where we can know the truth of an assertion independently of experience, and where it simultaneously is open to empirically conditioned revision. The interesting question for us, when we consider the principle of charity, is therefore whether or not it is open to empirically conditioned revision. Or, to turn the other side out: is the principle immune to empirical invalidation or not?

I consider the traits of radical interpretation that is not dependent on empirical reality, such as presuppositions about what a natural language is, and the psychological and physical reality of the creatures involved. These presuppositions are very general, and I argue that it is unlikely that it could turn out that we are genuinely mistaken about these traits. However, if it turns out that we are mistaken in making these assumptions, I maintain that the consequences for the status of the principle are obscure. It may be, however, that the principle can be empirically invalidated in some *other* way, but it is not easy to see how this could come about.

The question of empirical invalidation of the principle could, however, be read as a question about the status of the *beliefs* that we ascribe to a speaker with the help of the principle. This means that we really meant to ask whether the beliefs ascribed to the speaker in radical interpretation can be empirically invalidated or not. If we assume that the principle is a priori, could it be that the beliefs that the principle helps us to ascribe is also a priori? No, this can hardly be a consequence, since the beliefs that we ascribe to a speaker (according to the interpretational theory) are always hypothetical ascriptions, and besides, the speaker is normal. Hence, it is highly probable that some of his beliefs are false.

But it could also be that we meant to ask whether the beliefs we ascribe to the speaker by using the principle come out right. This means asking whether we, in employing the principle are able to get a correct theory of the speaker's beliefs and meanings. Could it be the case that another principle, or a completely different method, could do this job in such a way that a result would be a radically different theory about the speaker's beliefs and meanings, such that the speaker comes out with different beliefs?[20] In order to check this, we would have to be able to discover the beliefs of the speaker independently of the principle of charity, and then check whether the result obtained in this way is in agreement with the result we obtain by employment of the principle of charity. It is hard to imagine how we could perform such a check independent of the principle. For the theory of radical interpretation the principle thus seems to be necessary.

The principle could, of course, be indirectly invalidated if we were to replace the theory of radical interpretation with an entirely new theory, but then we could hardly say that the principle has been *empirically* invalidated.

I conclude that it *so far seems* that the principle of charity cannot be empirically invalidated. It seems that we can know the truth (or validity) of the principle independently of experience, and it does not seem to be open to empirically conditioned revision. Hence it does not seem to be a case of "non-conclusive a priori entitlement". So far it looks more like it is a case of "conclusive" a priori entitlement. I argue that the reason for this is more likely to be found in the character of radical interpretation than in any epistemologically decisive analysis of the relation between mind and world. Radical interpretation is a thought experiment, and since the principle of charity is a principle that has its place within

[20] This result is of course already a possibility within radical interpretation, due to the thesis of the indeterminacy of interpretation, but then the reason for the result is another.

this experiment, it is only doubtfully open to empirical revision. If we on this basis hold that the principle is a priori, it would not be directly false, but it would be misleading and perhaps also irrelevant. I moreover maintain that it is generally problematic to separate the empirical from the a priori in an epistemically significant manner within Davidson's theory, since further support for such a separation is lacking. On the contrary, Davidson's thinking has been built upon a denial of the epistemic importance of such dualisms. The realization of this point has been one of the main sources for my overall reading of Davidson as a naturalist. Having found that it is very doubtful that the principle of charity can be ascribed the status of synthetic apriori, or transcendental, the epistemic status of the principle is hence not a hindrance for a naturalistic doctrine of gradualism.

Furthermore, I find it characteristic that the thought experiment of radical interpretation is not about ideal and prefigurative interpreters and speakers. However, "The omniscient interpreter" is a figure once brought in by Davidson.[21] Later, he has considered this a big mistake.[22] Even so I would say that the fact that the figure could turn up in Davidson's thinking at all shows that his philosophy has developed in the reach between transcendentalism and naturalism. However, I consider the fact that the figure had a very brief appearance as a confirming fact of my overall reading.

I conclude that the principle does not constitute the beliefs of either speakers or interpreters as true beliefs. The principle helps the interpreter in his construction of a theory about the beliefs and meanings of the speaker; it helps bringing about hypotheses about belief and meaning, but it is not in position to issue a warrant on the *truth* of the result. The *consistency* of the beliefs is, however, a direct result of the method.

6. Second Order Belief and Naturalism

Article 6, "Ascription of Belief and Second Order Belief", addresses interrelated intentional concepts such as belief, second order belief, truth, objective truth, subjective

[21] In Davidson's "A Coherence Theory of Truth and Knowledge" (1983), pp137-153 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

[22] In Donald Davidson: "Reply to A.C. Genova" pp 192-194 in Lewis Edwin Hahn (Ed.): The Philosophy of Donald Davidson, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, p. 192. Davidson admits: "If the case can be made with an omniscient interpreter, it can be made without, and better."

judgment and error. This means that this article, to a greater degree than the others, takes an internal perspective on intentionality, and, in particular, it raises the question of how a concept of objective truth is partly constituted from within intentionality. For this constitution, the question of the relation between a notion of belief and a notion of second order belief is of particular interest. I discuss four of Davidson's criteria for ascription of attitudes and give an account of his notion of belief.

Davidson's first criterion for ascription of belief is that in order to ascribe belief to somebody that belief has to be one belief in a set of beliefs. I interpret this as a methodologically based restriction on interpretation. This is based on an opposition between methodological and ontological restrictions. Also his fourth criterion, that in order to have the concept of belief, one must have language, is a methodologically based restriction. The notion of belief is characterized in detail, and some of the most important features are that a belief is "non-Platonist", in the sense that it demands a subject, a believer, and we are therefore personally connected to it and we have different degrees of faith in each of our beliefs. But a belief is not yet subjective, as it has a sentence as its representative. Due to the way beliefs are constituted, we can know that most beliefs in a set of beliefs are true, but a belief is still always hypothetical in the sense that it can, in principle, never be once and for all confirmed as a truth for neither the believer, nor for his interpreter. We can identify a belief due to its truth conditions and due to its place in a network of belief. This means that beliefs belong both to human assessing subjects and the world.

The criterion that I pay the most attention to in article 6 is the criterion that second order belief is necessary for having belief. Research on autistic disorders anticipates that there exists speaker's who have beliefs but lack second order belief. Kathrin Glüer and Peter Pagin take these autistic speakers to be counterexamples to the criterion that belief requires second order belief, and also to the claim that one cannot have beliefs unless one has the concept of objective truth. I argue that in order to establish the results of the research in psychology as a counterexample, one has to show that the concepts of belief involved in the research and in Davidson are sufficiently similar for a comparison to hold. Glüer and Pagin has not taken any steps in this direction, and I take it that it is highly doubtful that the sophisticated and specified concept of belief employed in Davidson's theory will be the same as the one employed in psychological research, which probably will be correspondingly specified, albeit in other ways. Furthermore, when it comes to the lack of second order belief, I argue that deviant linguistic behavior (as shown by some autistic persons) is parasitic upon ordinary linguistic behavior reliant on second order belief (shown by non-autistic persons). As long as we take it that a community of autistic persons is unable to develop language from a pre-linguistic state, we must assume that the autistic persons benefit from having learned to speak in a non-autistic environment. Glüer and Pagin argue that if only autistic speakers of a linguistic community survive, the lack of counterexamples becomes a mere contingency. Contrary to this I argue that as long as the conditions for the development of language are not only metaphysical, but to a high degree biological, causal and demands an individual history of language acquisition in a society, this will be a part of our understanding of, and theorizing about, these speakers. I therefore disavow the claim that the existence of autistic speakers who lack second order belief is a counterexample to the criterion that belief requires second order belief, and I also disagree that this case is a counterexample to the claim that one cannot have beliefs unless one has the concept of objective truth.

Finally, I point out that as a criterion for ascription of propositional attitudes, second order belief is inevitable for interpretationism in epistemology. In order to interpret in accordance with the prescribed methodology, and in particular with the principle of charity, the creatures involved need to be able to assess their own and other people's beliefs. It is also vital for a constitution of the notion of objective truth from within intentionality.

When it comes to the question of the actual existence of second order beliefs in speakers, it could be said that, on the one hand; we should be able to know all that we can ever come to know about the beliefs of a speaker, from a third person point of view, according to the prescribed methodology of radical interpretation. On the other hand, Davidson claims that all believers *have* second order belief, but there is a question whether this could be available as a claim about the empirical existence of a particular state from a third person perspective. It seems to me that for an interpretationist it could only be available as a claim on a meta-level, in a theoretical reconstruction of interpretation, or, alternatively, as a theoretical presupposition in a thought experiment about interpretation. Therefore it could seem as if interpretationism as a method dodges the question of the *existence* of second order belief, and that interpretationism therefore must renounce naturalization in the form of a justified or verified empirical existence in this case.

In article 6 I have again raised the question of whether first, the criterion of predication and second, the criterion of the other creature's predicated response to my predication are sufficient for having evidence that the other creature has the awareness of error. (This was the starting question for article 5.) In article 5 I suggested that disagreement should be the criterion for ascription of awareness of error, as this would be the only methodologically sound criterion from a behavioral point of view. Now it seems that we in article 6 have gotten yet another confirmation of the presumption that smooth communication is insufficient for having evidence that the other creature is in possession of awareness of error. We can see that there exist language users who are able to smoothly communicate, but who simultaneously are not aware of the possibility of error. However, it has turned out that this is not necessarily a counter-argument against the criterion of second order belief as a criterion for having belief.

7. Radical Interpretation, Naturalism and Empiricism

In article 1, "Radical Translation and Radical Interpretation", I go into differences between Quine and Davidson as concerns epistemology. Though both of them pursue naturalized epistemology in some sense, the article points out that naturalism need not be empiricism. Davidson agrees with Quine that epistemology should be a study of the relation between input and output, but he changes the study of this relation from a study of radical translation to a study of radical interpretation. I find that a difference between radical translation and radical interpretation is that in interpretation our aim is to understand another person, via a truth-theory for the speaker, while the aim of translation is to produce a translation manual, thereby bringing out whatever there could be to "stimulus meaning". Davidson's ambition of reaching understanding of the speaker, via a theory of meaning and belief, brought out by a truth theory, far exceeds the use Quine made of his radical translation. This means that Davidson had to go beyond holophrastic translation, over to analytic interpretation that could bring out semantic structure.

Another perspicuous difference between the two approaches lies in their ways of relating to the intentional. While Davidson uses the model of radical interpretation as a way of bringing about understanding of intentional phenomena and also willingly employ intentional concepts from the outset, I argue that Quine sees radical translation as a move away from the metaphysically loaded intentional concepts. The intentional concepts are thus, with Davidson, brought back to the stage, but now with the aim of ridding them from their heavy metaphysical baggage. However, as radical interpreters our observational apparatus for taking in the world is not equipped with notions like meaning. Thus, in this regard Davidson is in line with Quine. As interpreters we cannot appeal to such notions when they are supposed to be the outcome of the interpretational process. But it seems to me that Quine was never much interested in bringing a notion like meaning back into the vocabulary of his theory however restored to dignity it may have been. Somewhat reluctantly he did go as far as crediting us with a notion of stimulus meaning. But later he turned away from even this limited notion of meaning.

Let's quickly recapitulate the main methodological differences between radical translation and radical interpretation: Quine prescribes a methodology that starts out with the observation of linguistic behavior related to the speaker's immediate environment, thereby establishing patterns of connections between behavior and surroundings. Roughly, the best possible outcome of the process will be that the translator finds stimulus-meanings for occasion sentences that have a high degree of observationality, and that he also is able to translate the logical connectives. The stimulus meaning for a person's sentence is the ordered pair of affirmative and negative stimulus meaning, and "the stimulus meaning of a sentence for a subject sums up his disposition to assent to or dissent from the sentence in response to present stimulations."[23] On basis of this, and assumptions about "stimulus-analyticity" and "stimulus-contradiction", the radical interpreter can translate some of the speaker's sentences by correlating these with sentences of his own. In this way the translator can reach the aim of giving a translation manual for a speaker of an unknown language. This will in effect be a correlation of synonymous sentences. In radical interpretation the aim is to devise a theory of truth for the speaker. The radical interpreter starts to search for patterns of sentences that the speaker *holds true* (or holds false) systematically dependent on changes in the environment. (The corresponding attitude in

[23] Willard Van Orman Quine: *Word and Object*, The MIT Press, Cambridge, Mass. 1960, p. 34.

radical translation would be prompted assent.) But since the interpreter has to go deeper into the structure of the language (and thus go *into* the sentences) in order to reveal the semantic structure of the language, he also has to employ more of his own logic, more of the structure of his own language, and more of his own system of belief in order to complete his mission than the translator had to do. Davidson therefore has to "apply the Principle of Charity across the board." When the interpreter has been successful, he has given a theory of truth for the language of the speaker, and the interpreter will then have an interpretation of the speaker's sentences, in the sense that he knows what the speaker means when uttering the sentences and he also knows the corresponding beliefs of the speaker. He thus understands the speaker, which means, among other things, that he can with some success predict his response to certain stimulations.

When trying to search out the epistemologically more *decisive* differences between Quine and Davidson, I have focused on their different views on the causes for beliefs. While Quine takes patterns of sensory stimulation to be the data on which translator and informant reacts, objects and events are the causes for beliefs and linguistic behavior in Davidson's theory. According to Quine the epistemological study of input and output would suffer if we take objects and events as the causes (or, as input), since the very process of reification is among the objects of his study. On the other hand, Quine eventually changed his mind about this in relation to translation and language learning. Quine came to hold the view that objects and events are the input in *these* connections, though not in the study of epistemology.

It has long been assumed that Davidson is a very loyal student of Quine's. Even though I, in article 1, agree with the claim, I argue that it is a truth with some very important and significant modifications. These modifications have motivated my article in the sense that I'm aiming at pinpointing the philosophical importance of the seemingly small and detailed differences between Quine's radical translation and Davidson's radical interpretation. I believe that if we want to understand what Davidson is up to, we must also understand what Quine is up to. But it would be a mistake to assume that Davidson had the same motivations and aims for his theorizing as had Quine. If we take Davidson's philosophy to be first and foremost a prolongation of Quine's, we shall be blind for Davidson's accomplishments. If we, on the other hand, fail to see Quine's enormous impact on

Davidson, we shall also be blinded in our reading of Davidson. My reading of some of the detailed differences of the two projects has revealed the substratum for Quine and Davidson's more decisively epistemologically different orientations. In particular, Davidson's focus on objects and truth (as opposed to Quine's focus on stimulation of nerve endings and evidence) has been at the core of their discussions over the years. I localize the origin for the epistemological differences between the two in Davidson's use of quantification in interpretation. However, to my knowledge, nobody has focused on the connection between this development in Davidson's theory, which had to come due to his shift to semantic theory, and the epistemological consequences of the shift. In article 1 I claim that the indeterminacy is less radical with Davidson's shift, and that ontological relativity to a large extent is precluded by the new approach to quantification. My contention is that the preclusion of ontological relativity is connected to Davidson's rejection of irritations on our sensory surface as the cause of our beliefs, or rather, his reliance on *socially* constituted objects and events.

There is a fundamental difference between Quine and Davidson when it comes to what is to count as a justification for a belief. For Quine, justification is what he calls "the tribunal of experience", which is something like a "court of justification" where experience is the jury and the judge. What we are given through our senses plays a particular role for our knowledge about the external world. From a Davidsonian perspective there is nothing wrong in letting our perceptions have a special position for our knowledge about the world. The problem, from the perspective of Davidson, is that Quine wants to give the causal influence from the world the status of epistemic justification for a belief. Davidson holds this to be impossible, and argues that only what is propositionally structured, i.e. a belief, can be a justification for something propositionally structured. What causally affects us is not propositionally structured, and cannot, according to Davidson, *justify* our beliefs.[24] This means that only what is or could be linguistically formed can count as justification of beliefs. The shift from sense perceptions to propositionally structured beliefs as justifiers for beliefs could seem to be a shift from a more naturalized to a less naturalized epistemology. It seems to be a turn away from non-normative empirical input to normative

[24] Roger Gibson argues against this in his "Quine on the Naturalizing of Epistemology", in P. Leonardi and M. Santambrogio (eds.): *On Quine*, Cambridge University Press, Cambridge 1995. He is offering an interesting argument for an interpretation of Quine that possibly could accommodate Davidson's requirement. See pp 96-98. However, it would take us too far off the track to give an account of this here.

and already-intentional input. If Davidson had upheld Quine's view concerning input as evidence, this would clearly be a problem, and Davidson's position would be one of idealism.

Even though radical translation was designed first and foremost for epistemological purposes, Quine also used his thought experiment for throwing light on language acquisition and for saying something about what meaning is *not*. In the article I point out that Quine eventually changed his mind about the proximal stimulus when it came to semantics, and agreed with Davidson that the distal stimulus were the relevant stimulus in this connection.

However, on the question of the *basis* for intersubjective similarity, Quine and Davidson differ yet again. While Quine holds that the similarity is based in pre-established harmony, and parallelism, Davidson holds that the similarity has its basis in common dispositions among humans, and in the exchange of linguistic descriptions of the world in the process of triangulation. My interpretation of Davidson in article 1 is neither congenial to idealism, nor to transcendentalism or empiricism. His position represents a non-reductive version of naturalism.

8. Naturalism and The Social

In article 2, "Davidson and the Role of Error" I took up Davidson's views upon non-conceptual aspects of experience, and especially the question of whether the non-conceptual can form a basis for knowledge. The approach to the problem was given from the occasion to which the paper were written, a symposium on "The Non-Conceptual Aspects of Experience" in Melbu in 2000. In the article I lay out Davidson's model of interpretation in triangulation, and I do this without presupposing that the reader is familiar with the model. There are examples that are of some value, and in the article I raise a number of issues that are treated more carefully in later articles. The article emphasizes that according to Davidson, the social relates to naturalism, and it stresses that the social is based in evolution and human nature.

When writing article 2 I took it for granted that in Davidson's model "the correction of wrong beliefs is going on all the time".[25] Later, when I wrote article 4, "The Explanation of Error", I had come to see that the kind of correction that I had been counting on, and, in particular, the correction from other speakers than the one interpreted, couldn't actually have a place in the model. In radical interpretation an interpreter interprets the language of a certain speaker at a given interval of time. A theory of truth for the speaker's language can be different from one day to the next, and any theory of truth for the language of a speaker will only relate to the particular speaker. The time-limited idiolect, structured in the form of a theory of truth, is the only form of a language that one can have in radical interpretation. My conclusion is that the radical interpretational process does not allow for social checks, and the social aspect is only made relevant in ostensive learning.

An important Davidsonian issue in a discussion of the role of the non-conceptual is the question of the conceptual versus the propositional. According to Davidson the conceptual is always already propositional, because we judge when we categorize parts of the world into this or that conceptual category. What is conceptual is also propositional. According to Davidson it is not "experience", but objects and events in the world that can cause beliefs in us. Beliefs are what knowledge consists of, and Davidson would say that conceptual knowledge is all that can be considered to be knowledge. I argue this means that Davidson has no place for something non-conceptual that can form a basis for knowledge, and neither can a notion of "experience" have a place in his theory - that is, if experience is supposed to be something intended as a basis for knowledge. Davidson would admit that there are non-conceptual aspects of our experience, but would most clearly deny that this can be said to form a basis for knowledge. It seems to me that the non-conceptual must be conceptualized in order to be relevant for conceptual knowledge in the interpretational model, and if it cannot be conceptualized, then it cannot be relevant, since it cannot be brought in as part of our network of belief.

I bring in the notion of error for the first time in article 2, and I connect it to Davidson's rejection of skepticism. I show how a notion of error is meaningful only where there are normativity or rules for correctness. This means that in order to employ a notion of error within the intentional, one must be able to connect it to truth and falsity, and to a notion of

[25] Article 2, p. 36.

belief. While belief is how we take things to be, truth is the way in which things are, and if we believe that things are different than they are, we have a false belief. We are thus mistaken, and we are in error. In order to be in error in this particular sense, one must be an intentional, linguistic creature, according to Davidson. In article 4 I have followed up this lead and I there provide a more thorough treatment of the problem of error.

Furthermore, I raise the question of how we can know whether or not a creature is capable of making error or mistake in our sense. If a creature is not in possession of a distinction between truth and falsity, it is not aware of the possibility of making a mistake of our kind. The difference is a difference between having a disposition and having a concept. Davidson has argued that awareness of the possibility of error is sufficient for ascription of thought. Simultaneously he has argued that exchange of propositional contents is needed in order for intentional creatures to be able to take advantage of the ability to triangulate their shared world. In the article I question the possibility that we can know that a creature is intentional unless it can disagree with us. Further investigation into this question is carried out in article 5.

In article 2, moreover, I defend Davidson's view on the relation between thought and world against criticism directed against this view by John McDowell. McDowell holds that Davidson's picture is that we cannot get outside our beliefs, and that Davidson thus is a coherentist at bottom. As long as Davidson admits only of the world's causal impact on us, McDowell holds, our empirical thinking is "engaged in with no rational constraint [] from outside"[26] This would imply that our interaction with the world is epistemologically insignificant. I argue that McDowell has overlooked the role that society plays in Davidson's model of triangulation. Thought and language are, according to Davidson, both interpersonal from the start, and without an upbringing in a society of language users, we would not come to be in possession of the awareness of error. Davidson and McDowell agree that if something should be in position to be a reason for a belief, this something must have propositional structure, and that "the relation of epistemic support requires that both relata have propositional content".[27] This means that (parts of) the world would have to be propositionally structured should it be in position to justify our beliefs about the

[26] Article 2, p. 42ff.
[27] Ibid, p. 46.

world. While McDowell affirms this, Davidson denies the possibility of a propositionally structured world.

In article 2 I stress the question of how the social relates to naturalism in Davidson's philosophy. I argue that the social, according to Davidson, is based in evolution and human nature, in the sense that we as humans are genetically disposed to discriminate the world in specific ways. Though this is the basis for discrimination, and thus for thought and language, we must, according to Davidson, also be taught to discriminate in more fine-grained ways in a linguistic society in order to be speakers. This means that *the basis* for our common ways of discriminating the world; ways that are socially (or intersubjectively) established, is our dispositions. The social establishment of belief and meaning, and thus, the cultural, has in this sense an empirical basis in human biology.

9. The Dialogical in the Social and Naturalism

Article 5, "Error and the Teacher-learner-object Triangle", emphasize the way that language and thought, and thus also the social, is genuinely dialogical. The article takes on a narrower approach to the problem of error than do any of the other articles. I address a situation of language learning and more specified, a triangular situation involving a full-fledged user of language who is communicating with a creature that is not yet a full-fledged language user. I'm asking what an explanation of *awareness of the possibility of* error (mistake from a person's own point of view) would amount to. An explanation of awareness of the possibility of error would be an explanation on the level of an idiolect, which means that it will be on the level of an individual language user. Awareness of the possibility of error in other minds, and the question of how to identify such awareness is therefore a more specified problem than the general problem of explaining error. How can we, from a third person point of view, justify to ourselves the difference between a learner who *is* and a learner who *is not* aware of error? And how could such a difference emerge in the triangular situation?

Davidson holds that creatures that intentionality inhabits are themselves aware of the possibility that their beliefs could be false, or in error, which means that they are aware that they could be deviating from the norm. He says that "The point is not to identify the norm, but to make sense of there being a norm, and this has been done if we can point to the differences between the preponderance of cases where the creatures respond alike and the deviant cases where they diverge. I insisted that it is not sufficient that a third party is able to observe or describe these two cases; I have claimed it is necessary that the existence of the contrast is available to the creatures themselves. This essential element enters the triangle when the creatures observe each other's reactions to the very phenomena they are observing".[28] This means that in order to ascribe intentionality and a grasp of the distinction between truth and falsity to another being, we have to know, or at the least we have to have some reasons for assuming, that the creature has a mind, and that he or she is able to read other minds. Both parties in the triangle thus have to possess awareness of the possibility of error. However, we do not find *evidence* that the other has a mind in the triangular situation itself. What we do find in triangulation are necessary *conditions* for creatures to be thinkers and mind readers. When or how could we then produce evidence that the other creature is thinking? Davidson assumes that when the other creature speaks such that we can understand, and thus predicates, we have a sufficient reason to assume that it has a mind. When we respond to the other's utterances and he or she can respond back such that we have good reasons for assuming they have understood us, we also have good reasons to assume that they are mind readers. According to Davidson, smooth communication is a criterion. But are these criteria, that of predication and that of the other creature's predicated response to my predication, sufficient for having evidence that the other creature has the awareness of error? The question is how we can identify awareness of error.

First, there is the question of the emergence of awareness of error for a creature that still hasn't got it, but whom we assume will be disposed for developing such awareness. I assume that a learner will be unable to become aware of the possibility that something predicated by him or her could be wrong only on basis of corrections from the teacher. My suggestion is to introduce a second teacher, such that the teacher-vertex includes two full-fledged speakers in discussion. Without a second teacher the learner will hold the teacher to be infallible, and the teacher will not be in position to function as a role model for the learner when it comes to awareness of error. When two people discuss, I maintain, it is

[28] Donald Davidson: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001, p. 5.

more likely that a learner will be in position to understand that one of them can be wrong, and I argue that this is the model for the assumption of error in oneself.

When two people discuss, for instance, what the right description of some phenomenon is, the child is able to understand that error of the meaning-kind is possible, as well as error of the belief-kind. When the two (for the child) authoritative persons use words in different ways, the child can get the idea that employment of norms is up for discussion. Admission to the disagreements of several authoritative persons is important for the learner. Here I presume that disagreement over norms can be apprehended before one is in possession of norms. Understanding that there are several possible understandings of the same phenomena means getting a grip on a hypothetical relation to world and language.

My second suggestion in article 5 is intended to meet the demand for evidence of the learner's awareness of the possibility of error. The suggestion is that in order for the teacher to know that the learner is in possession of error, the learner must be able to disagree with the teacher. He must be able to doubt and question the teacher's account. In order to be certain that this is what the learner does, the learner must be able to give descriptions and accounts that can compete with accounts given by the teacher.

Now, as long as Davidson demands the awareness of error to be internal to the creatures themselves in order for them to be in possession of error, one could always say that evidence from a third person point of view is not required. The criterion could in this way be said to be internal to each and every creature, and one could argue that behavioral evidence of awareness of error is therefore irrelevant. This is obviously a respectable stand, but the question is whether or not this is consistent with Davidson's overall strategy for epistemology. According to Davidson, the study of how we acquire knowledge should be carried out from a third person perspective. If this strategy is to be carried out in the triangular situation with teacher and learner, the teacher must, from his point of view, be able to effectively recognize the difference between learners in possession of the awareness and those without. This means that my suggestion takes this methodological predicament at face value, and pushes the methodological strategy further than Davidson's original suggestion. I will maintain that whatever the result may be, we will learn something from it. We could either find that the strategy has its limits in the sense that it becomes meaningless in actual cases, or that we must strengthen the methodology as a strategy for future studies. I maintain that my attempt here shows we can go on with the strategy, and that it is a useful tool that we can strengthen and apply to a wider range of cases.

There is also another way of approaching the question of awareness of error. Davidson holds that unless a creature has beliefs, neither can he make a mistake, from his own point of view. In his "Problems in the Explanation of Action", Davidson mentions an example with a tribe of monkeys.[29] The members of this tribe respond to the threat of danger by emitting a certain cry, and other members respond as to danger when they hear this cry. If now a monkey shouts out the danger cry when no danger is present, and another monkey reacts as to danger, *we* would surely say that a mistake occurred. Davidson's point here is that "...unless the monkeys *believe* there is danger when there is not, no error has been committed; they have simply responded to a stimulus that usually, but not always, accompanies danger." He thus holds that unless we can say that intention and belief are present, no concept of mistake can apply. In this regard I have said that we cannot be in a methodological position to ascertain ourselves that intention and belief are present in a creature unless we can communicate with it. And Davidson's point is that we don't have to attribute intentions or beliefs to the monkeys in order to explain their behavior. But as I read him, his point is *not* that the monkeys don't have beliefs. We simply cannot know whether or not they do have beliefs. Neither do we have to ascribe certain beliefs to them in order to understand their behavior. The concept of mistake does not, then, apply to a creature that could not itself be aware of the possibility of error. Another of Davidson's ways of putting this is the following: "So far as I can see, no account of error that depends on the classifications we find most natural, and counts what deviates from such as error, will get at the essence of error, which is that the creature itself must be able to recognize error. A creature that has a concept knows that the concept applies to things independently of what it believes."[30]

A theory of knowledge that can give an account of "thought as objective" has to give an account of error. This is a prerequisite for us to be able to distinguish between having a

[29] Donald Davidson: "Problems in the Explanation of Action" (1987) pp 101-116 in Donald Davidson: *Problems of Rationality*, Oxford University Press, New York 2004, p. 116.
[30] Donald Davidson: "What Thought Requires" (2001) pp 135-150 in his *Problems of Rationality*, Oxford University Press, New York 2004, p. 141.

concept (and thereby being able to recognize an error), and having a disposition. Creatures with only dispositions does not have a concept of objectivity, a distinction between truth and falsity, and a distinction between going on in correct versus incorrect ways. Such a theory will also have to give some directions showing *how* we can identify a mistake *as* a mistake. (One of the reasons for this is that we should expect a theory of the Davidsonian brand to give behavioral criteria for its theoretical entities / phenomena.) Telling how to identify a mistake *as* a mistake must be one of the aims of the theory, but this need not be more mysterious than explaining the reasons for our beliefs and thus also the reasons for specific mistakes.

"We are built to discriminate objects, to keep track of them, expect them to emerge from their holes or behind trees, and in some cases to feed or eat us. But these finely tuned abilities are not the same as thinking of things as objects and events. This demands the apparatus of propositional thoughts with truth conditions and the awareness of possible error. I don't think of the awareness of possible error as an abstract cautionary warning; it takes shape rather in the light of the reasons we have and can give for our beliefs, and therefore the explanations we can give of specific mistakes. Perception is propositional: when we look or feel or hear we believe."[31]

This also means that an account of error needs to be a part of a general unified theory of meaning and action. And here we have reached a simplification of my approach to the problem of error. Throughout my work on the articles I have related to Davidson's theory of interpretation in the form of a theory of meaning. Instead of the ascription of truth conditions to the speaker's individual sentences, as the theory of meaning instructs the interpreter to perform, the interpreter is in the unified theory of meaning and action to look for (the assumed observable notion of) the speaker's valuing the truth of one sentence more than the truth of another. Therefore I have not commented upon the ways in which my approach to the problem of error can work within a unified theory or not. As noted in the introduction, I have in general assumed that the extension of the theory is inessential to the problems that I am dealing with. However, in order to answer the question of error within the unified theory properly, one would have to pursue analyses further.

[31] Davidson in his "Reply to Dagfinn Føllesdal" in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, p. 731.

In article 5 there are two demands, or methodological assumptions, that we can directly relate to a Davidsonian version of naturalism. The first is the demand that we should be able to tell – from a third person point of view – whether or not the speaker has awareness of error. The general demand for a third person approach to epistemology is by Davidson spoken of as one out of two features of Quine's naturalized epistemology he wishes to continue to pursue. My demand in this article, connected to the first suggestion, is therefore clearly a demand that is naturalized in spirit, according to this mentioned understanding of naturalized epistemology. The other feature in article 5 that relates to this version of naturalism is the idea that disagreement (related to my second suggestion) is a condition for having a hypothetical relation to world and language. Such a hypothetical relation to world and language is another side of fallibilism, which I have deemed a core naturalistic conviction.

Now, we have summed up many of the main points, arguments and results of the 6 articles here presented, and more or less explicitly related it to naturalism and naturalized epistemology. I will not provide a general summary of what naturalism can be in Davidson's approach to knowledge, as I hope it is has become sufficiently clear by now. But a crucial point deserves focus, and that is the question of whether there are structures, conditions and presuppositions in Davidson's philosophy that are non-congenial to his version of naturalism. According to my analyses, there can be no structures, conditions or principles with a fundamentally different epistemic status within his philosophy, and this means that all structures, conditions or principles will have to be gradual and fallible. It thus seems that Davidson's philosophy has a clear profile in terms of epistemology.

There is, however, another feature in Davidson's thinking that has come to seem less congenial to his main lines of thought: the notion of truth. I have asked whether this is due to the metaphysical pretensions of truth, and this is in particular so for the notion of objective truth. The idea of objective truth seems to be transcending a naturalized approach to knowledge. Another way of talking about the difference between a metaphysically innocent notion of truth (here; just "truth") and a more heavily metaphysically loaded notion of truth (here; "objective truth"), is the distinction between "believed true" and "true". Thus, in the final three chapters of this synthesis we will look further into this question.

10. The Gap Between "Believed True" and "True"

In more recent publications Davidson has admitted to a connection between what he has termed "the distal theory of reference"[32] and the problem of explaining error. In his posthumously published *Truth and Predication*, he talks about this as "the crucial gap between what is believed true and what is true" and notes that the problem is crucial "since the distal theory directly relates truth to belief".

Davidson points out that "The ultimate source of both objectivity and communication is the triangle that, by relating speaker, interpreter and the world, determines the contents of thought and speech. Given this source, there is no room for a relativized concept of truth."[33] This means that Davidson holds that a non-relativized concept of truth, presumably then a concept of "objective truth", is to have its source in triangulation. He continues by stating that truth must also be related to "the attitudes of rational creatures" and that this relation springs from the nature of interpersonal understanding, in short; it is to be found in mutual interpretation. "Linguistic communication, the indispensable instrument of fine-grained interpersonal understanding, rests on mutually understood utterances, the contents of which are finally fixed by the patterns and causes of sentences held true."[34] In one sense, it *could* seem that the determination of causes of beliefs taking place in triangulation is, anyhow, the ultimate source for the concept of objective truth, according to Davidson's most recent text. He further notes that the "conceptual underpinning of interpretation is a theory of truth" and that truth therefore rests "in the end on belief". I take these statements to indicate that a concept of objective truth still relies on *both* its external constitution through triangulation and its internal constitution through the circle of interdefined intentional notions. But Davidson also points to truth in the form of an empirical theory of truth for speaker S in a language L at a time T, and such a theory can only be realized if believers exist in the first place. So, for one, *truth* will only exist where believers exist, and two, it seems the *objectivity* of truth ultimately rests on the determination of the contents of thought in triangulation. So, there is still somehow a gap between believed true and what is true. Hence, it seems that we, with Davidson's theory, have a good grip on "believed true" trough interpretationism, and this is what makes truth

[32] Donald Davidson: *Truth and Predication*, Harvard University Press, 2005, article 3, "The Content of the Concept of Truth", p. 64 and 65.
[33] Ibid, p. 75.
[34] Ibid.

interesting for us as knowers, but, on the other hand, "what *is* true" seems to be only *partially* secured trough interpretation in triangulation. I have argued that the reason for this is that the social check cannot be performed in interpretation in the triangular mood.[35]

However, as I noted in article 4, we could either give up the quest for objective truth, or we could give up the idiolect as the basis for interpretation. There are many reasons why Davidson could not give up the idiolect. For one, it is intrinsic to the model of radical interpretation, and it is intrinsic to Davidson's theory also in certain other respects. Time-limited idiolect as the notion of language to relate to in radical interpretation is connected to the thesis of the indeterminacy of translation. When we as interpreters restrict ourselves to a person and a time for interpretation, and we also operate alone as interpreters, we avoid the confrontation with the indeterminacy of translation. *Another interpreter* would be able to come up with another theory, we would on *another occasion* be able to come up with another theory about the beliefs and meanings of *the same speaker*, and we would certainly be able to come up with another theory for *another speaker* at *the same time*. This is why we have to make radical interpretation local, since going global means anarchy, that is, when we take it on the premises of interpretationism. The thesis of the indeterminacy of translation, which I believe is a crucial insight we have no choice but to accept, is therefore a second and separate reason why we could not give up the idiolect without devastating changes to the model of radical interpretation. It thus seems that there still has to be a gap between "believed true" and "true".

In order to look at the problem from another quarter, let's look at the basic situation of radical interpretation. In radical interpretation we start out with an assumption about what *some specific speaker* holds true and ends up with a theory of truth for the *language* he or she speaks, a theory of truth which serves as an empirical theory of meaning for the language. We start out knowing that:

(A) Kari holds true-in-L "It is raining" if and only if it is raining in her vicinity,

[35] In article 5 I suggested that an extension of the teacher-learner-object triangle with a third participant, a second competent person, would solve the problem of explaining how a learner could get the idea that error is possible. The possibility for a general extension of the triangle with two speaker-interpreters *could* be a way out of the predicament. However, in this synthesis there is no room for further research on this question.

This being our only "evidence" for the T-sentence:

(B) "It is raining" is true-in-L if and only if it is raining in the vicinity of the speaker [36]

The question is, of course, why we are entitled to infer (B) from (A), if we assume that the language L is a language such as Norwegian or German; in short, if we assume an ordinary everyday use of the term language. Even if it is true that Kari holds true the specific sentence (due to what she believes and what the sentence means) under the appropriate circumstances, how can we know that she is not mistaken? It seems that even if Kari believes a truth, sentence (B) could still be false. The inference from the fact of a speaker holding the sentence true in L to the fact of the sentence being true in L needs support, or at the least we have to explain how we got from (A) to (B). The problem has basically two "solutions": one is the principle of charity and the other is causality in triangulation. The principle of charity lets me assume that the speaker has the same beliefs as me (the interpreter), and since I believe the obvious, that it is raining when I see or feel rain, I assume that so does the speaker. In the situation of triangulation I can observe that the speaker is being caused to believe what she believes (which she then utters in a sentence) by the relevant object, event or situation present, and I can check the correctness of my ascription by further questions. But, as I have argued, the interpreter cannot reach out to the objective truth and thus neither to the intersubjectively established linguistic norms trough radical interpretation in triangulation.

However, according to Davidson, the interpreter can take in public standards in the form of a "plurality of private belief structures". (Ref) This is, according to him, "What makes a social theory of interpretation possible" and fortunately, "belief is built to take up the slack between sentences held true by individuals and sentences true (or false) by public standards." (Ref) But can belief solve the problem of error? Can it bridge the gap between idiolect and the correction or input from society? It does not seem so, since a plurality of private belief structures is not what the interpreter has to go on in *radical* interpretation. If interpretation were non-radical this could perhaps have been a possible solution.

[36] The phrasing of the example is borrowed from Simon Evnine's *Donald Davidson*, Stanford University Press, Stanford 1991, p. 101.

We could ask if there is a problem, from Davidson's point of view, in going from interpretation of idiolects or sets of idiolects to interpretation of whole languages? Is there even an entity that can, in an ordinary, everyday sense, be called a language in Davidson's theory? After all, "...there is no such thing as a language..." was the rhetorical punch line of his "A Nice Derangement of Epitaphs".[37] The concept of language in use is, however, quickly specified as "a clearly defined shared structure which language-users acquire and then apply to cases".[38] So, there is no such thing as a clearly defined structure to be acquired and then applied to cases. Rules are not static, objective or conventional. Rules are not what fix meanings. Or, to restate it with another of Davidson's articles, there is no scheme-content division to be had simultaneously with a view of the world as given. Schemes are not relative to content and concepts are not relative to scheme - at the least not while a constant and given world is simultaneously presupposed. Just as rules aren't fixing meanings, schemes aren't fixing worlds. Rules as well as schemes are in a state of relative and relatively slow flux. But, even though there is no language in the form of a clearly defined structure that could be acquired and applied to cases, it seems that we have to reach out for a concept of language that includes more than two speakers.

Davidson has a thesis concerning the impossibility of massive error about the world (which includes a thesis about the impossibility of massive error in our empirical beliefs), a thesis that has had a somewhat unclear status within Davidson's own thinking. He never had a watertight argument for the thesis, and it came to the fore due to work on other problems. In the beginning (especially in his "The Method of Truth in Metaphysics" from 1977) the thesis rested partly on the necessity for applying the principle of charity in interpretation.[39] This means that the principle of charity was a vital, although partial, basis for anti-skepticism about the world. Later, in the eighties and the nineties, with the development of a conception of triangulation, the thesis came to rest more on the circumstance that the content of our beliefs can be individuated based on their normal causes, in such a way that we have something close to a warrant that most of our empirical beliefs are true. This means beliefs are partly constituted by the causes to those beliefs. Simultaneously, the function of the principle of charity as constitutive for beliefs and for the thesis of the

[37] Donald Davidson: "A Nice Derangement of Epitaphs" pp 433-446 in E. LePore: *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford 1986, p. 446.
[38] Ibid.
[39] Donald Davidson: "The Method of Truth in Metaphysics" (1977) pp 199-214 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

impossibility of massive error is upheld. On the other hand, there is also the thesis that most of a speaker's beliefs are true; hence, truth in belief. The question is, again, whether a thesis about the impossibility of massive error, or alternatively, a thesis about the massive truth of the beliefs of speakers (whether this follows from the principle of charity or from triangulation) is consistent with Davidson's thorough hypothetical relation to the world. I would say that, in Davidson's approach, the closest we can come to truth is the ascription of "holding true" to a speaker, combined with the interpreter's following competent, but hypothetical assumption that what is *held* true, *is* true. This means that the speaker's and the interpreter's gradual subjective beliefs about their own beliefs is the arbitrator of the question of the extent to which something is true. The basis is, of course, roughly secured through mutual interpretation of common surroundings, but the truth of beliefs, in an absolute manner, cannot be secured through the methodology of interpretation.

11. Recapitulation of Conditions for Objective Truth

What does an externalist theory of knowledge have to do to show that there is an intrinsic connection between thought and the world (such that it can show that thought is objective)?

According to Davidson it has to:

1. Give an account of the contents of perceptual beliefs
2. Give an account of when two thoughts are the same
3. Give an account of how the content (however it is determined) is recognized as the same from one occasion to another
4. Give us a way to isolate the right cause, the cause of the perceptual belief. (The normal cause.)
5. Give an account of error, so that we can distinguish between having a concept and having a disposition
6. Tell us how we can identify a mistake as a mistake.

These are obviously what Davidson takes to be the necessary and sufficient conditions that an externalist theory of knowledge has to fulfill in order to show that there is an intrinsic connection between thought and the world, and to thus show that thoughts about the world can be objectively true. Which out of these conditions has his theory met, and which has the theory not been able to meet? He seems to have been able to give an account of (1.) the contents of perceptual beliefs, as he hold that the content is directly related to the objects and events that caused our perceptual beliefs, and we can identify their contents partly on this basis, and partly based on the location of the belief within the net of beliefs. How we know that (2.) two thoughts are the same is accounted for by the methodology of interpretation, and it is the principle of charity that describes how this is done. We take our own system of beliefs as a model, and we then hold the speaker's beliefs constant while solving for meaning. That (3.), the content is recognized as the same from one occasion to another, is accounted for by our dispositions to categorize empirical similarities similarly within the human race. This is something that we just do, as a statistical matter of fact, and it is therefore not normative. The way of isolating the right cause (4.) is the way of triangulation. Davidson seems to sometimes say that we *do* manage to isolate the right cause by crossing causal lines, but more often he seems to hold that we can do this *only* when we are speakers and thereby can correlate each other's reactions. The basis for our success in isolating the right cause is our disposition to categorize the world in similar ways, and our triangular social activity comes as a necessary "second step" (explanatorily, it is not a two-step process in time). Anyway, as we *are* speakers, we can take advantage of our ability to share reactions.[40] According to Davidson, he has not been able to (5.) give a complete account of error, and (6.) neither has he identified a way in which we can identify a mistake as a mistake. However, as concerns these two last conditions, he has pointed to society in the form of the second person.

I do share Davidson's intuitions about the solution to the problem of the externalist theory of knowledge. However, in the list of conditions to be fulfilled, Davidson does not list the demand that the theory can tell us how we can identify a *non-mistaken* belief. This demand could either be read as a demand that the belief is justified (that we have evidence for it) *or* that it is true. The reason why Davidson has not listed this demand together with the others is that he would say that the truth of a single belief couldn't be guaranteed, even though we can know that most of the beliefs of a speaker must be true. The demand for truth cannot be objective as seen from the inside of radical interpretation. The problem is, however, that in order to cash out a notion of knowledge we must demand minimally that the belief is

[40] Not all would agree that Davidson has been able to give a way to isolate the right cause of a conceptual belief. Kathrin Glüer and Ingar Brinck have on some occasions raised such misgivings, but I will here not go into their reasons for saying this.

justified, or, maximally, that it is true. And Davidson demands objective truth. He would hold that the process of content-determination gives as good a justification as we could get in this direction. But it does not give objective truth. And objective truth is most certainly what Davidson years for. Quine doesn't, and thus never run into this problem.

12. Quine, Davidson and Knowledge

Quine has rejected the view that knowledge is a mental state.[41] In opposition to this view, Quine holds that "knowing is a hybrid of warranted belief, which is mental, and truth, which is not."[42] Now, Davidson has no notion of "warranted belief". His notion of belief is instead a notion in which the warrant is intrinsic to the very concept. Davidson talks about this as "the veridicality of belief", which means that it is the nature of belief to be veridical, or "probably true". The veridicality of belief is, however, not present at the level of individual beliefs; it is rather a feature of a net of beliefs that most of the beliefs in it must be true. As far as I am able to figure out, Davidson uses the notion of "knowledge about the world" in the same ways as he uses the notion of "belief about the world". If this assumption of mine is correct, evidence, or warrant, is not a feature that can be analytically extracted from Davidson notion of knowledge, since it cannot be extracted from the notion of belief. But is it otherwise probable that Davidson could subscribe to Quine's view on what it is to know? Truth could not be reduced to warranted belief, according to Davidson. Davidson would agree that what is known must be true, and he holds that there must be some connection between belief and truth. More specifically, he holds that "belief can serve as the human attitude that connects a theory of truth to human concerns."[43]

It seems to me that Quine's expression of knowing as a hybrid goes well along with interpretationism. Thus the notion of knowledge has two sides: an internal consisting of what Quine labels mental, and which Davidson simply would call a belief, and an external side being truth. Truth is indeed not mental, even though Davidson *has* shown how to constitute the notion of truth partially from the inside of intentionality. Truth is partly internal, according to my distinction between the internal and the external side of

[41] In "Response to Gibson" in Alex Orenstein and Petr Kotatko (Eds.) *Knowledge, Language, and Logic: Questions for Quine*, pp 414-415, Kluwer Academic Publishers, Dordrecht 2000.
[42] Ibid, p. 415.
[43] Davidson, Donald: "Epistemology and Truth", (1988) pp 177-191 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001, p. 189.

intentionality, but it is not mental. Truth is, however, *also* partly external. The notion of truth, and indeed the notion of objective truth, is based on the attempt at an explanation of the notion itself via processual aspects of triangulation. I have argued that the model of triangulation in radical interpretation suffers from the absence of continual correction via the society since the societal influence can only come in by ostensive learning. This is due to radical interpretation's restrictive concept of language, the idiolect, and the connection between this restriction and the external constitution of a notion of objective truth. In order to establish a notion of objective truth via the model of triangulation, we need to be able to perform a social check for error. My contention hence is that this is impossible due to the methodological restrictions laid down in the individually oriented model of radical interpretation. We thus need a community-based notion of language, or alternatively, we could give up objective truth.

Davidson assumes, however, that an account of knowledge can be built upon an account of belief. An account of belief, or a theory of belief, is the fallout of the theory of meaning, which is given in the form of a theory of truth for a given speaker at a given time, as interpreted by a fellow speaker. Thus, according to Davidson, knowledge, like belief and meaning, is dependent on linguistic interaction between speaker-interpreters.[44] The theory of belief for a speaker, or a speaker's net of beliefs, is hence to be the basis for an account of knowledge. But how could such an account aspire to a substantial concept of knowledge without a substantial concept of truth? And, perhaps Davidson should be content with giving an account of belief. The question is thus whether Davidson is able to meet the demands an account of knowledge on his own terns would require. A description of the gradual confidence in our own and other people's belief, and other people's gradual confidence in their own and our belief, is of vital importance for understanding how we as humans relate to the world. And this might be all we can give in terms of a theory of

[44] In a short piece called "On Quine's philosophy" in *Theoria*, Vol. 60, 1994, pp 184-192, Davidson raises the question of consequences of the by Quine admitted irrelevance of stimulus meaning to translation. Davidson asks; "how are we now to think of the relation between theory of meaning and epistemology?" Davidson suggests that Quine, along with himself, "accept the account of belief that [] necessarily falls out of Quine's theory of meaning, and build the account of knowledge on this basis. In this case thought and knowledge, like meaning, would depend on linguistic interaction between agents." (p. 192) This advice from Davidson is of course based on the fact that Quine eventually had admitted that distal stimuli should be the basis for translation and language learning, while he still wanted to stick to proximal stimuli in the case of his epistemological study.

knowledge, at the least when we base ourselves on a methodological frame of interpretation.

13. Normativity and Regularity of Behavior

Triangulation takes away the illusion that the one part has got something that is definite, some "static semantic competence"[45]; in other words, that the one part has the definite key that the other one is trying to decipher. This is the reason why regularity of behavior is not equivalent to normativity. And this is the answer to Richard Rorty's question "What could normativity be, except regularities of behavior?"[46] I would say that the mental is normative because regularities of behavior can and will be interpreted in different ways by different people, and there is no way around this "fact of interpretation".[47] The reason why intentional action is normative is that we have to employ our own norms in order to grasp the attitudes of those creatures that we try to understand. That thoughts are assessable is a hindrance for saying that normativity *is* regularity of behavior.

What would Davidson say about Rorty's question? As I read him, normativity is not something *in addition to* action, causality, behavior and regularity of behavior. Normativity is all about making sense of intentional behavior; it is ultimately about the meaning that lies in such behavior and in this activity we cannot reach outside our own norms. My contention is that a Davidsonian understanding of the question of the relation between normativity and behavior is such that normativity is supervening on behavior. This means that regularity of behavior is necessary for normativity to emerge, but it does not mean that normativity is nothing but patterns of behavior.

It is also meaningful to ask what one possibly can mean by the term "behavior" as it is applied in the initial question that Rorty posed. Behavior as not interpreted or non-normative would amount to "movement" or "motion". When asking such a question one seems to presuppose that behavior is something that can be described in physicalist

---

[45] This is Bjørn Ramberg's formulation of it, in his *Donald Davidson's Philosophy of Language*, Blackwell, Oxford 1989, p. 78.

[46] Rorty posed this question as a comment upon my paper "Davidson, Naturalism and Constitutivity" in the seminar "Issues in Contemporary Metaphilosophy" in Tromsø, December 2001. It has haunted my thoughts ever since.

[47] This is one of the reasons why there will always be an indeterminacy of interpretation.

vocabulary, and of course it can. But he who asks such a question would also presuppose that human behavior is merely (or first and foremost) physical movement, and this is not compatible with Davidson's view.

What makes mere cow-reactions into a concept of cow? According to Davidson the transition takes place when the base line becomes as thick as language. I will answer Rorty's question on basis of what I have learned from my work on these articles. If there is regularity of behavior, assessed by an interpreter on basis of relevant similarity between speaker's reactions and interpreter's reactions, then assessment is already in the picture. Assessment that there is or is not regularity demands that we ascribe certain beliefs to the speaker, and this could not be done without also ascribing meanings, and we have therefore interpreted the speaker already when we judge that there is regularity of behavior. This demands that we have the ability to think about the beliefs of the others and ourselves. Mere cow-reactions can be regularly observed, but this cannot explain the *emergence* of norms. The norms are already there when we can observe something as regularity. In order to reach for *conditions* for normativity, we must go beyond the observations of regularity. Regularity of behavior in speechless creatures does fulfill two out of the three conditions that Davidson has pointed out for development of thought and language. The dispositions to categorize in certain ways, which are common for a species, are the first condition that is fulfilled. Next, there must be a statistically based normality present. This condition is fulfilled when we can observe regularity. The third condition is the possibility of error and an awareness of the possibility of error. If the creature in question produces a cow-reaction when no cow is present, we would normally ascribe an error to the creature, and we would perhaps say that his utterance were false. But as long as the creature is unaware of the possibility that something he utters can be a mistake, there is no way for us, as interpreters to actually operate with a notion of error, mistake or falsity, on a methodologically sound basis. We really would have no idea of why the creature uttered a cow-reaction when no cow was present.

References

Boghossian, Paul and Christopher Peacocke (Red.): *New Essays on the A Priori*, Oxford University Press, Oxford 2000

Brandom, Robert B.: "Knowledge and the Social Articulation of the space of Reasons" in *Philosophy and Phenomenological Research*, Vol. LV, No. 4, 1995.

Child, William: *Causality, Interpretation and the Mind*, Oxford University Press, 1994.

Davidson, Donald and Patrick Suppes: "A Finitistic Axiomatization of Subjective Probability and Utility" in *Econometrica*, Vol. 24, 1956, pp 264-75.

Davidson, Donald, Patrick Suppes and Sidney Siegel: *Decision Making, an Experimental Approach*, Stanford University Press, Stanford 1957.

Davidson, Donald: "Truth and Meaning" (1967) pp 17-36 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "Mental Events" (1970) pp 207-227 in *Essays on Actions and Events*, Oxford University Press, Oxford 1980.

Davidson, Donald: "Radical Interpretation" (1973) pp 125-139 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "Belief and the Basis of Meaning" (1974) pp 141-154 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "On the Very Idea of a Conceptual Scheme" (1974) pp 183-198 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "Psychology as Philosophy" (1974) pp 229-224 in *Essays on Actions and Events*, Oxford University Press, Oxford 1980.

Davidson, Donald: "Thought and Talk" (1975) pp 155-70 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "The Method of Truth in Metaphysics" (1977) pp 199-214 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "The Inscrutability of Reference" (1979) pp 227-242 in *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "Towards a Unified Theory of Meaning and Action" in *Grazer Philosophische Studien*, 11, 1980.

Davidson, Donald: "Rational Animals" (1982) pp 473-80 in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Blackwell, Oxford, 1985.

Davidson, Donald: "A Coherence Theory of Truth and Knowledge" (1983) pp 307-319 in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore, Blackwell, Oxford 1986, and pp137-153 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

Davidson, Donald: *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford 1984.

Davidson, Donald: "A Nice Derangement of Epitaphs" pp 433-446 in E. LePore: *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford 1986.

Davidson, Donald: "Afterthoughts", (1987) pp 154-157 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

Davidson, Donald: "Problems in the Explanation of Action" (1987) pp 101-116 in Donald Davidson: *Problems of Rationality*, Oxford University Press, New York 2004.

Davidson, Donald: "Epistemology and Truth", (1988) pp 177-191 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

Davidson, Donald: "Epistemology Externalized" (1990) in *Dialectica* 45 (1991) and pp 193-204 in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

Davidson, Donald: "The Structure and Content of Truth" (1990) *The Journal of Philosophy*, 87 (6), 1990, pp 279-328 and pp 1-75 in *Truth and Predication*, Harvard University Press, Cambridge, Mass. 2005.

Davidson, Donald: "Three Varieties of Knowledge" (1991) pp 153-66 in Phillips Griffiths, A.: *A. J. Ayer: Memorial Essays*. Royal Institute of Philosophy Supplement: 30, Cambridge University Press, 1991, and in *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

Davidson, Donald: "Reply to Jerry Fodor and Ernest LePore", pp 77-84 in Stoecker, R. (ed.): *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers*, De Gruyter, Berlin 1993.

Davidson, Donald: "On Quine's philosophy" in *Theoria*, Vol. 60, 1994, pp 184-192.

Davidson, Donald: "Pursuit of the Concept of Truth", in Leonardi, P. and M. Santambrogio (eds.): *On Quine*, Cambridge University Press, Cambridge, Mass., 1995.

Davidson, Donald: "The Problem of Objectivity" (1995) pp 3-18 in *Problems of Rationality*, Oxford University Press, New York 2004.

Davidson, Donald: "Externalisms" (1996) pp 1-16 in *Interpreting Davidson*, ed. P. Kotatko, P. Pagin and G. Segal, CSLI Publications, Stanford 2001.

Davidson, Donald: "The Centrality of Truth" (1996) pp 105-15 in *Truth and its Nature (if any)*, ed. by J. Peregrin, Kluwer, Dordrecht 1999.

Davidson, Donald: "Gadamer and Plato's *Philebus*", pp 421-432 in Hahn, L. Edwin: *The Philosophy of Hans-Georg Gadamer*, The Library of Living Philosophers, Open Court, Chicago and La Salle1997.

Davidson, Donald: "Seeing through language" pp. 15-27 in *Philosophy: the journal of the Royal Institute of Philosophy*, Cambridge University Press, Cambridge 1997.

Davidson, Donald: "Replies" in *Critica* 30, 1998: pp 97-112.

Davidson, Donald: "Intellectual Autobiography", pp 3-69 in Hahn, Lewis Edwin (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Davidson, Donald: "Reply to J.J.C. Smart", pp 123-126 in Hahn, Lewis Edwin (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Davidson, Donald: "Reply to A.C. Genova" pp 192-194 in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Davidson, Donald: "Reply to Carol Rovane" pp 480-482 in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Donald Davidson: "Reply to Dagfinn Føllesdal" pp 729-732 in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Donald Davidson: "Interpretation: Hard in Theory, Easy in Practice", in Mario De Caro (ed): *Interpretation and Causes. New Perspectives on Donald Davidson's Philosophy.* Kluwer, Dordrecht 1999.

Davidson, Donald: "Comments on Karlovy Vary Papers", in Petr. Kotatko, Peter Pagin and Gabriel Segal (eds.): *Interpreting Davidson*, CSLI Publications, Stanford 2001.

Davidson, Donald: *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford 2001.

Davidson, Donald: "What Thought Requires" (2001) pp 135-150 in his *Problems of Rationality*, Oxford University Press, New York 2004.

Davidson, Donald: "Quines epistemologier", in *Norsk filosofisk tidsskrift*, nr 1-2, 2004, translated by Eivind Balsvik from the English "Quine's epistemologies".

Davidson, Donald: *Problems of Rationality*, Oxford University Press, New York 2004.

Davidson, Donald: *Truth and Predication*, Harvard University Press, 2005.

Davidson, Donald: "The Content of the Concept of Truth", chapter 3 of his *Truth and Predication*, Harvard University Press, 2005.

Evnine, Simon: *Donald Davidson*, Stanford University Press, Stanford 1991.

Frith, Uta and Francesca Happé: "Language and communication in autistic disorders" in *Philosophical Transactions of the Royal Society*, Series B, 346, pp 97-104, London, 1994.

Føllesdal, Dagfinn: "Triangulation" in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, pp 719-728.

Gibson, R.: "Quine on the Naturalizing of Epistemology" in Leonardi, P. and M. Santambrogio (eds.): *On Quine*, Cambridge University Press, Cambridge 1995.

Glüer, Kathrin and Peter Pagin: "Meaning Theory and Autistic Speakers" in *Mind and Language*, Vol.18 No. 1 February 2003, pp 23-51.

Hahn, Lewis Edwin (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998.

Hahn, Lewis Edwin (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999.

Hanson, Norwood Russell: *Patterns of Discovery*, Cambridge University Press, London 1958, reprinted 1969.

Heal, Jane: "Radical Interpretation", in Bob Hale and Crispin Wright: *A Companion to the Philosophy of Language*, Blackwell, Oxford 1997.

Janvid, Mikael: *Naturalism and the status of epistemology*, Preprint no. 7, Department of Philosophy, Stockholm University, Stockholm 2001.

Kornblith, Hilary: "What Is Naturalistic Epistemology?" in Kornblith, Hilary (ed): *Naturalizing Epistemology*, The MIT Press, Cambridge, Mass. 1994, pp 1-14.

Kotatko, Petr., Peter Pagin and Gabriel Segal (eds.): *Interpreting Davidson*, CSLI Publications, Stanford 2001.

Kripke, Saul: *Wittgenstein on Rules and Private Language*, Oxford University Press, Oxford 1982.

Leonardi, P. and M. Santambrogio (eds.): *On Quine*, Cambridge University Press, Cambridge 1995.

McDowell, John: *Mind and World*, Harvard University Press, Cambridge, Mass. 1994.

Orenstein, Alex and Petr Kotatko (Eds.) *Knowledge, Language, and Logic: Questions for Quine*, Kluwer Academic Publishers, Dordrecht 2000.

Picardi, Eva: "Sensory Evidence and Shared Interest", pp 171-185 in Mario de Caro: *Interpretations and Causes*, Kluwer Academic Publishers, Dordrecht 1999.

Quine, Willard Van Orman: "Two Dogmas of Empiricism", p. 42-46 in *From a Logical Point of View*, Harvard University Press, Cambridge, Mass. 1953.

Quine, Willard Van Orman: *Word and Object*, The MIT Press, Cambridge, Mass. 1960.

Quine, Willard Van Orman: "Epistemology Naturalized" pp 69-90 in *Ontological Relativity and Other Essays*, Columbia University Press, New York 1969.

Quine, Willard Van Orman: *Ontological Relativity and Other Essays*, Columbia University Press, New York 1969.

Quine, Willard Van Orman: "Has Philosophy Lost Contact with People?" pp 190-193 in *Theories and Things*, Harvard University Press, Cambridge, Mass. 1981.

Quine, Willard Van Orman: "Reply to Roger F. Gibson, JR." in Hahn, Lewis Edwin (ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, rpr. 1988.

Quine, Willard Van Orman: *Pursuit of Truth*, Harvard University Press, Cambridge, Mass. 1990.

Quine, Willard Van Orman: "Response to George" in Orenstein, Alex and Petr Kotatko (Eds.): *Knowledge, Language and Logic*, Kluwer Academic Publishers, Dordrecht / Boston / London 2000,

Quine, Willard Van Orman: "Three Indeterminacies" in R. Barrett and R. Gibson (eds.): *Perspectives on Quine*, Oxford University Press, Oxford 1990.

Quine, Willard Van Orman: *From Stimulus to Science*, Harvard University Press, Cambridge, Mass. 1995

Quine, Willard Van Orman: "Reply to John Woods", in Lewis Edwin Hahn (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998.

Quine, Willard Van Orman: "Autobiography (continued)", in Lewis Edwin Hahn (Ed.): *The Philosophy of W. V. Quine*, in the series The Library of Living Philosophers, Open Court, La Salle 1986, reprinted 1988, extended version of 1998.

Quine, Willard Van Orman: "I, you, and it: an epistemological triangle" in Orenstein, Alex and Petr Kotatko (Eds.): *Knowledge, Language and Logic*, Kluwer Academic Publishers, Dordrecht / Boston / London 2000.

Quine, Willard Van Orman: "Response to Gibson" in Alex Orenstein and Petr Kotatko (Eds.) *Knowledge, Language, and Logic: Questions for Quine*, pp 414-415, Kluwer Academic Publishers, Dordrecht 2000.

Ramberg, Bjørn: *Donald Davidson's Philosophy of Language*, Blackwell, Oxford 1989.

Ramberg, Bjørn: "Naturalizing Idealizations: Pragmatism and the Interpretivist Strategy" at http://folk.uio.no/bjoerntr/Natideal.html

Rorty, Richard: Review of Davidson's *Problems of Rationality*, *Notre Dame Philosophical Reviews*, 2005-02-01.

Sinclair, Robert: "What is Radical Interpretation?" in *Inquiry*, 45, 2002, pp 161-84.

Solberg, Mariann: "Velvillighetsprinsippets epistemiske status", pp 61-76 in Norsk Filosofisk tidsskrift, nr 1-2, 2004.

Solberg, Mariann: "Davidson and the Role of Error" in Hallvard Fossheim, John Richard Sageng and Tarjei Mandt Larsen (eds.): *Non-Conceptual Aspects of Experience*, Unipub, Oslo 2003, pp 119-134.

Smart, J.J.C.: "Correspondence, Coherence, and Realism" in Lewis Edwin Hahn (Ed.): *The Philosophy of Donald Davidson*, in the series The Library of Living Philosophers, Open Court, Chicago and La Salle 1999, pp 109-122.

Wilson, N. L: "Substances without Substrata" in *The Review of Metaphysics*, Vol. 12, nr 4, 1959, pp 521-539.