

APPROPRIATENESS OF A COGNITIVE APPROACH TO DONALD
DAVIDSON'S MEANING THEORY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ESER AĞOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN
THE DEPARTMENT OF PHILOSOPHY

DECEMBER 2008

Approval of the Graduate School of Social Sciences

Prof. Dr. Sencer AYATA
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Arts.

Prof. Dr. Ahmet İNAM
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Arts.

Assoc. Prof. Dr. Erdinç SAYAN
Supervisor

Examining Committee Members

Assoc. Prof. Dr. Erdinç Sayan	(METU,PHIL)	_____
Assoc. Prof. Dr. Ayhan Sol	(METU,PHIL)	_____
Dr. Ceyhan Temürçü	(METU, COGS)	_____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name :

Signature :

ABSTRACT

APPROPRIATENESS OF A COGNITIVE APPROACH TO DONALD DAVIDSON'S MEANING THEORY

Ağoğlu, Eser

M.A., Department of Philosophy

Supervisor : Assoc. Prof. Dr. Erdinç Sayan

December 2008, 62 pages

The purpose of this study is to discuss the appropriateness of a cognitive approach to Donald Davidson's meaning theory. Davidson makes the bold proposal that a truth theory, modified for a natural language, may be treated as a meaning theory for that language. According to Davidson, a meaning theory is an empirical theory. Radical Interpretation is at the center of such an empirical inquiry which places restrictions on the truth theory to make it suitable as a meaning theory without appeal to semantic notions. Davidson's aim in presenting this bold proposal and radical interpretation is to shed light on the concept of meaning, not to define the actual semantic competence of language users. But what Davidson's project does not aim to define is the main thing that a cognitive approach must account for. Whether a truth theory can represent the semantic competence of language users is discussed in this work. It is concluded that, although there is no a priori reason for such a representation claim, the cognitive approach—with the right assumptions to make the claim testable—can lead to an empirical research programme.

Keywords: Meaning Theory, Theory of Truth, Radical Interpretation, Universal Grammar

ÖZ

DONALD DAVIDSON'IN ANLAM TEORİSİNE BİLİŞSEL BİR YAKLAŞIMIN UYGUNLUĞU

Ağođlu, Eser

Yüksek Lisans, Felsefe Bölümü

Tez Yöneticisi : Doç. Dr. Erdinç Sayan

Aralık 2008, 62 sayfa

Bu çalışmanın amacı Donald Davidson'ın anlam teorisine bilişsel bir yaklaşımın uygunluđunu tartışmaktır. Davidson, doğal bir dil için uyarlanmış doğruluk teorisine, o dil için anlam teorisi olarak yaklaşılabileceđine dair bir öneride bulunmuştur. Davidson'a göre bir anlam teorisi ampirik bir teoridir. Doğruluk teorisini sınırlayıp semantik kavramlara başvurmadan bir anlam teorisi olarak kullanılmasına uygun hale getiren radikal yorumlama kavramı bu ampirik araştırmanın merkezindedir. Davidson'ın önerisini ve radikal yorumlama kavramını sunuşundaki amacı gerçek dil kullanıcılarının semantik yeterliliđini tanımlamak deđil, anlam kavramına ışık tutmaktır. Fakat Davidson'ın projesinin tanımlamayı amaçlamadıđı bu yeterlilik, bilişsel yaklaşımın açıklaması gereken ana konudur. Bu çalışmada doğruluk teorisinin dil kullanıcılarının semantik yeterliliđini temsil edip edemeyeceđi tartışılmıştır. Böyle bir temsil iddiası için a priori neden olmamasına karşın, bilişsel yaklaşımın—iddiayı test edilebilir hale getiren doğru varsayımlarla— bizi ampirik araştırma programına yönlendirebileceđi sonucuna ulaşılmıştır.

Anahtar Kelimeler: Anlam Teorisi, Doğruluk Teorisi, Radikal Yorumlama, Evrensel Dilbilgisi

To İsmail Hakkı Özsulu

ACKNOWLEDGMENTS

I express my deepest gratitude to Assist. Prof. Dr. John H. Bolender for his guidance, advice, criticism, encouragements and valuable insight throughout the research. I am also thankful for Assoc. Prof. Dr. Erdinç Sayan for all his support and guidance. I would also like to thank Assoc. Prof. Dr. Ayhan Sol and Dr. Ceyhan Temürçü for their suggestions and comments.

TABLE OF CONTENTS

PLAGIARISM.....	iii
ABSTRACT.....	iv
ÖZ.....	v
DEDICATION.....	vi
ACKNOWLEDGMENTS.....	vii
TABLE OF CONTENTS.....	viii
CHAPTER	
1. INTRODUCTION.....	1
1.1) Theories of Meaning and Meaning Theories.....	1
1.2) A Brief Historical Introduction.....	2
1.3) Methodology and Organization.....	3
2. DAVIDSON’S MEANING THEORY.....	6
2.1 Compositionality as a Source of Davidson’s Meaning Theory.....	6
2.2) Tarski and Truth Theory.....	7
2.3) The Form of Theorems for a Meaning Theory.....	16
3. RADICAL INTERPRETATION.....	27
3.1) Behaviorism, Quine and Radical Translation.....	27
3.2) Davidson’s Revision of Radical Translation.....	30
3.3) Radical Interpretation and Truth Theory.....	31
4. THE COGNITIVE APPROACH.....	36
4.1) The Cognitive Revolution and Chomsky.....	36

4.2) Competence, the Cognitive Approach and Radical Interpretation	40
5. A COGNITIVE APPROACH TO DAVIDSONIAN SEMANTIC THEORY .	44
5.1) Objections against Davidsonian Semantics	44
5.2) A Cognitive Approach to Davidsonian Semantics	49
6. CONCLUSION	55
BIBLIOGRAPHY	56

CHAPTER 1

INTRODUCTION

1.1) Theories of Meaning and Meaning Theories

Semantics is traditionally defined as the study of meaning—or more specifically literal meaning. Although this definition seems to be vague as a result of containing an ambiguous concept, namely *meaning*, it serves as a categorical framework which draws borders for our study distinguishing it from syntax and pragmatics. They each take a different aspect of linguistic expressions as a subject matter. Linguistic expressions are formed by concatenation of morphemes and phrases. Syntax focuses on the study of the relation of these elements with each other in linguistic expressions independent from their meaning relevant properties. Linguistic expressions are also used to communicate messages different from their literal meanings whose study is in the domain of pragmatics. The border between the domains of semantics and pragmatics is not clear, since the term *literal* loses its sense if it is said that all there is to meaning is what is communicated.

Meaning is central to our linguistic practice; every language user has an everyday grasp of the meanings of expressions and uses such expressions for the sake of the fact that they mean something. So what is ambiguous about meaning? Much of the confusion is the result of the methodology of the study. Consider the following meta-level question: What is meaning? A study which takes such a metaphysical question to be central will try to analyze the concept of meaning or reduce it to more primitive concepts.¹ The term *theory of meaning* may be used to for

¹ For example, Frege (1948, 1956) gives an account of meaning in terms of the notions of *reference*, *sense*, *tone* and *force*, and Grice (1957) takes task of reducing sentence meanings to underlying psychological states.

methodologies whose aim is to give an account of the nature of meaning by this way. Theories of meaning do not exhaust the ways of studying semantics: “Another approach is to begin, pretheoretically, with phenomena that we think are connected to meaning, to construct a theory, to explain the phenomena, and to take as meanings what the theory attributes to expressions” (Davis & Gillon, 2004, p. 4). If we designate such theories as *meaning theories*² to emphasize the difference in methodology, then we may state that Davidson’s semantic theory is a kind of meaning theory. Davidson (1966/2001a) observes that natural languages consist of infinitely many meaningful expressions and also they are learnable, and assumes the task of identifying a theory, the truth theory restricted in a special way, to explain such phenomena. He defines truth conditions of sentences which are entailed by the theory as giving the meanings of such sentences.³ Meaning theories are also concerned with questions about the nature of meaning, but they are meant to shed light on them in an indirect way.

1.2) A Brief Historical Introduction

Davidson’s “bold proposal” (1999a) was that a truth theory may serve as a meaning theory. Such a proposal consists of a formal part and an empirical part. The formal part was a truth definition as developed by Tarski (1936/1956). Davidson offered to take such a formal structure as the core of meaning theory. Tarski developed his definition for formal languages and expressed pessimism about the applicability of such a definition to natural languages. But Davidson’s main interest was natural languages, so he outlined a programme which aims to reveal the logical form of natural language sentences and thus make them compatible with the formal

² The terms *theory of meaning* and *meaning theory* are adapted from Lepore and Ludwig (2005, pp. 19-20) as they suggested for clarifying the terminology. Note that Davidson uses *theory of meaning* freely to refer to them both. An alternative would be *analytic theory of meaning* for the former and *constructive theory of meaning* for the latter as indicated by Glock (2004, p. 142).

³ Note that Davidson does not identify the meanings of sentence with their truth conditions. The relationship between meaning and truth is a loose one. See section 2.3.

structure of truth theory. While Davidson's programme was adopted and developed by several philosophers, another discipline which also addressed the structure of linguistic expressions was flourishing. Chomsky (1957, 1965) was revolutionizing linguistics with his *generative grammar*. Early generative grammar assumes two level of architecture of *surface structure* (SS) and *deep structure* (DS). The DSs were said to be formal representations which are formed by the syntactic rules and supply input to the semantic component which forms the semantic interpretation of sentences. The grammaticality of a sentence was defined as being autonomous from the semantic properties of its components. The SSs were said to be formed from the deep structures by meaning preserving *transformation rules* and as supplying input to the phonetic component which forms the phonetic interpretation of sentences. The power and fruitfulness of the formal structure of generative grammar together with its relative weakness in semantics motivated linguists and philosophers to unite two projects. Their main proposal was that the deep structures and the logical forms of sentences can be identified (Davidson & Harman, 1972). Although the influence of this endeavor slowly faded away with the success of the extended versions of generative grammar within the Chomskian framework, together with Davidson's presentation of radical interpretation, it did not completely disappear. In the 90s, another attempt for unification of truth conditional semantics with of linguistic theory started to emerge (Higginbotham, 1997; Larson & Segal, 1995).

1.3) Methodology and Organization

The purpose of this thesis is to discuss and evaluate the appropriateness of a cognitive approach to Davidson's meaning theory. To accomplish this, two tasks must be completed; that is, (a) the clarification of Davidson's project on semantics and (b) the exposition of what a cognitive approach is. Davidson's work on semantics is spread through a collection of papers which were published over an extended period of years. Davidson also admitted that his views underwent several modifications during the defense of his bold proposal against criticisms. So the task of clarification will include a presentation of Davidson's semantic theory and a discussion of the nature of such modifications together with the evaluation of

generally accepted claims about them. The central position and importance of radical interpretation in Davidson's theory will be clarified and it will be argued that many objections against the bold proposal were raised, only due to misunderstandings or underestimations of its role in the overall project. Davidson's position as here clarified will be labeled the *philosophical approach* in contrast to the cognitive approach. The task of expounding the cognitive approach will be completed through the presentation of its origin and main principles in a separate chapter. It will turn out that any cognitive approach to Davidson's theory must abandon the philosophical project together with radical interpretation. But, without it, a truth theory cannot be said to serve as a meaning theory. So it will be concluded that any cognitive approach has to (a) find a way to cope with criticisms that are raised against the bold proposal and (b) legitimize the use of truth theory as a meaning theory. Larson and Segal's (1995) attempt which aims at a unification of truth conditional semantics with generative linguistics will be evaluated as an example of the cognitive approach.

Chapter 2 will focus on Davidson's early works on meaning theory—which were meant to be published before “Radical Interpretation” (1973/2001e). The compositionality of language and Tarski's truth definition, which are the main sources of Davidson's meaning theory, will be presented in separate sections. The rest of the chapter will consist of Davidson's discussion of the form of the meaning theory. This discussion will reveal that the empirical application of truth theory is central to Davidson's semantic theory.

Chapter 3 will introduce Davidson's procedure of radical interpretation. Davidson considers radical interpretation to be a modification of Quine's similar thought experiment of radical translation. Quine's radical translation, together with its relation to behaviorism and the content of Davidson's modification of radical translation, will be presented in the first two sections. The Last section will focus on the relation between radical interpretation and formal truth theory in light of morals drawn from chapter 2.

Chapter 4 will begin with an introduction to the cognitive revolution, which is a paradigm change in psychology, against the dominant view of behaviorism. The cognitive revolution influenced different disciplines, and its underlying motive became an interdisciplinary theme including psychology, linguistics, neuroscience,

and even philosophy. The first section will focus on the cognitivist approach of Chomsky in linguistics. The next section discusses the compatibility of such a cognitive approach with radical interpretation.

Chapter 5 will discuss the appropriateness of a cognitive approach to Davidson's meaning theory. Several objections are raised against Davidson's bold proposal that a truth theory may serve as a meaning theory. The first section will present such objections and responses to them in light of radical interpretation. Some objections will be directed against the cognitive approach and the test of relevancy of truth will be presented to evaluate the appropriateness of a cognitive approach to Davidsonian semantics. The test will be applied to Larson and Segal's case and it will turn out that their approach fails to pass the test. A revision in their main assumptions will be proposed to avoid the failure.

CHAPTER 2

DAVIDSON'S MEANING THEORY

2.1 Compositionality as a Source of Davidson's Meaning Theory

“Theories of meaning and Learnable Languages” (Davidson, 1966/2001a) is a presentation of the compositionality requirement which is at the center of Davidson's overall project on meaning. The *learnability argument* reveals the need for a compositional meaning theory: Natural languages are learnable, and any meaning theory which fails to give an account of this fact will not be adequate. Davidson invokes a contrast between the infinitude of language and the finiteness of its users⁴ through the learnability argument.

Any natural language consists of an infinitely many sentences; for example, it is always possible to form a new sentence from a present one by adding the phrase *it is not the case that* in front of it:

1. The weather is cold.
2. It is not the case that (the weather is cold).
3. It is not the case that (it is not the case that (the weather is cold)).

and so on. Human beings are finite in nature; they may learn only a limited number of semantic elements through their lifespan. But humans learn to entertain a potentially infinite number of sentences in their early childhood. So, the limited

⁴ This observation is not new. Humboldt claims that language makes “infinite use of finite media” (as cited in Abler 1989, p. 1). But Davidson uses such observation as a constraint on semantic theories. Also see Postal (2004, chap. 6) for the critique of the view that language consists of finite number of semantic primitives.

number of elements that are learned is sufficient for using and understanding the infinite number of sentences. It is only possible if the meanings of complex expressions are determined by the features of their elements and their mode of composition.

Davidson does not just adopt Frege's so-called *compositionality principle* as a dictum.⁵ Compositionality is the conclusion of Davidson's learnability argument, which is an empirical argument due to its assumptions (Davidson, 1966/2001a, pp. 8-9). This presentation is important, because it exhibits the nature of Davidson's approach to the problem of meaning. He starts with a pre-theoretical claim which is connected with meaning—learnability—and asks for a theory which gives an account of this.

If language is compositional, then a meaning theory must be compositional as well. If a meaning theory for a language which gives the meanings of its expressions is not compositional, then it will include an infinite number of semantic primitives, and the language which it aims to account for will not be compositional. Therefore, it would not be a learnable natural language which we want to account for. In the light of compositionality, Davidson (1966/2001a) states that a meaning theory is "a constructive account" (p. 3) which (a) gives the meaning of sentences of a language by appeal to sentence parts and (b) generates the meaning of any given arbitrary sentence. Davidson adds that the structure of a Tarski style truth theory is "identical with or closely related to" (p. 8) the structure just outlined and can serve as a meaning theory.

2.2) Tarski and Truth Theory

Truth is a primitive concept about which people have pre-theoretical intuition. In this respect, the concept of truth is similar to the concept of meaning: Despite the fact that both of them seem to be unproblematic and clear, varied and conflicting answers given to the questions about them form one of the most controversial areas

⁵ Frege's principle states that the meaning of a sentence depends on the meanings of its parts and their mode of combination (Pelletier, 2001).

of philosophical investigation. The scene was already populated by several forms of correspondence, coherence and pragmatic theories of truth, which deal with such problems, when Tarski set out to give a *definition of truth*. As a mathematician, his aim is to give a definition of truth which may be employed in “deductive sciences” (Künne, 2005, p. 176)—namely, in the formal languages used to state them (Tarski, 1936/1956, p. 166). According to Tarski, the available theories were neither scientifically respectable nor free of paradoxes. So, he identified sentences as truth bearers and presented two conditions that must be satisfied by a definition of *truth predicate*—namely, *is true*—for a language: “[The] task is to construct—with reference to a given language—a *materially adequate and formally correct definition of the term ‘true sentence’*” (Tarski, 1936/1956, p. 152).

Material adequacy is the condition which will ensure that the predicate which is defined amounts to our *intuitive notion of truth*. Although Tarski roughly appeals to the “classical conception of truth”⁶ (Tarski, 1936/1956, p. 153; 1944, pp. 342-343) to pre-theoretically identify the intuitive notion, he also claims that such formulations are not “sufficiently precise and clear” (Tarski, 1944, p. 343). Let *x* be a sentence and consider the formulation “*x* is true if and only if *x* corresponds to reality.” This invokes the ambiguous notions of “correspondence” and “reality” (Gupta, 1998, p. 265) which will confuse one as to whether the definition “actually fulfills its task” (Tarski, 1944, p. 341). Tarski (1936/1956) proposes another form which is capable of capturing our intuitions about truth: “‘it is snowing’ is a true sentence if and only if it is snowing” (p. 156).

It is a truism to claim that the sentence *it is snowing* turns out to be true when it is snowing. Tarski considers the conditions when a sentence is true or false: If what a sentence means is the case, then it is true and if what the sentence means is not the case, then it is false. The sentences of a language have truth values because they *mean* something. Consider the following schema which contains the truth and meaning predicates:

⁶ Tarski uses the terms freely to refer to correspondence theories of truth as well as Aristotle’s original formulation: “To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true...” (as cited in Tarski, 1944, p. 343)

Schema TM. If X means in L that P , then X is true in L iff P . (Soames, 1999, p. 68)

where X is the (structural descriptive) name of the sentence and P is that sentence. For the sentence *it is snowing*, the schema will entail the below:

If “it is snowing” means in English that it is snowing, then “it is snowing” is true in English iff it is snowing.

By assuming a prior grasp of meanings of English sentences, one can derive:

“It is snowing” is true in English iff it is snowing.

Convention-T⁷ is proposed as a criterion for material adequacy. Satisfaction of it by any truth definition will ensure one that the defined predicate amounts to our intuitive concept of truth. A truth definition for a language must entail all and only the instances of Schema-T which satisfy Convention-T:

Schema T. X is T in L iff P . (Soames, 1999, p. 68)

where X is the structural description⁸ of a sentence, and P is that sentence or the translation of it in the meta-language (Tarski, 1936/1956, pp. 187-188). For every instance of Schema-T, there will be a corresponding instance of Schema-TM. So the predicate which is defined for that language will be “extensionally equivalent” with our intuitive truth predicate: “ X is T in L iff X is true in L ” (Soames, 1999, p. 68).

Formal correctness specifies the characteristics of the languages through which the definition is given. Natural languages are “semantically closed” (Tarski,

⁷ When compared with Tarski’s original presentation (Tarski, 1956, pp. 187-188), Convention-T is revised for the sake of simplicity and conformity with the philosophical literature.

⁸ The structural description of an expression “describes the expression as a concentration of elements from a finite list (for example of words or letters)” (Davidson, 1967/2001b, p. 18). Through this thesis, double quotation marks are used to indicate that the quoted expression is a structural description.

1944, p. 348) or said to have the characteristic of “universality” (Tarski, 1936/1956, p. 165) which means that the semantic terms of the language can be used to refer to sentences of the language itself. This characteristic allows for semantic paradoxes such as the *liar’s paradox*. Consider the following claim:

This sentence is false.

If the sentence is false, then the falsity claim is false too and the sentence is true, but the sentence explicitly states its falsity. Tarski (1944) distinguishes the *object language* and the *meta-language* to avoid such paradoxes. The object language is the language *for which* the truth predicate is defined and the meta-language is the language *through which* the truth predicate is defined. The object language is forbidden to have names for its sentences or any semantic terms used to refer them, in contrast to the meta-language. The structure of the object language must be previously specified which means that the sense of every meaningful expression must be determined by its form so as to avoid any ambiguity. The object language for which truth is defined cannot be a natural language, hence it does not have the characteristics specified by criteria of formal correctness. As a result, Tarski claims that the truth predicate can be defined for formal languages, or strictly restricted parts of natural languages. The same specification also applies to semantic terms—such as *satisfaction* or *denotation*—which appear in the meta-language. Any undefined term is said to violate formal correctness (pp. 346-351).

Suppose that a language L_1 contains only a single sentence *it is snowing*. So the truth predicate may be defined by D_1 :

(D_1) “it is snowing” is true in English iff it is snowing.

Let L_2 contain one more sentence in addition to L_1 :

(D_1) “Snow is white” is true if and only if snow is white.

(D_2) “Snow is cold” is true if and only if snow is cold.

For L_1 , D_1 is the definition of truth for that language but for L_2 , D_1 and D_2 are both *partial definitions*. The General definition⁹ for them will have a universally quantified form:

(GD) (s)(s is a true sentence in L_2 iff ((s = “Snow is white” and snow is white) or (s = “Snow is cold” and snow is cold)))

It is obvious that Convention-T is satisfied and the definitions are materially adequate in the above cases. But such definitions are only applicable to languages which have a finite number of sentences. The main problem is to find a way to generalize the definition which is given in D_1 for a language with infinitely many sentences: GD will not work for such cases. Tarski claims that such a generalization is not possible, and proposes constructing a *recursive definition*: “A true sentence is a sentence which possesses such and such structural properties (...) or which can be obtained from such and such structurally described expressions by means of such and such structural transformations” (Tarski, 1936/1956, p. 163).

Now, suppose that a language L_3 contains two sentences of L_2 and a logical constant. Contrary to the above definitions, the language will contain infinitely many sentences due to the concatenation of an atomic or compound sentence with another by the logical constant *and* recursively. The recursive definition of *true in L_3* may be given after specifying the syntax of the object language as:

(D₁) “Snow is white” is true in L_3 if and only if snow is white.

(D₂) “Snow is cold” is true in L_3 if and only if snow is cold.

(R) For any two sentences of S_1 and S_2 , “ S_1 and S_2 ” is true if and only if S_1 is true and S_2 is true.

⁹ The definition of truth for languages with finite sentences presented here is derived from Künne (2005, p. 192). Similar generalizations may be found in Soames (1999, pp. 69-70) and Platts (1997, pp. 16-17) which are variants of Tarski’s original formulation (Tarski, 1936/1956, p. 188). To avoid any confusion which may arise from the term *general*, it should be noted that it is a general definition of *truth for sentences of finite number*, not a general definition of truth.

Although the adequacy of the definition may be tested for any given arbitrary sentence of L_3 , it is evident that it will logically entail sentences which are the required kind of sentences by Convention-T. But D_1 , D_2 and R form a set of axioms, not an explicit definition. The axiomatic theory may be transformed into an explicit definition¹⁰ as:

(s)(s is true in L_3 iff there is a set T_3 such that s is a member of T_3 and for all sentences z of L_3 , z is a member of T_3 iff (i) z = “Snow is white” and snow is white; (ii) z = “Snow is cold” and snow is cold; (iii) z = “ S_1 and S_2 ” for some sentences of L_3 , both of which are members of T_3 .)

Such a definition will be adequate as long as the language contains infinitely many sentences formed by unrestricted use of logical constants over its atomic and compound sentences. But a language may contain infinitely many sentences which are formed by quantification over “open sentences”¹¹ (Kirkham, 1992, p. 152) such as the sentence “Someone cuts my lawn and prunes my roses” (Platts, 1997, p. 19). The quantified form of the sentence is *for some x(x cuts my lawn and x prunes my roses)* and *x cuts my lawn and x prunes my roses* are open sentences. It is obvious that open sentences cannot take truth values, and this blocks the way for giving the definition of truth for a sentence by considering their truth values. Tarski (1936/1956) proposes to define a more general concept called *satisfaction* which “is applicable to sentential functions” (p. 189) and gives a definition of truth recursively by means of such a predefined concept. Given that a quantified sentence may contain infinitely many variables, they may be satisfied by infinite sequences of objects. Suppose that L_4 contains two open sentences (*n is black* and *n is white*), two names¹² (*snow* and *coal*)—defined as terms—the existential quantification symbol (EQ), infinitely many variables ($x_1, x_2, x_3\dots$)—also defined as terms—and the logical

¹⁰ See Soames (1999, p. 74).

¹¹ Tarski uses the term “sentential function” (1936/1956, p. 189) instead of the term *open sentence*.

¹² Tarski’s original formulation does not contain names as terms but only variables.

constant of conjunction. So the truth definition¹³ for L_4 will be composed of the following axioms:

The Reference Axioms:

(RA.1) $(s)(\text{ref}(\text{"snow"}, s) = \text{snow})$

(RA.2) $(s)(\text{ref}(\text{"coal"}, s) = \text{coal})$

(RA.3) $(s)(k)(\text{ref}(\text{"x}_k", s) = s_k)$

where *ref* amounts to “*reference-relative-to-a-sequence*,” “‘s’ ranges over sequences,” “‘k’ ranges over natural numbers and s_k is the k^{th} element of s ” (Neale , 2001, pp. 34-36)

Satisfaction Axioms:

(SA.1) $(s)(n)(s \text{ satisfies “} n \text{ is white” iff } \text{ref}(n, s) \text{ is white.})$

(SA.2) $(s)(n)(s \text{ satisfies “} n \text{ is black” iff } \text{ref}(n, s) \text{ is black.})$

(SA.3) $(s)(\alpha)(\beta)(s \text{ satisfies “} \alpha \text{ and } \beta” \text{ iff } s \text{ satisfies } \alpha \text{ and } s \text{ satisfies } \beta)$

(SA.4) $(s)(k)(\alpha)(s \text{ satisfies “} (\text{EQ}_{x_k})\alpha” \text{ iff there is at least one sequence differing from } s \text{ at most in the } k\text{th place that satisfies } \alpha)$

where n ranges over terms—names or variables—and α and β over open or closed sentences.

Definition of Truth:

(DT) $(S)(S \text{ is true iff it is satisfied by every (some) sequence.})$

where S is a closed sentence.

What it is for a sequence to satisfy a closed sentence such as $(x_1) (x_1 \text{ is white})$. A sequence S_1 satisfies a universally quantified sentence if and only if S_1 satisfies the

¹³ The Truth definition for L_4 is originally adapted from Neale (2001, pp. 34-36). Although the definition is presented as an axiomatic theory of truth, it can be transformed into an explicit definition by appealing to set theory—just like the transformation presented in L_3 .

open sentence which is formed by the removal of the quantification, and S_1^* satisfies the same open sentence where S_1^* is a set of all sequences identical with S except for the first member. So, the set of the first members of the sequences of S_1^* will include all objects which can be members of a sequence. So S_1 satisfies $(x_1) (x_1 \text{ is white})$ if and only if everything (every object) is white. The same condition applies to some other sequences S_2, S_3, S_4, \dots . Suppose that S_5 fails to satisfy the open sentence, this means there is an object that is not white so it is not the case that everything is white. If one sequence satisfies $(x_1) (x_1 \text{ is white})$, it is clear that all other sequences satisfy that quantified open sentence. The same condition is applied to existentially quantified and merely closed sentences which contain no variables (Kirkham, 1992, p. 156-157). A sentence is true if it is satisfied by all sequences or false otherwise (Tarski, 1956, pp. 193-195). Informally, *snow* satisfies $x \text{ is white}$ if *snow is white* is true, but such a formulation is not open to Tarski because he aims to define truth by using the notion of satisfaction.

The truth conditions of any object language sentence of L_4 can be stated from the axioms of the above presented axiomatic truth theory for L_4 . Consider the sentence *snow is white and coal is black* and by the use of standard first order logic (UQI: Universal Quantifier Instantiation, REP: Replacement Scheme, SUB: Substitution of Identicals)

- (1) “snow is white and coal is black” is true iff it is satisfied by every (some) sequence. (DT)
- (2) “snow is white and coal is black” is true iff “snow is white and coal is black” is satisfied by at least one sequence, s . (DT)
- (3) “snow is white and coal is black” is true iff s satisfies “snow is white” and s satisfies “coal is black” (SA.3, UQI)
- (4) s satisfies “snow is white” iff $\text{ref}(\text{“snow”}, s)$ is white. (SA.1, UQI)
- (5) s satisfies “coal is black” iff $\text{ref}(\text{“coal”}, s)$ is black. (SA.2, UQI)
- (6) “snow is white and coal is black” is true iff $\text{ref}(\text{“snow”}, s)$ is white and $\text{ref}(\text{“coal”}, s)$ is black. (3, 4, 5, REP)
- (7) “snow is white and coal is black” is true iff snow is white and coal is black. (6, RA.1, RA.2, SUB)

The Material adequacy of the theory may be tested by checking whether or not T-theorems (theorems which are entailed by the theory) satisfy Convention-T. But no matter how many arbitrary theorems are checked, the evaluation would not be conclusive. But in the formulation of reference and satisfaction axioms, metalanguage translations of object language sentences are used systematically. Tarski clearly states this fact as follows: “*For all a, a satisfies the sentential function x if and only if p....we insert for ‘p’ not the sentential function itself, but the expression of the metalanguage having the same meaning...*” (Tarski, 1936/1956, p. 190). If sameness of meaning is presupposed, the truth theory will logically entail T-Theorems which are instances of Schema-T and the totality of them will ipso facto satisfy Convention-T. This property of the axioms of the truth theory exhibits a very important aspect of Tarski’s truth definition: It is a truth definition for a specific language, not a general definition of truth. There will be a finite list of axioms and such axioms will be stated exclusively for each predicate and name of the object language. So the resulting T-theorems will have the form:

“snow is white” is true in English iff snow is white.

“Schnee ist weiss” is true in German iff snow is white.

For a sequence of languages <German, Chinese, Turkish...> there will be a corresponding sequence of truth definitions <Truth Definition for German, Truth Definition for Chinese, Truth Definition for Turkish...> where the meta-language is supposed to be English. They may all satisfy Convention-T and turn out to be extensionally equivalent to one’s intuitive notion of truth, but Convention-T is a criterion which is stated in the meta-theory to check the adequacy of the stated theories. It cannot unify different theories. So one’s intuitive notion of truth is used pre-theoretically in a loose sense. Otherwise Tarski’s definition may be accused of presupposing the notion which it aims to define.¹⁴

¹⁴ Although it is controversial whether Tarski’s theory is a form of the correspondence theory, such discussion is not presented here (Field, 1972; Soames, 1984; Etchemendy, 1988; Davidson, 1990).

2.3) The Form of Theorems for a Meaning Theory

Davidson (1966/2001a) claims that the structure of a Tarskian truth theory is “identical with or closely related to” (p. 8) the structure of a meaning theory. But Davidson does not assert directly that a truth theory may serve as a meaning theory prior to discussion about what requirements a predicate for meaning theory must satisfy to count as adequate. Davidson (1967/2001b) considers two alternatives (indicated in M_1 and M_2 below) which are proposed to serve as the predicate (indicated in P below) and decide that they cannot serve as a form of a meaning theory.

(P) s (predicate) p

(M_1) s means p

(M_2) s means that p

M_1 may be defined as the “referential account of meaning” (Glock, 2004, p. 147) such that p is an entity which serves as meaning of s . Davidson (1967/2001b) rejects this account, because meanings as entities have “no demonstrated use” (p. 21). Davidson’s argument against M_1 proceeds by showing that appealing to meanings as entities is “neither sufficient nor necessary” (Ludwig & Lepore, 2003, p. 40) for a compositional meaning theory. According to the compositionality requirement, the meaning of a sentence must be given by appeal to its constituents. Consider the following sentence:

“Brutus killed Caesar” means p_1

It is required that each word in the sentence has an entity as its meaning, and the meaning of the sentence p_1 is determined by them. Let the entities Brutus and Caesar be assigned as meanings for the words *Brutus* and *Caesar* respectively, and let the relation (property) of killing be assigned as the meaning of the word *killed*. The same words can be rearranged to form a new sentence, namely:

“Caesar killed Brutus” means p_2

Despite the fact that each sentence contains all and only the same words (same set of meanings as entities), the meanings of p_1 and p_2 are intuitively different. To overcome this difficulty, a new relation of *concatenation* of words may be included in the theory. But one must also treat such a relation as a meaning, which leads to an “infinite regress” (Davidson, 1967/2001b, p. 17).¹⁵ So appealing to meanings as entities is not sufficient to determine the meaning of a complex expression, since it only leads us to a mere list of constituent elements.

Davidson (1967/2001b) next considers the Fregeian treatment of predicates as “unsaturated entities” which also applies to “complex singular terms” (p. 17). Suppose that the singular term *the teacher of Aristotle* stands for a meaning, and suppose further that the entity of Aristotle stands for the meaning of the word *Aristotle* and an unsaturated entity stands for *teacher of*:

“Aristotle” refers to Aristotle.

“The Teacher of Aristotle” refers to the teacher of Aristotle.

The question is how the meaning of the *teacher of Aristotle* depends upon the meanings of its parts. In this case, it is evident that such an unsaturated entity adds nothing to the determination of the reference of the expression except stating that the referent must be the teacher of the entity referred to by the singular term. Assigning it an entity adds nothing to its “explanatory function” (Davidson, 1967/2001b, p. 17). Davidson offers a *recursive rule* which functions as such entity:

[R] For all referring terms α , the concatenation of “the [teacher] of” with α refers to the [teacher] of what α refers to. (Lepore & Ludwig, 2005, p. 46)

If one knows the rule then (by R and universal quantifier instantiation):

¹⁵ This regress is known as the Bradley’s regress (Bradley, 1916, pp. 20-22).

(R1) The concatenation of “the teacher of” with “Aristotle” refers to the teacher of what Aristotle refers to.

In addition if one knows that *what Aristotle refers to* is Aristotle, or *Aristotle* refers to Aristotle, then,

“The teacher of Aristotle” refers to the teacher of Aristotle.

Still *the teacher of Aristotle* is a referring term and the rule can be reapplied to it:

“The teacher of the teacher of Aristotle” refers to the teacher of the teacher of Aristotle.

Although the formulation uses part of the sentence for which reference is given, it is compatible with the compositionality requirement anyway. Compositionality does not require giving the meanings of all elements, but requires giving the meaning of the expression by giving the meanings of its parts (Davidson, 1967/2001b, p. 18). The recursive character of such a sample theory enables it to give the references for an infinite number of expressions from a finite base. Davidson concludes that meanings as entities are not necessary for a meaning theory.

Davidson assumes that the meaning of an expression is its reference to reach that conclusion. But is such a treatment right in the first place? Davidson invokes the *slingshot argument*¹⁶ against M₁:

The slingshot argument aims to show that any compositional meaning theory in which sentences refer to their meanings is unworkable, because once we assume sentences are singular referring terms, we are forced to the patently false conclusion that all sentences alike in truth value are synonymous. (Lepore & Ludwig, 2005, p.50)

¹⁶ The slingshot argument is also known as the “Frege argument” (Evnine, 1991, p.180; Neale, 2001, p. 8) in the literature. In Fact, Neale’s book *Facing Facts* (2001) is entirely devoted to evaluating the slingshot argument. Only a primitive version of it will be presented uncritically in this work to show how Davidson uses it against meanings as entities.

The success of the argument depends on main assumptions which Davidson (1967/2001b) presupposes: (A1) “Logically equivalent singular terms have the same reference” and (A2) “a singular term does not change its reference if a contained singular term is replaced by another with the same reference” (p. 19). Suppose that p and q are *Ankara is more populated than İzmir* and *İzmir is less populated than Ankara* respectively. They are singular terms and *logically equivalent*, so according to A1, they have the same reference. Both of them consist of atomic singular terms which have their own references. Now, *Ankara* and *The capital city of Turkey* are co-referential. According to A2, the singular terms of *Ankara is more populated than İzmir* and *the capital city of Turkey is more populated than İzmir* have the same referents and so they have same meaning. In the same line, Davidson further claims that *Ankara is more populated than İzmir* and the sentence of the form *(the x such that x = x, and Ankara is more populated than İzmir) = (the x such that x = x)* are logical equivalences because “it is impossible for them to differ in truth value” (Evnine, 1991, p. 181). Let x be the singular term *Davidson*, so slingshot for these terms can be stated as:

- (1) “Ankara is more populated than İzmir”
- (2) “(the x such that x = Davidson, and Ankara is more populated than İzmir) = (the x such that x = Davidson)”
- (3) “(the x such that x = Davidson, and Tarski is a mathematician) = (the x such that x = Davidson)”
- (4) “Tarski is a philosopher”

(1) and (2) are logical equivalences, and according to A1 they have the same meanings. Also (3) and (4) are logical equivalences and they have same meanings. If *Ankara is more populated than İzmir* and *Tarski is a mathematician* are true, then *(the x such that x = Davidson, and Ankara is more populated than İzmir) = (the x such that x = Davidson)* and *(the x such that x = Davidson and Tarski is a mathematician) = (the x such that x = Davidson)* have the same reference—namely, *Davidson*. So according to A2, they have same meanings. The result is that all these singular terms refer to same meaning. The same line of argument applies to all true

sentences, and as a result of this, they all refer to the same meaning. According to Davidson, it is “an intolerable result” (1967/2001b p. 19); so we must abandon the view that sentences refers to their meanings.

It is worth noting that Davidson does not only use slingshot argument against meanings as entities. He states that “the argument does not depend on any particular identification of the entities to which sentences are supposed to refer” (1967/2001b, p.19). Accordingly, Davidson (1969/2001c) invokes the slingshot argument against the view that true sentences correspond to facts and against the claim that “declarative sentences ... refer to *propositions, states of affairs, or situations*” (Neale, 2001, 9). For the former the argument shows that all sentences correspond to one fact; and for the latter they refer to *The True* or *The False*.

M_1 may be modified by replacing *means* by *means that* in which p can be a sentence instead of an entity:

(M_2) s means that p

where s is the structural description of a sentence and p is a sentence which gives the meaning of s . Davidson (1967/2001b) argues that M_2 would not be the form of the theory because the predicate *means that* is “non-extensional” (p. 22) and one need to invoke intensional logic for the deduction of meaning theorems. So it would be a mistake to use *means that* in the theorems of a meaning theory because its logic presupposes meaning (Glock, 2004, p. 148; Kölbel, 2001, pp. 614-615): “For it is reasonable to expect that in wrestling with the logic of the apparently non-extensional ‘means that’ we will encounter problems as hard as, or perhaps identical with, the problems our theory is out to solve” (Davidson, 1967/2001b, p. 22).

After rejecting alternative predicates, Davidson (1967/2001b) concludes that the “success of our venture depends not on the filling but on what it fills” (pp. 22-23) and seeks a suitable notation which will fill the blank at “ $p\dots s$.” According to Davidson, the desired structure may be sustained by two alternatives:

(Case₁) The meta-language contains the object language

(Case₂) The meta-language does not contain the object language

For Case₁, *s* is the structural description of the sentence and *p* is the sentence itself. For Case₂, *s* is the structural description of the sentence and *p* is the translation of it into the meta-language. Davidson offers a connective in first order predicate logic and a predicate which transforms the structural description of a sentence into a sentence to fill the blank:

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

The structure of the meaning theory is formed by this last move, and what comes next is to clarify in which way *p* gives the meaning of *s*:

- (a) What we require of a theory of meaning for a language is that without appeal to any (further) semantical notions it place enough restrictions on the predicate 'is T' to entail all sentences got from schema T when 's' is replaced by a structural description of a sentence of L and 'p' by that sentence. (Davidson, 1967/2001b, p. 23)

Tarski's truth theory enters the scene at this point. Requiring the theory to entail all sentences got from schema T when *s* is replaced by a structural description of a sentence of L and *p* by that sentence is identical with requiring the theory to satisfy Convention-T. If the theory satisfies Convention-T, then the predicate *is T* is extensionally equivalent to *is true*. If we ignore for a moment *how* such a theory can be constructed, or be sure it satisfies Convention-T, it is clear that a truth theory which qualifies to serve as a meaning theory *has to* satisfy Convention-T. So I suggest adding this condition to Davidson's claims about the relationship between a truth theory and a meaning theory, and truth condition of a sentence and meaning of a sentence:

- (b) ...a theory of meaning for a language L shows 'how the meanings of sentences depend upon the meanings of words' if it contains a (recursive) definition of truth-in-L [which satisfies Convention-T.] (Davidson, 1967/2001b, p. 23)
- (c) ...and to give truth conditions is a way of giving the meaning of a sentence [if and only if it is an instantiation of Schema-T.] (Davidson, 1967/2001b, p. 24)

Tarski gives the truth definition for a fully interpreted formal language. Now in Davidson's case, the language for which a truth theory will give an account is neither formal nor interpreted but an un-interpreted natural language. Tarski (1944) claims that it is impossible to construct a truth theory for a natural language. So Davidson's programme has to deal with two branches of problems. One branch of technical problems requires answers to such questions as, (a) how the universality of language and the presence of ambiguous terms can be accounted for, and (b) how structures which are present in natural language such as indexicals, demonstratives, "sentences that attribute attitudes, modalities, quantifiers like 'most' and so on" (Davidson, 1973/2001e, p. 132) are dealt with. Another branch of problems will include (c) how an empirical account of a truth theory can be given.

Although most of the technical part may be left incomplete by pointing to a programme whose aim is to solve the technical problems, giving an empirical account is crucial. Davidson defines a vague predicate (*is T*) and claims that such a predicate can be said to be a meaning predicate if and only if it serves to give the meaning of an object language sentence for which it is predicated. Suppose that we have a language with only two sentences (*snow is white* and *coal is black*). So Davidson requires his predicate to entail the T-theorems

"Snow is white" is T in English if and only if snow is white

"Coal is black" is T in English if and only coal is black

A meaning theory is a theory which entails all and only these T-theorems. If the theory does not entail correct theorems or entail wrong theorems (such as "*Snow is white*" is T if and only if *coal is black*), then the theory is not a meaning theory and *is T* is not qualified to be a meaning predicate. So, the predicate *is T* is a meaning predicate if and only if it satisfies Convention-T. If such a theory satisfies Convention-T, then the meaning predicate is at least co-extensive with the truth predicate. As Davidson has no idea how to construct a meaning theory without appeal to semantic concepts, he proposes to construct a truth theory which may serve as a meaning theory. And such a claim is justified only if it is possible to construct an empirical truth theory which satisfies Convention-T.

Two interlocking consequences of Davidson’s argumentation about the form of a meaning theory can be stated as: (a) There may be other ways to construct a meaning theory, and (b) a truth theory is not a meaning theory. Davidson does not claim that it is the only form a meaning theory could take, but his proposal is a technical maneuver which he appeals to because he has “no other idea how to turn the trick”(Davidson, 1967/2001b, p. 23). In fact, Davidson’s overall argumentation would fail if he had said that the predicate *is T* is necessarily the truth predicate, since there is no such necessity. Davidson (1976/2001h) explicitly states that “no matter how well selected,” a truth theory “is not a theory of meaning” (p. 179). But, at least, knowledge of a truth theory, which is constructed by the lines Davidson proposed, will suffice to interpret utterances of the speaker. These consequences reveal the nature of Davidson’s meaning theory and will be revisited in section 4.2 during the discussion of whether interpreters actually construct truth theories to understand speakers or whether they have *tacit knowledge* of the theory.

Although it is clear that a truth theory for a natural language is an empirical theory and there is no way to construct one a priori, Davidson often seems to talk about a theory which is already constructed and to consider how to verify it. But such considerations may confuse one by leading one to think that Davidson advocates the possibility of constructing a formal truth theory for a natural language, or that he inaccurately uses Tarski’s translation criterion in doing so. But Davidson never makes either claim. Being empirical is the very nature of a truth theory for a natural language and he emphasizes this along with his bold proposal. So, how can such an empirical programme carried out? The only hint is that the theory must be “recursively characterized” (1967/2001b, p. 23), and Tarski uses translations of a finite number of primitives in his axioms freely through the recursive characterization of truth theory:

(s)(n)(s satisfies “n is white” iff ref(n, s) is white.)

(s)(n)(s satisfies “n ist weiss” iff ref(n, s) is white.)

Call axioms of this kind “interpretive” axioms. If the truth theory is composed of such interpretive axioms—so that it ipso facto satisfies Convention-T—then call it an

“interpretive truth theory” (Lepore & Ludwig, 2005, p. 72). But such a procedure is not available for Davidson, as he does not want to use semantic notions such as translation, meaning, or synonymy. In principle, evidence that can be used to *verify* the consequences of an (empirical) theory may be used to *construct* the theory. I suggest to interpret Davidson’s talk of *verifying* and *testing* theory as *constructing* theory where necessary.¹⁷ Davidson claims that theory can be verified by checking whether T-theorems are true (1967/2001b, pp. 25, 27; 1970/2001d, pp. 61-62). So, in this initial case the only way to construct a truth theory is to determine the axioms of the theory by assigning temporary axioms according to true T-sentences and correcting them according to further entailments. Although such an empirical theory is “doomed to be to some extent incomplete and schematic” (1970/2001d, p. 59), in principle, for infinitely many sentences, it will fix axioms of the theory as interpretive.

It is said that what Davidson wants from a truth theory is just extensional adequacy: “Davidson proposes that a merely extensionally adequate truth theory for a natural language (i.e. one that is simply true) would thereby meet Tarski’s Convention T or an analog for natural languages” (Lepore & Ludwig, 2005, p. 75). This claim seems true, but is misleading. Consider the quotation (b) and quotation (c) above. If Lepore and Ludwig’s claim is true, then it is a mistake to add further conditions to them. But it is clear from the quotation (a) that Davidson considers a truth theory which satisfies Convention-T to be qualified to serve as a meaning theory, no matter *how* such a theory satisfies it.¹⁸ One must seek for the extensional adequacy of the truth theory in empirical applications, not at the statement of the

¹⁷ Direct textual evidence for such claim may be found in “Reply to Foster” (1976/2001h) where Davidson considers his initial claims about truth and meaning as “...I imagined the theory to be known by someone who had constructed it from the evidence...” (p. 173). Also note that the term *construction* did not refer to an actual construction but a hypothetical one.

¹⁸ It is also hard to see how one can agree to the further condition in quotation (c) but reject the one at (b), which is the Lepore and Ludwig’s case: “We must understand the expression ‘truth conditions’ so that a T-form sentence does not give truth conditions of a sentence unless the sentence used on the right hand side of the biconditional interprets the sentence mentioned on the left” (Lepore & Ludwig, 2005, p. 95).

relationship between the truth and meaning theories. In fact, mere external adequacy cannot fix the axioms of theory as interpretive without indexicals and demonstratives which can be invoked only in empirical applications.

Natural languages contain demonstratives and indexical elements. These elements are vital for constructing an empirical truth theory and it can be said that without them such an endeavor will be sentenced to failure.¹⁹ In empirical application, it is not the sentences but utterances of such sentences that take truth values, and they take different truth values at different times and for different persons. *I am tired* is true if the person who utters it is tired at that moment. So Davidson (1967/2001b) proposes to view “truth as a relation between a sentence, a person, and a time” (p. 34) and to revise the form of the theorems which a truth theory is meant to entail as follows:

‘I am tired’ is true as (potentially) spoken by p at t if and only if p is tired at t.

‘That book was stolen’ is true as (potentially) spoken by p at t if and only if the book demonstrated by p at t is stolen prior to t. (p. 34)

Of course, axioms which entail such a form for T-theorems, and Convention-T itself, must be revised accordingly. Although Davidson does not attempt to give an explicit presentation of such modifications, one can restate Convention-T as: A truth theory is materially adequate if and only if it entails all and only instances of the following schema:

For any speaker *S*, time *t*, *s* for *S* at *t* is T in L iff *p* (Lepore & Ludwig, 2005, p. 84)

where *s* is the structural description of the object language and *p* is a meta-language sentence which gives the meaning of *s* for *S* at *t*.

A truth theory which satisfies this modified Convention-T may serve as a meaning theory. The reference axiom for indexical elements will be of the form

¹⁹ The importance of indexical elements is consistently emphasized in Davidson’s paper’s even before his “Radical Interpretation” (1973/2001e): “Sentences with demonstratives obviously yield a very sensitive test of the correctness of a theory of meaning, and constitute the most direct link between language and the recurrent macroscopic objects of human interest and attention” (1970/2001b, p. 35). They correspond to Quine’s *observation sentence*. See section 3.1.

$(x)(e)(t) \text{Ref}("I", s, e, t) = e$, $(x)(e)(t) \text{Ref}("that", s, e, t) = \text{the object demonstrated by } e \text{ at } t$ and accordingly $(x)(e)(t) \text{Ref}("snow", s, e, t) = \text{Snow}$ where e is the utterer, and t is the time of the utterance. The Satisfaction axioms will be revised as $(s)(n)(s \text{ satisfies}(e, t) \text{ "n is white" iff } \text{ref}(n, s, e, t) \text{ is white})$.²⁰ Davidson's extensional adequacy condition will rule out such un-interpretive axioms as $(x)(e)(t) \text{Ref}("snow", s, e, t) = \text{Milk}$, which result in the theory entailing true but un-interpretive T-theorems such as "*Snow is white*" is true if and only if milk is white because the same theory will include the axiom $(x)(e)(t) \text{Ref}("that", s, e, t) = \text{the object demonstrated by } e \text{ at } t$, where the demonstrated object is *snow*. The axioms of the truth theory will be fixed for the totality of T-theorems and the truth theory will entail T-theorems which satisfy the (modified) Convention-T.

²⁰ Revised axioms are modified from Neale (2001, p. 41) and Larson and Segal (2005, pp. 85-86).

CHAPTER 3

RADICAL INTERPRETATION

It is a fact that we can understand what others' utterances mean, and Davidson (1973/2001e) raises two questions about this ability: (a) "What could we know that would enable us to do this?" and (b) "How could we come to know it?" (p. 125). Davidson's (1994) answer to the former is that we *could* know a truth theory "constructed more or less along the lines of one of Tarski's truth definitions" (p. 126)—a recursive one—and such a theory *would* enable us to understand the meaning of the utterances of others. The answer to the second question addresses the procedure of how such a construction may be carried out. In his early papers, Davidson claims that the procedure must be something like Quine's *radical translation* (Davidson, 1967/2001b, pp. 25, 27; 1970/2001d, pp. 61-62). Later he acknowledged the influence of Quine over his own version—namely, radical interpretation.

3.1) Behaviorism, Quine and Radical Translation

Behaviorism is a doctrine which has different forms and which dominated psychology and philosophy from the 1920s to the end of the 1960s. In the domain of psychology, it can be seen as an attempt to develop a scientifically respectable methodology for psychology whose main tool through its investigations has been introspection, which seems far from objective. Accordingly, "psychology is the science of behavior" (Gallistel, 1998, p. 696), and the study of such behavior can be carried out upon observable behavior without appeal to the mental. In the domain of philosophy, it corresponds to the thesis that mental concepts can/must be reduced/eliminated in favor of behavioral ones (Graham, 2007). The underlying

motive of both types of behaviorism is to establish a scientific ground for their course of study, namely, the third-person point of view which is the position of a scientist constructing and testing theories.

The same motive is central to Quine's approach to philosophy. According to Quine, "philosophy and science are continuous with one other" (Smart, 1969, p. 3). Quine (1951) rejects a status for philosophy's statements as a priori, and attacks the analytic/synthetic distinction to weaken the "supposed boundary between speculative metaphysics and natural science" (p. 20). Quine claims that analytical truths are true by virtue of synonymy (sameness of meaning) such as *bachelor* is synonymous with *unmarried man* in the so-called analytic truth *All un-married men are bachelors*. But the notion of synonymy is as problematic as that of analyticity, because the definition of such a concept presupposes the notion of analyticity itself. Grice and Strawson (1956) raise the following question: "Is all talk of correct or incorrect translation of sentences of one language into sentences of another meaningless? It is hard to believe that it is" (p. 146). They defend this so-called dogma in virtue of meaning. Quine (1988) answers their question through his influential indeterminacy thesis, which is illustrated by radical translation.

Quine adopts a behaviorist stance towards language. All the available evidence for a pre-linguistic infant to acquire such an ability is the observable verbal behavior of others, and the correction of his behavior by others:

Language is a social art. In acquiring it we have to depend entirely on intersubjectively available cues as to what to say and when. Hence there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable stimulations. (Quine, 1988, p. ix)

Quine (1988) develops the thought experiment of *radical translation* in line with this behaviorist approach. The aim of Quine's radical translator is to construct a translation manual for a completely unknown language—for example, for a language of a recently discovered jungle tribe, namely Jungle—which will include a dictionary and a grammar for translation of Jungle sentence. Although the radical translator already possesses a language, he can mimic first language acquisition because the

evidence available is (supposed to be) the same with such a case. For example, in the presence of a rabbit, a tribesman says *Gavagai* and the translator notes that it can be translated as *Rabbit*. In further cases in the presence (and also in the absence) of a rabbit, the translator says *Gavagai?* and observes the *assent* or *dissent* behavior²¹ of the tribesman to test the appropriateness of the translation. Quine claims that what makes the tribesman assent or dissent is not the rabbit but the rabbit stimulation. In this case the stimulation is the pattern of his activated nerve cells of his eyes which is caused related with the presence of a rabbit in his field of view. An “observation sentence”²² is translated into another as having the same “stimulus meaning,” because what is common in both tribesman and translator seems to be such stimulations or patterns of stimulations. Quine (1988) defines stimulus meaning as an “ordered pair” (p.33) of affirmative and negative stimulus meanings. The former is the stimulation patterns which cause the tribesman to assent and the latter is the stimulation patterns which cause the tribesman to dissent. Although observation sentences sustain empirical evidence for the translation of a set of jungle sentences, there will be many others which cannot be translated with the guidance of assent/dissent behavior. If we assume that Jungle consists of infinitely many sentences, then the manual has to give meaning of each Jungle sentence according to its parts and their mode of combination. Quine’s (1988) proposal is to construct “analytical hypotheses” (p. 68) which couple Jungle words and phrases with those of the translator’s language. These analytical hypotheses “constitute the linguist’s jungle-to-English dictionary and grammar” (p.70). This matching will be guided by the parts of the observation sentences which were translated before. Once they are determined, the analytic hypotheses can be used to translate sentences of Jungle.

Quine (1988) extracts his famous thesis of the *indeterminacy of translation* from the procedure of radical translation: “Manuals for translating one language into

²¹ Even the assent or dissent behavior of the tribesman must be discovered through observation. See Quine (1988, pp. 29-30)

²² Quine(1988) defines observation sentences as those “occasion sentences whose stimulus meanings vary none under the influence of collateral information” (p. 42) where the occasion sentences are sentences “which command assent or dissent only if queried after an appropriate prompting stimulation” (pp. 35-36).

another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with one other” (p. 27). There are two sources of this conclusion. One is that evidence available may vary from one translator to the other and the translators may form different analytical hypotheses from different sets of observation sentences. The other is the *inscrutability of reference*. Although stimulus meanings can be used to determine translations of Jungle sentences, sameness of stimulus meaning for observation sentences does not guarantee sameness of the extension of words which are parts of analytical hypotheses:

For, consider [the term] “gavagai.” Who knows but what the object to which this term applies are not rabbits after all, but mere stages, or brief temporal segments, of rabbits? In either event the stimulus situations that prompt assent to [the observation sentence] “Gavagai” would be the same as for “Rabbit”. (Quine, 1988, pp. 51-52)

So there will be alternative analytic hypotheses which fit the overall evidence from behavior equally well.

3.2) Davidson’s Revision of Radical Translation

At the very beginning of this chapter, it was noted that Davidson uses some form of Quine’s radical translation thesis for the empirical application of a truth theory, before developing his own version. But such a close relation must not lead one to think that one is just a mere restatement of the other. Quine and Davidson present radical translation and radical interpretation for different purposes. Quine’s aim in radical translation is to eliminate intensional notions in favor of behavioristic counterparts so as to attain a scientific language. Davidson (1967/2001b) does not want “to change, improve, or reform a language, but to describe and understand it” (p. 29) and radical interpretation is part of Davidson’s endeavor to give a meaning theory for a natural language that answers the question of how we come to have knowledge which suffices for understanding others. Despite the presence of several technical differences, they both share a general third person point of view in approaching language affirming that all evidence is open to public observation. Respectively, important consequences of both thought experiments such as

indeterminacy of translation/interpretation and inscrutability of reference conjoin. Technical differences are due to Davidson's departure from Quine's behaviorism. Davidson (1999b) prefers (a) "distal stimuli" instead of Quine's "proximal stimuli," (b) the attitude of holding a sentence true instead of "assent or dissent." There is a "causal chain from a thing or event in the external world to the act of utterance by which the subject reports it" (Quine, 1999, p. 74). Recall that what makes Quine's tribesman to assent or dissent is stimulation of that thing or event—which is the proximal stimulus of that thing. Contrarily, Davidson prefers to use the cause of this proximal stimulus which is the object or event itself—namely, a distal stimulus.²³ Davidson departs from Quine's behaviorism because he counts the attitude of holding a sentence true as evidence which is a belief of the speaker.

3.3) Radical Interpretation and Truth Theory

Recall that Tarski's purpose is to define a predicate (*is T*) and show that it is the truth predicate (*is true*). He presupposes translation as a criterion to ensure that the predicate applies all and only to the true sentences of a language and without it there is no reason to call the predicate *is T* the truth predicate.

Platts (1997) criticizes Davidson's bold proposal for presenting the relationship between the truth and meaning predicates in an illegitimate way: If one aims to construct a meaning theory and all which one has a null meaning predicate *is X*, then without translation criteria one cannot decide whether it is the truth predicate or not. So, Davidson's bold proposal is wrong because one cannot "identify the predicate required for the theory of meaning, 'is X', with a Tarskian truth-predicate" (p. 56). Also consider the following objection which is along the same line: "...without knowing the meanings of sentences of L, one could not determine whether 'T in L' is a truth predicate for L ... Similarly, without knowing 'T in L' is a truth predicate...one could determine nothing about the meaning of sentences of L" (Soames, 1999, p. 105).

²³ The reason for Davidson's preference is beyond the scope of this thesis. Briefly, Davidson rejects Quine's proximal stimuli because he rejects what he calls the third dogma of empiricism which assumes a scheme-content dichotomy. See Davidson (1974/2001g).

Now, I want to claim that the main problem is not with stating a relationship between truth and meaning in an illegitimate way, as per above objections. Davidson never invokes translation criteria for devising a truth theory which may serve as a meaning theory, and Davidson's bold proposal is independent of questions about how such a theory can be devised (see section 2.3). The problem is in the empirical application of a truth theory to an actual language. Davidson uses the truth predicate freely to construct an extensionally adequate theory as presented in section 2.3, which satisfies Convention-T after fixing its axioms as interpretive. But the oddity of this suggestion is obvious: One cannot presuppose the concept of truth in defining truth. Davidson (1973/2001e) accepts his mistake²⁴ and proposes to "reverse the direction of explanation: assuming translation, Tarski was able to define truth; the present idea is to take truth as basic and extract an account of translation or interpretation" (p. 134). Although Davidson explicitly presents this methodology in "Radical Interpretation" it is even available in earlier works. Davidson's aim in "Radical Interpretation" is to justify his methodology, not to change the old one by declaring it to be wrong.

What is it to say that a sentence or utterance is true for someone? In the simplest case, someone holds his sentence to be true because he knows what this sentence means and he believes that it is the case. Also, someone holds a sentence false because he knows what this sentence means and believes it not to be the case. So, "knowing that he holds the sentence to be true, and knowing the meaning, we can infer his belief; given enough information about his beliefs, we could perhaps infer the meaning" (Davidson, 1973/2001e, p. 134).

Let *S* be a speaker who holds the sentence *snow is white* as true. If you know *snow is white* means that snow is white in *S*'s language then you can infer that *S* believes that snow is white (*S* would hold *snow is white* to be false if he had believed that snow is not white). Again, *S* holds the sentence *snow is white* to be true. If you know that *S* believes that snow is white, and that *S* is rational and holds the sentence

²⁴ "My mistake was to think we could both take a Tarski truth definition as telling us all we need to know about truth and use the definition to describe an actual language. But even in the same essay [Truth and Meaning] I (inconsistently) discussed how to tell that such a definition applied to a language" (Davidson, 1990, p. 286).

true because of his belief, then you can infer that *snow is white* means that snow is white in *S*'s language. But suppose that you do not know what he believes about his environment, and you do not know what his utterances mean. The only evidence available is your ability to identify the utterances which the speaker holds true and your own beliefs about the environment. This is the position of a radical interpreter. Davidson claims that belief and meaning are interdependent, and the radical interpreter has to construct a meaning theory and a theory of belief at the same time. Now, an example of such evidence can be formulated as:

[Type-1] Kurt belongs to the German speech community and Kurt holds true 'Es regnet' on Saturday at noon and it is raining near Kurt on Saturday at noon. (Davidson, 1973/2001e, p.135)

Kurt and other people belonging to the same speech community may utter such a sentence at different times and the interpreter may produce an inductive formulation based on such evidence as:

[Type-2] (x)(t)(if x belongs to the German speech community then (x holds true 'Es regnet' at t if and only if it is raining near x at t)). (Davidson, 1973/2001e, p.135)

Davidson invokes the *principle of charity*²⁵ which states that people's beliefs about their environments are usually true. In the light of this principle, the above generalization may be used to justify the following claim:

[T-Sentence] 'Es regnet' is true-in-German when spoken by x at time t if and only if it is raining near x at t. (Davidson, 1973/2001e, p.135)

Davidson takes the notion of truth as a primitive. Through the fact that Tarski aims to give a definition of truth, this *explicit definition* cannot be used in Davidson's programme. Davidson proposes to cancel Tarski's set theoretic step in his definition (see section 2.2) to treat the truth definition as an axiomatic theory. The undefined

²⁵ Although Lepore and Ludwig (2005) formulated this version of the principle of charity as: "For any speaker S, time t, belief b, *ceteris paribus*: b is a belief of S's at t about and prompted by S's environment iff b is true" (p. 189).

notions in such a theory will be reference, satisfaction and truth. Even in Tarski's case, when the theorizer's aim is to define the truth, a pre-theoretical grasp of truth is presupposed in Convention-T, and Davidson takes truth to be basic in relation to satisfaction and reference for the sake of pre-theoretical insight (Davidson, 1990, p. 299). Although the truth of sentences in constructing and testing the theory is taken as basic, the appeal to the concepts of reference and satisfaction is necessary. Without them, one cannot satisfy the compositionality requirement. But Davidson (1973/2001e) claims that such notions must be treated as "theoretical constructions, beyond the reach of direct verification" (p. 133), and they exhaust their role through the entailment of T-theorems.²⁶

The procedure which is followed by a radical interpreter is similar to the procedure of a radical translator. The interpreter starts to devise a truth theory by determining type-1 sentences from the behavior and utterances of the speaker. He tests the correctness of these sentences by uttering them in relevant environmental conditions and observing speaker's attitude of holding that sentence true. After the confirmation of the correctness of type-1 sentences, the interpreter infers type-2 sentences from them and he constructs T-sentences from type-2 sentences in light of the principle of charity. Next, he identifies the axioms of the theory which best fit the totality of T-sentences, which allows derivation of the T-sentences from themselves. The correctness of the axioms is tested indirectly by observing speaker's attitudes of holding T-theorems true which are derived from the axioms. After indefinitely many cases of testing and correction, the theory will suffice to enable the interpreter to understand the object language sentences (Davidson, 1973/2001e, pp. 136; 1974/2001f, pp. 150-152).

If the formal criterion of translation is used in constructing the theory, the T-theorems which are entailed from it will be logical truths and there will be no room for indeterminacy. But, in empirical application, there will remain a number of theories which can be constructed from the same evidence.

If we consider the relationship between a truth theory and a meaning theory in light of the procedure of radical interpretation again, it adds nothing to the initial

²⁶ Compare with Field (1972) where reference and satisfaction are treated as primitives.

claim about what kind of truth theory is qualified as a meaning theory, namely an interpretive one. Davidson does not aim to develop theorems which give the meanings of sentences directly. Empirical evidence is collected in the form of truth biconditionals, not meaning theorems like “‘Es regnet’ in Kurt’s language is translated as ‘It is raining’ in mine” (Davidson, 1973/2001e, p. 129). So the process presupposes that the meaning theory will take the form of a truth theory. As discussed in chapter 2.3, the role of radical interpretation is to show that it is possible to construct an interpretive truth theory without appeal to translation criterion.

Till now, it has been considered how one can justify the claim that a truth theory may serve as a meaning theory. But how a truth theory gives the meaning of utterances has been left untouched. At first sight, it seems that the T-theorems which are entailed from the theory give the meanings of the object language sentences, while Davidson claims that “...to give truth conditions is a way of giving the meaning of a sentence” (1967/2001b, p. 24). But T-theorems are just material biconditionals. There also exist enumerable un-interpretive biconditionals as well as interpretive ones. So entailment from the theory must be addressed before saying that a T-theorem gives the meaning of a sentence. Such an entailment condition can be satisfied by requiring the proof from the axioms of the theory in addition to T-theorems. Of course, it is assumed that the theory which entails such T-theorems is interpretive, so one also knows that the theory satisfies formal and empirical constraints which are easy for one who constructed the theory according to these constraints (Davidson, 1973/2001e, p. 138-139). If these conditions are satisfied, then the predicates *means that* and *is true if and only if* are *salva veritate* and it can be said that giving the truth conditions of a sentence is a way of giving the meanings of it.

CHAPTER 4

THE COGNITIVE APPROACH

4.1) The Cognitive Revolution and Chomsky

Miller (2003) defines the cognitive revolution as a “counter-revolution” (p. 141) against the behaviorist movement whose aim was to eliminate any talk of common sense psychological concepts such as belief and desire. Recall that behaviorism declared such concepts unscientific by insisting that all there is to study is stimulus and observable behavior relations. But ironically, behaviorism blocked scientific development in psychology:

So long as behaviorism held sway—that is, during the 1920s, 1930s, and 1940s—questions about the nature of human language, planning, problem solving, imagination, and the like could only be approached stealthily and with difficulty, if they were tolerated at all. (Gardner, 1985, p. 12)

The resources of behaviorism turn out to be too poor to explain complex behavioral patterns of human beings. Advancements through different areas such as neurology, artificial intelligence, linguistics and the explanatory success of the information processing method in experimental psychology (Gardner, 1985, chap.1; Miller, 2003, p. 142) also encouraged the appeal to internal mechanisms. This paradigm shift is now labeled as the cognitive revolution.

Chomsky can be regarded as one of the founding fathers of cognitive science due to his radical reformation of linguistics. He challenged the behavioristic approach to language through his influential review of Skinner’s (1957) book *Verbal*

*Behavior*²⁷ by claiming that Skinner's attempt to extend principles which are determined from animal behavior to human language mastery and learning is not justified and that such an approach is a methodological dead end because of the *species specific* nature of language:

One would naturally expect that prediction of the behavior of a complex organism (or machine) would require, in addition to information about external stimulation, knowledge of the internal structure of the organism, the ways in which it processes input information and organizes its own behavior. These characteristics of the organism are in general a complicated product of inborn structure, the genetically determined course of maturation, and past experience. (Chomsky, 1959, p. 27)

Chomsky does not deny the fact that theory must be in accordance with empirical evidence and observation, but objects to the claim that theory is all about evidence alone. Evidence must be used to discover “hidden and underlying laws” (Searle, 1972, ¶ 3) which manifest themselves by producing such evidence—the methodological suggestion of behaviorism is equivalent to calling “physics a science of meter reading” (¶ 2): “The cognitive perspective regards behavior and its products not as the object of inquiry, but as data that may provide evidence about the inner mechanisms of mind and the ways these mechanisms operate in executing actions and interpreting experience” (Chomsky, 2000, p. 5). The behaviorist approach was the core of American structuralism, together with logical positivism, which was the dominant view in linguistics. According to American structuralism, linguistics is “a descriptive and taxonomic science, like zoology, geology, and astronomy” (Harris, 1993, p.27) and all mental considerations are out of the scope of science. Success of Chomsky's new conception of science and his refutation of structuralist claims systematically made generative grammar a new orthodoxy.

Chomsky (1965) makes a distinction between *linguistic competence* and *performance*. Competence is what the speaker knows and performance is how the speaker uses this knowledge. Suppose that a boy could once ride a bike, but he

²⁷ Chomsky, in a letter which he wrote to the author of his biography, admits that the real object of his criticism is not Skinner but Quine: “It wasn't Skinner's crazy variety of behaviorism that interested me particularly, but the way it was being used in Quinean empiricism and naturalization of philosophy” (Barsky, 1998, p.99).

cannot at present because he broke one of his legs, or suppose that he used to read but he cannot now because it is too dark. It is absurd to claim that the boy does not know how to ride or read in these cases. There are many factors, both internal and external, of the speaker which affect his linguistic performance. The object of scientific study is the competence of the speaker,²⁸ which governs linguistic performance in addition to such factors. An emphasis on competence identifies the scope of the study as knowledge of an ideal speaker—knowledge of an “internalized ... grammar, a system of rules and principles that assigns structural descriptions to linguistic expressions” (Chomsky, 1981, p. 223). Anyone is a competent speaker of a language if the knowledge of such grammar is represented in his mind/brain (Chomsky, 1980, p. 5). Chomsky uses the term *tacit knowledge* to identify the relation between the competent speaker and the grammar which contrasts with linguists’ propositional knowledge of the grammar as being unconscious.

If a newborn is exposed to English utterances, during a critical period, he will acquire the grammar which will manifest itself as English just as in the cases of other languages such as Turkish, German, etc. Does this mean that the infant *learns* the grammar of such a language from exposure to utterances where learning indicates the infant’s grasp of rules and conventions from scratch? Arguments against such a view can be categorized as “poverty of the stimulus” (Chomsky, 1980, p. 34). On such an argument, the quantity and quality of the infant’s linguistic exposure is not sufficient to explain the uniformity of the stages of development of language acquisition and the infant’s apparent rule following verbal behavior in contrast to un-grammatical stimuli (Lund, 2003, pp. 53-54). Chomsky claims that human beings possess an innate grammar which takes the shape of particular grammars through language acquisition:

We may think of universal grammar as, in effect, the genetic program, the schematism that permits the range of possible realizations that are the possible human languages. Each such possible realization is a possible final steady state, the grammar of a specific language. Universal grammar is a system that is genetically determined at the initial state, and is specified,

²⁸ Note that the object of scientific study is linguistic performance rather than linguistic competence, according to behaviorism. See Chomsky (1969).

sharpened, articulated, and refined under the conditions set by experience, to yield the particular grammars that are represented in the steady states attained. (Chomsky, 1980, p. 234)

Just as human beings are genetically determined to have legs, arms, and a circulation system, they are also determined to have a *faculty of language*²⁹ which is an internal component of the mind/brain specialized for language mastery. The scope of linguistic study now can be extended to universal grammar. Specifically linguistic study is the construction of theories which characterize universal grammar in addition to grammars of particular languages (Chomsky, 2006, p. vii) where universal grammar is the *initial state* of the faculty of language and particular grammars are about the *steady states* of such a language organ (Chomsky, 2002, p. 64). Chomsky (1995) presents the “principle and parameters approach” (p. 17) to explain universal grammar and its variations as particular grammars. Principles are defined as the initial state which is innate, species specific and common in all languages and parameters are defined as such principles’ flexibility to form different grammars with different but limited variations (pp. 14-19).

Chomsky’s discussion of universal grammar leads him to the nativist tradition of Descartes which states that human beings possess tacit and innate knowledge. Now the question is how much of our language knowledge is innate? The principles and parameters approach indicates a set of possible forms from which the syntax of language is chosen and fixed. But what about lexical items? Although early Chomsky identifies the questions about “universal semantics” as inconclusive and does not give a clear account of whether the meanings of lexical elements are included in universal grammar (Katz, 1980), he later strengthens his nativism by claiming that the meanings of lexical items are innate too (Chomsky, 1997). He claims that the meanings of lexical items are restricted by universal grammar just like the restriction and identification of possible particular grammars. The argument from the poverty of the stimulus, together with the apparently rapid acquisition of meanings of words by infants leads him to the idea that “the concepts are already available, with all or much of their intricacy and structure predetermined, and the

²⁹ See Hauser et al. (2002) and Fitch et al. (2005) for recent discussions.

child's task is to assign labels to concepts, as might be done with very simple evidence" (Chomsky, 1997, p. 29).

From above considerations, we arrive at a bold position which may be labeled as the *cognitive approach*. The cognitive approach is used in a broad sense to refer to any position which appeals to mechanisms which are internal to the mind/brain in its explanations of behavior. The philosophical implications of such an approach are very diverse and open to controversy (Horst, 2005; Pitt, 2008). In a more restricted sense—which accords with the subject matter of this thesis—the cognitive approach refer to a Chomskian position which includes the following claims:

1. The tacit knowledge of a semantic theory constitutes the semantic competence of speakers of a language.
2. The theory is represented in the mind/brain of the language user.
3. Language acquisition of an infant includes acquisition of the knowledge of the theory.³⁰

4.2) Competence, the Cognitive Approach and Radical Interpretation

Recall the two questions which Davidson used to guide his investigation of meaning theory and his answers to such questions about our understanding of other's utterances (1) "What could we know that would enable us to do this?" and (2) "How could we come to know it?" (Davidson, 1973/2001e, p. 125). Davidson's answer to the former is that we *could* know a recursive truth theory "constructed ... along the lines of one of Tarski's truth definitions" (Davidson, 1994, p. 126). This *would* enable us to understand the meanings of the utterances of others. The answer to the second question addresses the procedure of how such a construction may be carried out, namely by radical interpretation. Davidson claims that former of such answers does *not* state what an interpreter *actually* knows, and the latter does not state how

³⁰ The first two of these conditions are necessary in the sense that any Chomskian cognitive approach must contain. The approach also has to contain a third claim about the acquisition of language or the growth of language organ. The acquisition model, which is defined as learning under the restrictions of universal grammar, is preferred in this thesis.

language acquisition takes place or how an actual field linguist works. I call the collection of these questions and Davidson's proposed answers *the philosophical approach*. Now compare these questions with their revised versions:

(1*) What do we know that enables us to do this?

(2*) How do we come to know it?

The answers to such questions reveal the aim of the cognitive approach. The meaning theory must represent the semantic competence of language users. This means that the theory must be about an *actual* mechanism that is internal to the mind/brain. In other words, language users must possess tacit knowledge of such a theory. The answer to second question will indicate how interpreters acquire their languages.

It has been said that early Davidson seems to describe the linguistic competence of an actual language user (Glock, 2004, pp. 244-249; Lepore & Ludwig, 2005, p. 122). However, later he explicitly rejects such an endeavor at the very beginning of "Radical Interpretation" (1973/2001e) and adopts the philosophical approach. What is the aim of the philosophical approach? Davidson claims that there are concepts which resist philosophers' attempts to reduce them to more basic ones, or to give analyses/definitions in terms of more basic ones. He introduces an alternative methodology which is based on two morals which he draws from his early work on decision theory:

One was that by putting formal conditions on simple concepts and their relations to one another, a powerful structure can be defined ... The second lesson was that a formal theory says nothing about the world. It defines an abstract structure. There remains the question whether that structure is realized in the world, or even in some abstract domain of numbers or sets. (Davidson, 1999a, p. 32)

Davidson (2004) takes *belief*, *desire* and *meaning* as instances of concepts which invoke endless controversies when taken as subjects for traditional philosophical methodology. He defines a *unified theory* which relates abstract structures that are assigned to such notions—truth theory for meaning and decision

theory for belief and desire—with the principle of charity to illuminate such problematic concepts. Such abstract structures must be realized in the actual world and Davidson defines their realization as the procedure of radical interpretation. Evidence is restricted to a degree such that one does not presuppose the notions to be illuminated. Now, if the theory is realized in the actual world, it must be realized in the *actual* interpretation of the verbal behavior of language users. Or actual interpreters must construct the theory. But Davidson rejects this idea and reiterates a sufficiency condition: “If an interpreter had explicit propositional knowledge of the theory, he would know the truth conditions [meanings] of utterances of the speaker” (Davidson, 1990, p. 312). So one must read Davidson’s claim of *is realized in the world* as *is realized in the world in principle*. Davidson’s insistence seems puzzling for some commentators; even leading them to ask “what is the interest of this project?” (Glock, 2004, p. 249). Consider the following statements:

- (1) An interpreter actually possesses the knowledge of the theory.
- (2) An interpreter in principle can possess the knowledge of the theory.

If one wants to claim (1), then one has to set a condition which attributes the necessity to relationship between truth and meaning theories. Davidson does not advocate (1) because a necessity claim will lead his overall programme to failure. Recall that Davidson commits a technical maneuver at the very beginning of his discussion about the form of the meaning theory which states that the truth theory may serve as a meaning theory. But such argumentation is valid only if he allows other predicates (in place of *is T*) to be used. Radical interpretation justifies (2) and the claim that a truth theory may serve as a meaning theory. The theory is not about how an interpreter interprets but what makes an interpreter an interpreter. If it is possible for the theory to be realized in the actual world through defined evidence, then the theory can be said to be successful in showing what could make thought and language interpretable (Davidson, 2004).

Consider the following quotation from Davidson while keeping in mind that the theory Davidson refers to is not the theory of a cognitive approach but of the philosophical approach.

In any case, claims about what would constitute a satisfactory theory are not ... claims about the propositional knowledge of an interpreter, nor are they claims about the details of the inner workings of some part of the brain. They are rather claims about what must be said to give a satisfactory description of the competence of the interpreter ... It does not add anything to this thesis to say that if the theory does correctly describe the competence of an interpreter, some mechanism in the interpreter must correspond to the theory. (Davidson, 2005, p. 96)

Chomsky (2000) objects that “from the standpoint of natural sciences” Davidson’s claim “is utterly wrongheaded.” According to Chomsky, with the attribution of some internal mechanism, “the natural scientist is then in a position to test the theory in terms of a wide array of evidence, for example, evidence drawn from other languages ... or evidence from pathology or the brain sciences or biochemistry” (pp. 56-57) but Davidson’s theory will lack all these features. Davidson replies what he “said was, and was intended as, a tautology: if a pill puts you to sleep, it adds nothing to say something about the pill had the power to put you to sleep” (Davidson, 2004, p. 132). Davidson discusses and defines what it is to be an interpreter—namely, to be capable of having an interpretive truth theory. Adding an internal mechanism will add nothing to such a definition. Davidson does not seem to reject any abstract, neural or physical internal component. Rather he expresses his anxiety by claiming that “it would be vastly interesting to know more about the nature of our linguistic abilities, and the mechanisms underlying them. Who would deny it?” (2004, p. 132). But in Davidson’s case there is no hint if such an underlying mechanism corresponds to semantic competence.

CHAPTER 5

A COGNITIVE APPROACH TO DAVIDSONIAN SEMANTIC THEORY

5.1) Objections against Davidsonian Semantics

It is often said that Davidson revises his programme on semantics to overcome *extensionality*, *determination* and *information problems*. These revisions are said to include (a) the introduction of radical interpretation as a *constitutive element* of a truth theory, (b) the presentation of the canonical proof procedure and (c) the assignment of a law-like nature to T-theorems (Glock 2004; Lepore&Ludwig 2005; Kölbel 2001). In previous chapters, it was shown that radical interpretation is a constitutive element of a truth theory for a natural language even in “Truth and Meaning” (1967/2001b).³¹ If this claim is true, then so-called revisions of (b) and (c) turn out to be just explicit statement of its consequences.

There are several points where Davidson (2001, p. xvi) modifies his semantic programme, but these interconnected modifications do not signal a departure from his initial position. It has been stated that Davidson's mistake is not assuming meaning through the construction of theory but through the grasping of truth. Later, Davidson identifies the details of radical interpretation to justify the construction of truth biconditionals. In his early papers, Davidson considers two cases. For case-1, where the meta-language contains the object language, the interpreter is assumed to have direct insight into the truth value of a truth biconditional. For case-2, where the meta-language does not contain the object language, the interpreter is assumed to have the ability to correlate the speaker's true sentences with his own. Davidson

³¹ “That empirical restrictions must be added to the formal restrictions if acceptable theories of truth are to include only those that would serve for interpretation was clear to me even when I wrote ‘Truth and Meaning’” (Davidson 1976/2001h, p. 173).

discards consideration of case-1, after explicit discussion of radical interpretation, by accepting Quine's dictum that "translation begins at home" (Quine, 1968, p. 198): "All understanding of the speech of another involves radical interpretation" (Davidson, 1973/2001e, p. 125). Davidson requires that the interpreter identify T-sentences from sentences which speakers hold true and use them to construct an axiomatic truth theory which entails *truth theorems*. The presentation of the details of radical interpretation also reveals the law-like nature of the truth theorems.

Now, in light of above considerations, we are in a position to evaluate objections raised against Davidson's semantic programme. The so-called problem of extensionality is that requiring of a truth theory for a natural language to be simply true does not ensure the entailment of *only* interpretive theorems of that theory. It has been said that a theory of this sort may entail the following sort of theorems and still be extensionally adequate:

- (1) "Snow is white" is true if and only if coal is black.
- (2) "Snow is white" is true if and only if it snow is white and LT.³²
- (3) "Tarski is a renate" is true iff Tarski is a cordate.³³

Davidson is aware of instances of type-1, even when he presents his bold proposal, and he claims that indexical elements will rule out such un-interpretive theorems (see section 2.3). Instances of type-2, when considered as a part of the extensionality problem, are results of the un-interpretive axioms: "The requirement that every T-form sentence be an extensionally adequate biconditional requires only that every axiom be an extensionally adequate biconditional, and not that it be interpretive as well" (Lepore & Ludwig, 2005, p. 102). Consider the following axiom of the theory:

(s)(n)(s satisfies(e, t) "n is white" iff ref(n, s, e, t) is white and LT)

³² LT is used to abbreviate *logical truth* such as *grass is green or grass is not green*.

³³ The term *renate* means creature with kidneys and the term *cordate* means that creature with a heart.

It is hard to see how such an axiom will be part of the empirically constructed theory. Axioms are taken to be theoretical constructions whose roots are hold true attitudes of speakers, and such an axiom is possible only in the case of the following empirical evidence.

Jack belongs to an English speech community, and Jack holds true "that is white" on Saturday at noon, and the object demonstrated by Jack on Saturday at noon is white and LT.

Recall that the interpreter is supposed to form particular sentences which the speaker holds true according to his own beliefs about the environment and the behavior of the speaker. But LT has nothing to do with the present circumstances of the environment. The same LT can be added to all other sentences which the speaker holds true. So if it is assumed that the interpreter is a rational being, he would not use instances of such pointless³⁴ evidence.

According to the determination problem, instances of type-2 may be derived from interpretive axioms. T-theorems are just material biconditionals. If it is possible to derive "*Snow is white*" is true if and only if it snow is white from axioms of the theory, it is also possible to derive "*Snow is white*" is true if and only if it snow is white and LT because both are logical consequences of theory³⁵ (Lepore & Ludwig, 2005, p. 109). Davidson (1973/2001e) proposes the "canonical proof" procedure (p. 138) to overcome such a problem, according to which a T-theorem is derived as a sequence of biconditionals where derivation is started from axioms and continues with biconditionals that are derived from them.³⁶ On that proposal, Davidson restricts

³⁴ The term *ceteris paribus* used at the formulation of the principle of charity will rule out the addition of LT. See footnote 25 at section 3.3.

³⁵ Consider the inference rule of substitution of equivalents (Larson & Segal, 1995, p. 35): For any formula α , β , $F(\alpha)$: $F(\beta)$ may be inferred from $F(\alpha)$ and $(\alpha \text{ iff } \beta)$.

³⁶ Lepore and Ludwig (2007) define a canonical proof as "...a proof characterized formally which is designed so as to draw only on the content of the axioms of the theory to prove T-form sentences in which we have eliminated the semantic

the rules of inference which can be used to derive T-theorems in accordance with radical interpretation.

A consideration of type-3 theorems may clarify Davidson's answer to the determination problem by giving his reason to appeal to the canonical proof procedure. Now, it is possible to substitute a co-extensive, but not synonymous, term without disturbing the truth value of the biconditional. Consider the following T-theorems:

“Tarski is a renate” is true iff Tarski is a renate.

“Tarski is a renate” is true iff Tarski is a cordate.

Davidson claims that T-theorems are law-like and must support counterfactuals:

If “Tarski is a renate” were false, then Tarski would not be a renate.

If “Tarski is a renate” were false, then Tarski would not be a cordate.

Although the former of the above sentences supports counterfactuals, the latter does not. So it is not law-like and it cannot be considered as an interpretive T-theorem. What is Davidson's point when he claims that T-theorems and axioms of the theory are law-like? T-theorems are derived from the axioms of the theory, and these axioms are determined from T-sentences. T-sentences are formed as a result of interaction with the environment and speakers in light of the principle of charity. Davidson's main point is that the T-theorems do not stand alone for the truth conditions of the object language sentences, but they are the products of the procedure of radical interpretation. The interpreter aims to reveal the compositional structure of the object language sentences. One needs an additional premise of the sort “all and only renates are cordates” (Lycan, 2006, p. 143) to derive the T-theorem of the form “*Tarski is a renate*” is true iff *Tarski is a cordate*. Also the interpreter must make additional inferences which have nothing to do with the compositional

vocabulary introduced by the theory” (p. 36). See also Davidson (1970/2001d, p. 61) which includes an early statement of the canonical proof procedure.

nature of object language sentences to form type-2 theorems. It is clear that there is no room for such a premise or inference rule in the procedure of radical interpretation.

The information problem was first formulated by J.A. Foster (1976) and strictly advocated by S. Soames (1984, 1999, 2003) later. According to the information problem, the knowledge of a truth theory does not suffice for interpreting utterances of the speaker. Soames (2003) claims Davidson's proposal is wrong because (a) "one might know that which is stated by a translational truth theory without knowing that the theory is translational" (p.305), and (b) "the fact that the truth theory is translational ensures that one of the T-sentences it generates ... is translational, but it does not tell us which T-sentence that is" (p. 308). According to the first of these objections, one might know that a T-theorem "*snow is white*" is true if and only if *snow is white* which is entailed by a truth theory. But one would nonetheless not use such knowledge to understand the object language sentence—to form a meaning theorem of the form "*snow is white*" means that *snow is white*—because nothing in the theory says that it is an interpretive theory. Davidson's response to this criticism is straightforward. He admits that a truth theory is sufficient for an interpreter to understand a speaker if and only if the interpreter has knowledge of the theory and he knows that the theory is of a special kind, i.e., that it satisfies "specified empirical and formal constraints" (1976/2001h, p. 172).³⁷ The formal constraint is the compositionality requirement and the empirical constraint is the construction (verification) of the theory by the procedure of radical interpretation. What is crucial is that Davidson does not develop a new argumentation to face such criticism, but he simply adds an explicit statement to his sufficiency condition.

³⁷ It has been stated that a truth theory which satisfies Convention-T—namely an interpretive one—would enable us to understand what other's utterances mean. So it may seem puzzling that Davidson does not just say that someone who knows an *interpretive* truth theory, and also knows that the T-theorems are entailed from such an *interpretive* truth theory, is in a position to understand utterances of others. In the empirical application of the theory, the interpreter cannot know that his theory is interpretive if the term *interpretive theory* refers to a theory which is constructed by the translation criterion. But the interpreter could know his theory is interpretive if the term *interpretive theory* refers to a theory which is constructed by means of radical translation.

Davidson claims that he “imagined the theory to be known by someone who had constructed it from evidence, and such a person could not fail to realize that his theory satisfied the constraints” (1976/2001h, p. 173). The addition of such a condition is not new for him. Also, canonical proofs were already available in Davidson’s programme to cope with objection (b).

5.2) A Cognitive Approach to Davidsonian Semantics

I want to offer a criterion, namely *the test of relevancy of truth*,³⁸ which can be used to evaluate the appropriateness of any cognitive approach to semantics. According to this test, a cognitive approach which proposes the replacement of a meaning theory with a truth theory (a) must give a legitimate reason for that proposal and, (b) must prohibit further replacements. The test simply asks the question of *why truth matters*. A cognitive approach to Davidson’s meaning theory must contain the following three claims:

1. The tacit knowledge of a truth theory constitutes the semantic competence of speakers of a language.
2. The truth theory is represented in the mind/brain of the language user.
3. The language acquisition of an infant includes acquisition of the knowledge of the truth theory.

According to the first of these conditions, a cognitive approach must identify the truth conditions of the sentences of the object language with their meanings. If a cognitive approach fails to pass the test of relevancy of truth, then it turns out that the notion of truth is irrelevant to the semantic model proposed and there is no reason to identify a truth theory with a meaning theory, or truth conditions of a sentence with meanings of sentence.

³⁸ See Soames (2003) for the roots of the test of relevancy of truth.

Knowledge of meaning (Larson & Segal, 1995) is the most systematic attempt to develop a cognitive approach to Davidson's meaning theory. The rest of this section will evaluate their work according to the test of relevancy of truth and discuss their answers to the objections which are presented in the previous section.

A cognitive approach will require the identification of truth conditions with the meanings of sentences to represent the semantic competence of a language user. But it is hard to see how such an identification is justified. Davidson prefers to use truth theories because he wants to formulate the meaning theory in fully extensional terms, but there is no reason why knowledge of an extensional theory represents our competence. The best starting point is to reveal the underlying motive:

With the development of Chomskyan linguistics and the rise of cognitivism in the latter part of the 20th century, a number of theorists adopted Davidson's insight that a T-theory (or some elaboration of one) was the right sort of theory for semantics, but deployed the insight within a more cognitivist perspective. (Segal, n.d, p. 18)

Segal's statement of "Davidson's insight" seems at least misleading. Davidson does not claim that the truth theory is the right sort of theory for semantics, if the words "right sort" indicates a necessity, because such a claim will rule out other possible theories. But another reading of the word "right" is available which refers to a scientific research programme. Larson and Segal (2005)³⁹ are well aware of this, so they make a controversial assumption and claim that the truth of this assumption depends on its explanatory power and verification of its consequences by empirical evidence: "*The T hypothesis* A speaker's knowledge of meaning for a language L is knowledge of a deductive system (i.e., a system of axioms and production rules) proving theorems of the form of (T) that are interpretive for sentences of L" (p. 33).⁴⁰ Does LS's assumption pass the first part of the test of relevancy of truth? Note that the test does not require an a priori condition but a

³⁹ The abbreviation *LS* will be used to refer Larson and Segal (1995) hereafter.

⁴⁰ The T-form theorems are T-theorems of Davidson's semantic theory.

loose one such as a *legitimate reason*. Although LS's assumption fails from a philosophical perspective, it passes the test from a scientific perspective.

The cognitive approach must cope with extensionality, determination and information problems, which are raised against its philosophical antecedent, to secure its internal consistency.

Suppose that as a matter of biological endowment (that is, of universal grammar), [A1] humans are designed to acquire T theories. These theories...are represented internally in the brain. In the course of learning a language, speakers fix axioms and production rules yielding T theorems as outputs. Suppose further that [A2] humans are designed to treat whatever T-theory they acquire as interpretive...Finally, [A3] suppose that events conspire to give speakers an interpretive T theory in the course of development. (Larson & Segal, 1995, p. 39)

A1 includes an answer to the determination problem. Recall that if the logic of the meta-language is not restricted, it is always possible for the truth theory to entail type-2 theorems from axioms of the theory. LS specify a system of production rules, in addition to axioms that are both part of the internalized truth theory, to derive T-theorems⁴¹ and these production rules prevent the truth theory being "overproductive" (p. 34). Davidson requires that a truth theory entail *all and only* interpretive theorems. Although type-2 theorems are ruled out by the introduction of production rules, entailment of type-1 theorems is still possible, as they are a result not of improper logic but of improper axioms. But for a moment, suppose that the axioms of the theory are fixed as interpretive. Now, only one condition of Davidson's requirement from an interpretive theory is secured. The theory will entail *only* interpretive theorems, but the other side of the problem questions whether the theory is "sufficiently productive" (p. 34). Davidson undertakes two tasks when he proposes to develop a truth theory for a natural language. One task is to overcome

⁴¹ Structural descriptions of object language sentences, which are posited by Larson and Segal, are composed in accordance with phrase structure which is used to indicate syntactic rules. As a result, axioms and production rules which are developed through the cognitive programme differs from the classical axioms and logic which are adopted by Davidson. But the main lines of argument and problems are same. See Larson and Segal (2005, chap.3).

the problem of the empirical application of the theory, and the other is to reveal the logical form of different structures which are present in natural language and to integrate it into the truth theory. Questions about sufficiently productive theory correspond to the latter of these tasks, which is the technical one. Although Davidson outlines a research programme to deal with natural language with a level of optimism, such optimism is not suitable for the cognitive approach. Davidson's programme is philosophical, and its justification does not crucially depend upon the technical part but the success of a cognitive approach directly rests upon it.

According to A3, it is assumed that people acquire interpretive truth theory during language acquisition. An interpretive theory is a theory whose axioms are interpretive. So, LS claim that "universal grammar" and "the very context in which natural language is acquired" (p. 40) ensures that the axioms of the theory are interpretive. Language acquisition proceeds by making conjectures about the axioms of the theory under the guidance of universal grammar, testing and revising them according to the available evidence. Consider the axiom $(s)(ref("dog", s) = dog)$. Universal grammar may be said to rule out un-interpretive axioms like $(s)(ref("dog", s) = dog \text{ and } carpet)$ when the parent points out a dog which sits on a carpet and says *that's a dog*. This axiom will be tested by checking whether it makes the meanings of further utterances sensible according to the context of communication. LS claim that, after a process of trial and error, the axiom will be fixed as interpretive. So the theory will not entail type-1 theorems.

LS admit that knowing that a truth theory is interpretive is necessary for treating theorems of the theory as giving meaning. Davidson can defend his position by claiming that the theorizer constructs the theory in accordance with presupposed constraints and will know that it satisfies them. But Davidson's position is valid only for the philosophical approach, and that is the reason why Davidson continuously reminds us that his theory does not describe the actual competence of an interpreter. A2 is LS's answer to the information problem. The information problem requires knowledge of an interpretive truth theory, and A2 just assumes it. LS suppose that T-theorems are produced in a semantic module and other modules (or performance systems) "that need information about the literal meanings of sentences receive T theorems as their inputs and then proceed under the assumption that the RHSs of

these T theorems give the meanings of what's mentioned on the left"⁴² (p. 40). While the axioms are ensured to be interpretive by A3, LS claims that a truth theory represents our semantic competence.

Does LS's formulation pass the second part of the test of the relevancy of truth? Note that in A2, all required is to treat a mentalese sentence as giving the literal meaning, and any other notion may do as well as the notion of truth. Consider the revised versions of LS's assumptions where *F* is false or falsity:

1. Humans are designed to acquire an F theory.
2. Humans are designed to treat any F theory they acquire as interpretive.
3. In the course of development, humans learn an F theory that is interpretive in fact.

According to these assumptions, an F theory can also be used as a meaning theory equally as well as a truth theory. It seems LS's approach fails to pass the test of relevancy of truth. A falsity theory may represent the semantic competence of a language user. The correctness of the T Hypothesis depends on its explanatory power and the verification of its consequences by empirical evidence but an F Hypothesis will have the same explanatory power, and empirical evidence which is used to verify T Hypothesis can be used to verify an F Hypothesis as well.

I want to claim that the information problem does not arise through the cognitive approach and A2 is redundant. Suppose that the object language of a truth theory is German and the meta-language is English and suppose further that this truth theory is given to an English speaker who does not know German. According to the information problem, the English speaker cannot use the theory to interpret the utterances of Germans because there is nothing in the theory which tells him to do this. But how can this be a problem for the cognitive approach? The speaker does not consciously decide to use a truth theory to interpret utterances of others, but he can interpret the utterances because he has the right sort of truth theory. The theory itself makes a competent speaker of a language a competent speaker. It can be objected

⁴² The term *RHS* is abbreviation for *right hand side*.

that there may be other truth theories which are not interpretive and the speaker may know them too. But only one theory will be known whose axioms and production rules are fixed during the acquisition. Of course, the speaker may know the T-theorem "*Snow is white*" is true iff *coal is black* but he can know it because he knows "*Snow is white*" is true iff *snow is white* and "*Coal is black*" is true iff *coal is black*, which are entailed by the truth theory which represents the speaker's competence. So I propose to omit A2 and revise the T Hypothesis as: A speaker's knowledge of meaning for a language L is knowledge of a deductive system (i.e., a system of axioms and production rules) proving theorems of the form of (T) for sentences of L. Note that after such revisions, the cognitive approach will pass the test of relevancy of truth. The evidence which can be used to verify the revised T Hypothesis will be different from an F Hypothesis. In LS's version, there is no room for the notion of truth in language acquisition because the axioms are said to be fixed according to the infant's grasp of the meanings of utterances. But, according to revised version, the grasp of truth is taken as evidence to fix axioms. So the revised version also suggests a research programme for examining the role of the grasp of truth in language acquisition.

CHAPTER 6

CONCLUSION

The purpose of this work is to discuss and evaluate the appropriateness of a cognitive approach to Davidson's meaning theory. It has been stated that two tasks must be completed to accomplish this. The first task is the clarification of Davidson's project on semantics and the second task is the exposition of what a cognitive approach is. Two separate approaches are identified by the completion of such tasks. Davidson's bold proposal is that a truth theory may serve as a meaning theory. The clarification revealed the nature of such proposal and Davidson's position is named as the philosophical approach. Chomsky's views on linguistics are presented and three conditions which a cognitive approach must contain are derived from it. It is shown that the cognitive approach must abandon the philosophical project together with radical interpretation. According to the cognitive approach, it must be possible to keep the core of Davidson's meaning theory distinct from radical interpretation. The hope is that a truth theory can be used to represent the semantic competence of a language user. But such hope turns out to be questionable in the light of the previous discussion about the nature of Davidson's bold proposal. It is considered that (a) the relationship between a truth theory and a meaning theory is justified by radical interpretation and (b) such a relationship is a loose one in the standards required by the cognitive approach. It is claimed that any cognitive approach has to identify truth conditions of sentences with their meanings. Larson and Segal's case is introduced as an example of the cognitive approach. The test of relevancy of truth is presented to evaluate the appropriateness of a cognitive approach to Davidsonian semantics and it is applied to Larson and Segal's case. It turns out that Larson and Segal's approach fails to pass the test and a revision in their main assumptions is proposed to avoid the failure. It is concluded that the cognitive approach can lead us to an empirical research programme with right assumptions which makes its claim testable.

BIBLIOGRAPHY

- Abler, W.L. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures*, 12, 1-13.
- Barsky, R.F. (1998). *Noam Chomsky: A life of dissent*. Cambridge. London: The MIT Press.
- Bradley, F.H. (1916). *Appearance and reality* (6th ed.). London: George Allen & Unwin Ltd.
- Frege, G. (1948). Sense and reference. *The Philosophical Review*, 57(3), 209-230.
- Frege, G. (1956). The Thought: A logical inquiry. *Mind, New Series*, 65(259), 289-311.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1959). [Review of the book *Verbal behavior*]. *Language*, 35(1), 26-58.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, M.I.T. Press.
- Chomsky, N. (1969). Some empirical assumptions in modern philosophy of language. In S. Morgenbesser, P. Suppes & M. White (eds.), *Philosophy, Science, and Method* (pp. 260-285). New York: St. Martin's Press.
- Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell Publisher Ltd.
- Chomsky, N. (1981). Knowledge of language: Its elements and origins. *The Philosophical Transactions of the Royal Society*, 295, 223-234.

- Chomsky, N. (1995). Language and nature. *Mind*, 104 (413), 1-61.
- Chomsky, N. (1997). Language and problems of knowledge. *Teorema*, 16(2), 5-33.
- Chomsky, N. (2000). *New horizons in the study of language and mind* (5th ed.). Cambridge: Cambridge University Press.
- Chomsky, N. (2002). Language and the brain. In A. Belletti & L. Rizzi (eds.), *On nature and language* (pp. 61-91). Cambridge: Cambridge University Press.
- Chomsky, N. (2006). *Language and mind*. Cambridge: Cambridge University Press.
- Davidson, D. (1990). The Structure and content of truth. *The Journal of Philosophy*, 87, 279-328
- Davidson, D. (1994). Radical interpretation interpreted. *Philosophical Perspectives*, 8, 121-128.
- Davidson, D. (1999a). Intellectual autobiography. In L. E. Hahn (ed.), *The philosophy of Donald Davidson* (pp. 3-70). Chicago: Open Court.
- Davidson, D. (1999b). Reply to Quine. In L. E. Hahn (ed.), *The philosophy of Donald Davidson* (pp. 80-85). Chicago: Open Court.
- Davidson, D. (2001). *Inquiries into truth and interpretation*. Oxford: Oxford University Press.
- Davidson, D. (2001a). Theories of meaning and learnable languages. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 3-15). Oxford: Oxford University Press. (Reprinted from *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, by Y. Bar-Hillel, Ed., 1966, Amsterdam: Nord-Holland Publishing Company.)
- Davidson, D. (2001b). Truth and meaning. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 17-36). Oxford: Oxford University Press. (Reprinted from *Synthèse*, 17, 304-323, 1967.)

- Davidson, D. (2001c). True to the facts. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 43-54). Oxford: Oxford University Press. (Reprinted from *Journal of Philosophy*, 66, 748-764, 1969.)
- Davidson, D. (2001d). Semantics for natural languages. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 55-64). Oxford: Oxford University Press. (Reprinted from *Linguaggi nella Società e nella Tecnica*, 1970, Milan:Comunità.)
- Davidson, D. (2001e). Radical interpretation. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 125-139). Oxford: Oxford University Press. (Reprinted from *Dialectica*, 27, 313-328, 1973.)
- Davidson, D. (2001f). Belief and the basis of meaning. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 141-153). Oxford: Oxford University Press. (Reprinted from *Synthese*, 27, 309-323, 1974.)
- Davidson, D. (2001g). On the very idea of a conceptual scheme. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 183-198). Oxford: Oxford University Press. (Reprinted from *Proceedings and Addresses of the American Philosophical Association*, 47, 1974.)
- Davidson, D. (2001h). Reply to Foster. In D. Davidson (Ed.), *Inquiries into truth and interpretation* (2nd ed., pp. 171-179). Oxford: Oxford University Press. (Reprinted from *Truth and Meaning: Essays in Semantics*, G. Evans & J. McDowell, Eds., 1976, Oxford University Press.)
- Davidson, D. (2004). Could there be a science of rationality?. In M. Cavell (Ed.), *Problems of rationality* (pp. 117-134). Oxford: Oxford University Press. (Reprinted from *International Journal of Philosophical Studies*, 3, 1-16, 1995.)
- Davidson, D. (2005). A nice derangement of epitaphs. In M. Cavell (Ed.), *Truth, language and history* (pp. 89-107). Oxford: Oxford University Press. (Reprinted from *The New Yorker*, 1995, April 4.)
- Davidson, D., & Harman, G. (eds.) (1972). *Semantics of natural language*. Dordrecht: Reidel.

- Davis, S., & Gillon, B. S. (2004). Introduction. In S. Davis, & B. S. Gillon (eds.), *Semantics: A reader* (pp. 3-100). Oxford: Oxford University Press.
- Etchemendy, J. (1988). Tarski on truth and logical consequence. *The Journal of Symbolic Logic*, 53(1), 51-79.
- Evnine, S. (1991). *Donald Davidson*. Stanford: Stanford University Press.
- Field, H. (1972). Tarski's theory of truth. *The Journal of Philosophy*, 69(13), 347-375.
- Fitch, W.T., & Hauser, M.D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97, 179-210.
- Foster, J. A. (1976). Meaning and truth theory. In G. Evans, & J. McDowell (eds.), *Truth and Meaning: Essays in Semantics* (pp. 1-32). Oxford: Clarendon Press.
- Gallistel, C.R. (1998). Behaviorism, methodological and scientific. In *Routledge encyclopedia of philosophy* (Vol.1, pp. 696-699) London: Routledge.
- Gardner, H. (1985). *The mind's new science: A history of cognitive revolution*. New York: Basic Books Inc.
- Glock, H. J. (2004). *Quine and Davidson on language, thought and reality* (2nd ed.). Cambridge: Cambridge University Press.
- Graham, G. (2007). Behaviorism. In Edward N. Zalta (ed.). *The Stanford encyclopedia of philosophy* (Fall 2007 Edition). Retrieved August 30, 2008 from <http://plato.stanford.edu/archives/fall2007/entries/behaviorism>
- Grice, H.P., & Strawson, P.F. (1956). In defense of a dogma. *The Philosophical Review*, 65(2), 141-158.
- Grice, H.P. (1957). Meaning. *The Philosophical Review*, 66(3), 377-388.

- Gupta, A. (1998). Tarski's definition of truth. In *Routledge encyclopedia of philosophy* (Vol.8, pp. 265-269) London: Routledge.
- Harris, R.A. (1993). *The linguistics wars*. Oxford: Oxford University Press.
- Hauser, M.D., & Chomsky, N., & Fitch, W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569-1579.
- Higginbotham, J. (1997). Reflections on semantics in generative grammar. *Lingua*, 100, 101-109.
- Horst, S.(2005). The computational theory of mind. In Edward N. Zalta (ed.). *The Stanford encyclopedia of philosophy (Fall 2005 Edition)*. Retrieved August 30, 2008 from <http://plato.stanford.edu/archives/fall2005/entries/computational-mind>
- Katz, J.J. (1980). Chomsky on meaning. *Language*, 56(1), 1-41.
- Kirkham, R.L. (1992). *Theories of truth: A critical introduction*. Cambridge, Mass.: M.I.T Press.
- Kölbel, M. (2001). Two dogmas of Davidsonian semantics. *The Journal of Philosophy*, 98(12), 613-635.
- Künne, W. (2005). *Conceptions of truth*. Oxford: Clarendon Press.
- Larson, R., & Segal, G. (1995). *Knowledge of meaning: An introduction to semantic theory*. Cambridge, Mass.: M.I.T. Press.
- Lepore, E., & Ludwig, K. (2003) Truth and meaning. In K. Ludwig (ed.), *Donald Davidson* (pp. 35-63). Cambridge: Cambridge University Press.
- Lepore, E., & Ludwig, K. (2005). *Donald Davidson: Meaning, truth, language and reality*. Oxford: Clarendon Press.

- Lepore, E., & Ludwig, K. (2007). *Donald Davidson's Truth-Theoretic Semantics*. Oxford: Clarendon Press.
- Lund, N. (2003). *Language and thought*. East Sussex: Routledge.
- Lycan, W.G. (2000). *Philosophy of language*. London, New York: Routledge.
- Miller, G. A. (2003). The cognitive revolution: A historical perspective. *TRENDS in Cognitive Sciences*, 7(3), 141-144.
- Neale, S. (2001). *Facing facts*. Oxford: Oxford University Press.
- Pelletier, F. J. (2001) Did Frege believe Frege's principle? *Journal of Logic, Language, and Information*, 10, 87-114.
- Pitt, D. (2008). Mental Representation. In Edward N. Zalta (ed.). *The Stanford encyclopedia of philosophy (Fall 2008 Edition)*. Retrieved August 30, 2008 from <http://plato.stanford.edu/archives/fall2008/entries/mental-representation>
- Platts, M.D. (1997) *Ways of Meaning: An introduction to a philosophy of language* (2nd. Ed.). Cambridge, Mass.: M.I.T. Press.
- Postal, P.M. (2004). *Skeptical linguistic essays*. Oxford: Oxford University Press.
- Quine, W.O. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60(1), 20-43.
- Quine, W.O. (1968). Ontological Relativity. *The Journal of Philosophy*, 65(7), 185-212.
- Quine, W.O. (1988). *Word and object* (15th ed.). Cambridge, Mass.: M.I.T. Press.
- Quine, W.O. (1992). *Pursuit of truth*. Cambridge, Mass.: M.I.T. Press.

- Quine, W.O. (1999). Where do we disagree? In L. E. Hahn (ed.), *The philosophy of Donald Davidson* (pp. 73-79). Chicago: Open Court.
- Searle, J.R. (1972). Chomsky's revolution in linguistics. *The New York Review of Books*, 18(12). Retrieved August 30, 2008 from <http://www.nybooks.com/articles/10142>
- Segal, G. (n.d.). *Truth and meaning*. Retrieved August 30, 2008 from King's College London, Department of Philosophy Web Site <http://www.kcl.ac.uk/kis/schools/hums/philosophy/staff/truth&meaning.pdf>
- Skinner, B.F. (1957). *Verbal Behavior*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
- Smart, J. J. C. (1969). Quine's philosophy of science. In D. Davidson, & J. Hintikka (eds.), *Words and objections* (pp. 3-13). New York: Humanities Press.
- Soames, S. (1984). What is a theory of truth? *The Journal of Philosophy*, 81(8), 411-429.
- Soames, S. (1999). *Understanding truth*. New York : Oxford University Press.
- Soames, S. (2003). *Philosophical Analysis in the Twentieth Century* (Vol.2). Princeton: Princeton University Press.
- Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3), 341-376.
- Tarski, A. (1956). The concept of truth in formalized languages (J. H. Woodger, Trans.). In *Logic, semantics, metamathematics* (pp. 152-278). Oxford: Clarendon Press. (Reprinted from *Studia Philosophica*, 1, 261-405, 1936.)