Reconciling Internalism and Externalism Through Use

WORAH, PALLAVI

Reconciling Internalism and Externalism Through Use

Thesis submitted for the Degree of

Doctor of Philosophy

By

Pallavi Worah

Department of Philosophy

Durham University

2014

Abstract

This thesis stems from two related problems within the Internalism-Externalism debate. First is the problem of intentionality: how is it that we are able to reach out to the world through language. The answer I believe is through use. The second problem relates to the possibility of scientific enquiry: whether or not it is possible to scientifically study the object of externalist inquiry, the use of language for instance. The kind of theory I try to point towards, both incorporates how we are able to reach out to the world through language; i.e. explains use, and also aims to do so scientifically. I arrive at this through various steps, the first of which is an interpretation of Wittgenstein's use of 'grammar' as both constitutive of a proposition and involving an element of use. Further through an analysis of Horwich, I attempt to see whether it is possible to systematize a use theory of meaning. I move on to externalist referential semantics, specifically to minimal semantics, as it appears to incorporate both internalist as well as externalist elements. For a theory that was able to scientifically explain content ascription, I then turn to Davidson. Davidson's decision theoretic account of what he called triangulation presents itself as a candidate to better explain how we ascribe a particular content to a particular word or sentence. This framework, I contend is exemplified in Pietroski et al.'s "Meaning of 'Most'" experiments. I argue that Horwich's use-properties could form the test statements for Pietroski-like experiments, which would yield a possible way to relate the use of a sentence to meaning and truth. And this result is only achieved with the important assumption that we are situated in a world which we share with others much like ourselves.

Table of Contents

List of Figures

List of Abbreviations

Works by Wittgenstein:

TLP	Tractatus Logico Philosophicus [TLP], Tr. D.F. Pears and B.F. McGuinness. International Library of Philosophy and Scientific Method. 1974
BBB	Blue and Brown Books, Ed. Rush Rhees, Oxford: Basil Blackwell. 1958a
PI	Philosophical Investigations [PI], Ed. G.E.M. Anscombe and R. Rhees, tr. G.E.M. Anscombe. Second edition. 1958b
Z	Zettel, G.E.M. Anscombe and G.H. von Wright, (Eds.) Oxford: Basil Blackwell. 1967
LA	*Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*. Ed. Cyril Barrett Oxford: Basil Blackwell.1966

Acknowledgements

I am extremely grateful to Wolfram Hinzen, every interaction with whom has been intellectually stimulating. His passion for his work was a constant inspiration. I thank him for his guidance, understanding, patience and simply for putting up with me.

I would also like to thank Nirmalangshu Mukherji for all his support, encouragement and motivation. Thanks goes to Sophie Gibb for meeting me, and commenting on a part of my thesis, to Jonathan Lowe for crucial feedback in the Mind Language Metaphysics meeting, and to Nick Zangwill and Ian Kidd for helpful advice during PhD review sessions, and to Mitchelle Sheehan for taking the time to read and comment on some of my work. I have also benefitted from interaction with Paul Horwich, Emma Borg, Paul Pietroski and John Collins. Each of their work has inspired me in various ways and has found a place in my arguments.

Being a part of the Language and Mind Research Group was an extremely enriching experience. Thank you Ulrich Reichard, David Kirkby, Andrew Woodard, Tom Hughes, Alex Malt and James Miller, for all the stimulating discussions and feedback.

I also thank Durham University, and the Department of Philosophy for the Durham Doctoral Fellowship and the Overseas Research Student Award, which gave me the opportunity to undertake this work. I would also like to thank participants at the MLM meetings for feedback at the early stage of this work.

And, of course, a huge thank you to my family. I am grateful to my parents and in-laws for helping with the baby at the final stages of my thesis. Being away from Durham and my Research group for an extended period, my husband was, for the most part, the one I went to for discussions and bouncing off ideas. He was not just an amazing listener, but also gave me a different perspective on this work, which I greatly value. Above all, his love and support is what got me through these years. Lastly, my little boy Taran has been the final motivation for me to complete this work.

Chapter 1: Introduction

An enquiry begins with a conundrum, a problem that seeks a solution or a simplification. The problem that puzzled me constantly and which led me to this work is: how do symbols acquire meaning. This is a central question in philosophy of language (among other fields of study as well) and one which is multifaceted, with many possible directions of inquiry. The intuition that has guided my thoughts regarding this broad and general question is that meaning must already exist in some form within the symbol itself (with relation to the symbol system) for it to mean anything. The idea is that we cannot give something meaning unless it already possesses an element of meaning or the possibility of it. There is certainly a sense in which everything possesses the possibility of meaning. To put it differently, we can decide to give meaning to anything: we can call a rock by a name and worship it and decide that trees symbolise growth and prosperity. Similarly the same painting could mean various things, a solemn sunset for some, and a furry cat for others. This way of ascribing meaning, as the term ascribing itself suggests, involves ascribing to the object or symbol (or even the system of symbols, an element (i.e. meaning) which is external to it. So the intuition suggests that we could only ascribe such external meaning if the system of symbols was internally significant in some way. We are thus seeking a theory of meaning which is capable of accounting for both the internal significance of language as well as the aboutness of language, or in other words, that language has external significance.

Why should one believe that the task is an important one to begin with? Isn't there a way to understand language as systematic, in a way that allows for ascription of meaning? A sentence is made up of significant parts, nouns, verbs, prepositions and connectives and these are combined in definite ways to form noun phrase, verb phrase, determiner phrase etc., and then each of these elements can systematically be ascribed a definite meaning. This way of understanding language could easily be aligned with Chomskian Biolinguistics (see section 1.4 below). Further, as Hinzen (2007) argues, we can even articulate the particular structures that enable reference. Reference isn't achieved by words alone[1], and we require structures of a certain degree of complexity in order to achieve reference. But what we seek is an explanation

[1] Contrary to traditional referential theories (e.g. Mill 1973) of meaning which state for instance that proper names refer to objects.

for these phenomena, although we are not yet certain of the form of this explanation, or whether an explanation is even available in this regard**.**My intuition is that we need to be able to reconcile the two aspects of meaning: internal content and external significance.

As an answer to the question: 'how does language reach out to the world?',I would intuitively say that this is achieved through the use of language. In Chapter 2, we seek motivation from Wittgenstein towards reconciling internal meaning (grammar) and external significance (use). In chapter 3, through an analysis of Horwich, we attempt to see whether it is possible to systematize a use theory of meaning. Then in chapter 4 we move on to externalist referential semantics, specifically to minimal semantics, and focus on Borg, as she appears at first glance to have a theory which incorporates both internalist as well as externalist elements. Finally, for a theory that is able to scientifically explain content ascription, we turn to Davidson in chapter 5. Davidson's decision theoretic account of what he calls triangulation presents itself as a candidate to better explain how we ascribe a particular content to a particular word or sentence. In this chapter, below, I hint at my underlying motivations and then clarify some terms under discussion: symbols as Chomskian internalism.

1.1 Motivations from Advaita Vedanta

'Tat tvam asi' or 'thou are that' the all-encompassing dictum of the culmination of the Vedas refers to the identification of the self with Brahman (or the Ultimate Reality). The idea is that there is no Other, there is just the singularity of Brahman, and all perceptible differences and differentiations are a part of that singularity. We perceive them as differences, and are able to distinguish one thing from another simply because we are unable to perceive that all that there is, is a manifestation of that Unity. All dichotomies, including what is internal and what is external, i.e. including the separation of the self and not-self is a manifestation of our ignorance of (the ultimate) Reality, or Brahman.

The above was a perspective from Indian Philosophy to suggest that dichotomies, especially of the self and that which is beyond the self; or of what is internal and what is external to the mind/brain; or what is physical and what is mental; are merely apparent. Chomsky (2000)[2] argued that the problem of mind-body dualism was no longer one that can be formulated, since the concept of the physical is no longer distinguishable from the mental.[3] He

[2] See also Mukherji 2010 for an overview of this argument.
[3] I am not entirely convinced that because the Cartesian concept of the physical body is no longer one that can be upheld, mind body dualism no longer applies. Crudely speaking we can easily distinguish between the physical

claimed that Descartes two-substance theory "was put to rest by Newton, who showed that one of the two substances does not exist: namely body." (Chomsky 2000:17)

The current work partakes of this debate in terms of the dichotomy between externalist and internalist approaches to the study of language. The reason we enter this debate is essentially in terms of how it informs the problem of intentionality in language. As an aside, the reason I became interested in this problem is that the Indian philosophical perspective, namely Advaita Vedanta, abolishes dualities and dichotomies, but it appears that language has inherent in itself the possibility of a dichotomy between the speaker and the non-speaker or that which is talked about. In other words the problem of intentionality in language, the aboutness of language, presupposes (though it need not state) the existence of an object or person that is talked about, apart from the speaker. It appears that the mind (or the speaker) reaches out to the world through language.[4] But language, for ancient Indian grammarians like Panini (c. 450 BC) and Bhartrhari (c. 450-500 AD) was itself an embodiment of the ultimate reality, Brahman. In the Indian philosophical tradition culminating in Vedant (end of the Vedas) 'Word' (or language) has throughout maintained an important epistemological significance. Through the various traditions, Nyaya, Vaisesika, Mimamsa, Yoga, Dvaita and Advaita Vedanta, the idea remains that knowledge is linguistic in an important sense, and whatever is knowable, is linguistic. Anything that is knowable can be described in language. For the Indian tradition culminating in the Vedanta school of thought, all that exists, Reality is one[5]. To describe it would be to admit of distinctions and differentiations within the singularity. Brahman, which is indescribable. But anything that is describable, is describable through words. So at the level of existence where we have not as yet realized Brahman, we can access Brahman through words (more specifically through the word 'Om').

To say that we only have language to describe reality with, it may appear that language in a sense constitutes reality or our modes of reference to it[6]. But language could also be understood as a way of getting at reality, which is not independently available in descriptive

---

body and consciousness, which could be equated with the mind, which is non-physical (given that we do not subscribe to a reductionist view: that everything can be reduced to the physical.) In other words, while physics may not give us a coherent notion of the physical, our intuitive understanding does. But I am not making this point or arguing for it here.

[4] Though for Chomsky the most important function of language is the expression of our own thoughts. However, even this expression, and language itself presupposes at least another person if not a community of speakers. (Here we are following Wittgenstein's *Philosophical Investigations* in assuming that there is no private language).

[5] Rather, it would be more appropriate to call it singular, as the term 'one' admits of a contrast with many; whereas Reality (or Brahman) admits of no such contradictions.

[6] This is the Un-Cartesian position discussed in section 5.6.

form. We can either experience it (reality or Brahman) as indescribable (i.e. beyond descriptions) or through word (Shabda Brahma). In other words, apart from experiencing, or realizing reality, the only way that we can access reality or truth is through language. Language does not create reality, but gives us a way to refer to it, access it in a sense. But then, how does it relate to the external world? How is it that we can refer and reach out to things in the world through language? This is the problem of this thesis. We should note here that the fact that we do in fact refer to this world through language is one that needs an explanation. And the explanation we seek is not just of the mechanisms involved; i.e. our explanation must not be limited to the particular sentential structure or formal element that is required in order for us to achieve reference, but should attempt to address how it (*any structure*) *can* do so.[7]

1.2 Symbols and Signs

Gallistel claims however, that the "essence of a symbol is that it denotes or refers to something other than itself." (Gallistel 1998: 7) We now briefly discuss this claim and the nature of a symbol (as opposed to a sign) which it involves.

This claim is made in the context of his idea that a symbol-processing system has two components: a set of symbols, and a set of operations that can be performed on the values of the symbols. Each symbol stands for something, "symbols for distance, symbols for positions... and so on. The symbols take on different values at different times, depending on the values of the variables that they denote." (1998: 7) We may note here that even in calling 'd' the symbol for distance for instance, it is we who are defining the usage and restricting the set of possible values it can take. Thus even here it appears that symbols do not refer to anything other than themselves, in the sense that they have this *external* meaning attached to them only through our ascription. The fact that symbols can be mapped on to a system of meaning or external significance is important in that perhaps it indicates that the symbol system must itself be (internally) significant in some sense, in order to systematically be given external significance.

A symbol system according to Harnad (1990) may be understood as a set of 'physical tokens' (marks on paper, holes on a tape) manipulated on the basis of a set of rules, which are also strings of such tokens. The combination of tokens is purely syntactic and based solely on

[7] This way of formulating the problem, at least limits the explanatory role that an analysis of the linguistic structure of the sentence can provide. Thus we have mentioned the negative, i.e. the kind of explanation that will not be enough, but we have not yet mentioned the positive, or the kind of explanation that we do in fact seek. We hope that this will become clear as we progress in this work.

the shape of the symbol. Further the entire system is ‘semantically interpretable’. It is only within the structured framework of these rules that these marks or holes have meaning.

Symbols also have to be distinguished from a sign or signal (that which indicates something other than itself). A large range of research (discussed for instance in Hauser et al 2002) indicates that non-human animals communicate with each other using signals of a variety of sorts, e.g. Vervet monkey alarm calls that indicate the presence of the particular predator. Even highly complex bird songs are not a system of musical phrases significant in themselves and then mapped compositionally on to some notion of meaning. They convey only such information as announcing the presence of a sexual partner or guarding of territory. The individual units of birdsongs do not combine to generate more complex meaningful units. (see e.g. Fitch 2005, Mukherji 2010). Thus it appears that one feature distinguishing a sign from a symbol is the presence of a natural relation between the sign and the meaning given to it, as opposed to Saussurean arbitrariness in the case of symbols, where there is no natural relation (rather an arbitrary one) between the symbolic object and the meaning given to it. Perhaps the way we here understand a symbol, restricts our application of the term to systems similar to human language itself – composed of letters and their combinations into words and sentences in unlimited ways, which are arbitrarily attached to a system of meanings.

This line of argument may be considered an assumption by those who believe that animal communication systems must also be considered language. The assumption however is reasonable considering crucial differences between human language and animal communication systems (Hauser et al 2002). For instance, although some animals can be taught symbols, their use of these acquired symbols is limited in at least two ways: firstly, they seem to lack generative capacity – the ability to combine the limited symbols in order to generate unlimited expressions – the property of discrete infinity. Secondly, and perhaps derivatively, the usage is limited and more or less directly linked with the world, i.e. they lack something like the creation of possible worlds. In other words, animal communication is not stimulus independent.

We thus have, intuitively speaking, the following conception of language as a starting point: a language is a symbol system, which consists of a set of physical tokens used and understood in accordance with a set of rules. Language is also stimulus independent and has infinite generative capacity.

A reasonable assumption we could add here is that this symbol system is grounded in human biology, and is not merely a social construct, or a kind of conditioned behavior like for instance what we find in Skinner (1957). There is a lot of research available to support this claim (Chomsky 1959, Goldin-Meadow and Mylander 1998, Crain and Pietroski 2001). But even if we do not accept this last assumption, we have a conception of language as a set of physical tokens (sounds or marks on paper), which have corresponding to them, some mental representations (required for understanding them).[8]

1.3 Internalism and Externalism

Here we are led to the debate between externalists and internalists. The debate between internalism and externalism arises with respect to varied domains of inquiry. In the philosophy of mind and language, the debate arises with respect to the individuation of linguistic meaning, mental content, and propositional attitudes (Goldberg 2007). Thus one could argue that linguistic meaning, or the mental content of our thoughts or our propositional attitudes are internalistically individuated, i.e. without reference to the environment of the speaker, and alternatively, that these are individuated with reference to the external environment of the speaker. For example, according to Egan, "Externalism holds that mental states are individuated by reference to features of the subject's environment or social context. Internalism denies that external properties have any individuative significance, holding that two subjects in the same internal (neural) states are in the same mental states, whatever their contexts" (Egan 2003: 90) Goldberg (2007) points out that traditionally, "psychological internalism is the thesis that psychological properties supervene on the intrinsic, non-relational features of the physical body of the subject instantiating those properties;… and… psychological externalism… [is] the denial of this supervenience thesis." And epistemic internalism is the thesis that "epistemic justification supervenes on what is reflectively accessible, whereas epistemic externalism would be the denial of this" (Goldberg 2007: 2)

Thus there are various dimensions of the debate between internalists and externalists. Simply put, we could understand this debate as follows. The philosophy of language attempts to explain the mind-language-world connection. If the course of the explanation focuses on the

[8] Note here that we have claimed that the mental representations are associated with sounds and marks on paper as opposed to being associated with things (objects, people, places) in the world. And the term 'mental representations' is used here loosely, simply to mean that in order to understand or make sense of these physical tokens, we need some brain-internal processing, and such processing will need to be over mental states which in some way represent the physical tokens.

language to mind connection, then the explanation could be termed an internalist one, whereas, if the explanation focuses on the language-world connection, then it is an externalist explanation. Note that we mentioned 'focus' as opposed to 'accept', leaving room for compatibility between externalists and internalists as neither need to deny the other (unlike the language used by Goldberg (2007) above. Given the questions raised in the previous section concerning intentionality, the aboutness of language, the particular kind of externalism that concerns us in this work is content externalism, or "the view that the intentional contents of certain of a subject's thoughts are to be individuated with essential reference to her environment" (Majors and Sawyer 2005:257). Framed in this manner, the obvious example that comes to mind is the Twin-Earth thought experiment by Putnam (1975). Here, we were asked to imagine another planet, Twin Earth, which is exactly like earth, except that the substance that is present in its lakes and rivers etc. is not H20 but XYZ. However, XYZ is in all other ways indistinguishable from H2O, and is used in all the same ways, and is even called water. Earthlings and Twin-Earthlings, when they utter the term water, they think of the substance in the oceans and lakes, which quenches thirst etc., and both water and Twin-water are identical in that they both have these properties. But Putnam argues that given that the substance is not H2O, it is not the same as our water (although it fills the oceans and lakes, quenches thirst, freezes into ice and snow etc.). This is because the term 'water', the word as well as the thought, derives its intentional content from the environment; thus leading to semantic or content externalism. So although water and Twin-water may be indistinguishable in thought, they are different because the elements that form them are different, i.e. as they are found in the natural environment, they occur differently constituted. Thus, given that the intentional contents of our thoughts (especially about natural kinds) are individuated with reference to the environment, water and Twin-water cannot be the same. Oscar and Twin Oscar uttering the term water may be in identical psychological states, but their referents are different. Thus referents are not determined by the speakers psychological state, and "'*meanings' just ain't in the head!"* (Putnam 1973: 704). But Millikan argues that this is an oversimplification and "Externalism should not be so obviously true": "If we explain the externalist idea in this crude way, however, it becomes hard to see how anyone could deny it. If the question were, merely, how are the referents or extensions of thoughts determined, it seems patently obvious that nothing inside someone's head could, by itself, determine that anything in particular existed outside the head". (Millikan 2004: 229) If only mind internal states were to determine meaning, there would be no objective way of identifying truths from falsehoods. For (the later) Putnam, however, this is an important starting point: that an individual in isolation cannot grasp any

concept, and needs to interact with other speakers and the world (Putnam 1996). Putnam in a recent article explains the development of his externalist position as one starting from an 'internalism', which he later came to believe was false. The internalist position he puts himself in opposition to (and which he claimed was one he held in early work) was that "the meaning of a word could be given by 'semantic rules'" (2013: 196) which are "implicitly known by each speaker" (2013: 192). Arriving at his version of externalism, Putnam (2013) claims, involved the realisation that "what fixes the meanings of a speaker's words is not just the state of her brain; the reference of our terms is generally fixed by two things that classical philosophy of language either ignores or mentions only as an afterthought: *other people and the world.*" (2013: 197) There is in fact nothing controversial about this statement. Internalists do not claim that the entirety of meaning can be known by implicit semantic rules, or states of the brain. Also, fixing reference is a matter of interacting with the world as well as others in the community of speakers; not many would deny that. Here Putnam himself notices the shift from meaning to reference, but claims that the two are inseparably connected, at least with regard to natural kind terms: "For natural kind words, e.g., names of biological species and names of substances, difference in reference *counts as* difference in meaning." (Putnam 2013: 198 italics in original) This identification of meaning and reference is the point of contention. We will take this up again in chapter 4 when we discuss Minimal Semantics, an externalist position that seeks to identify what is minimally required for semantics.

The kind of internalism that Putnam seems to be opposing is for instance that described by Newman (2005): "Internalism about mental content holds that microphysical duplicates must be mental duplicates full stop. Anyone particle-for-particle indiscernible from someone who believes that Arisotle is wise, for instance, must share that same belief."[9] This, however, is not the kind of internalism Chomsky holds. As Lohndal and Narita (2009) aptly point out: "the core claim of internalists who, like Chomsky, seek a naturalistic theory of language is that the proper object of a serious linguistic science should be organism-internal aspects of human language (namely I-language). Internalists never deny that there are phenomena broadly related to language (in particular to language use) that are beyond the narrow confines of the architecture of the human mental grammar (I-language). Such phenomena would surely include prescriptive pressures from the linguistic community, speakers' intentionality for

[9] Newman goes on describe externalism as follows: "Externalism instead contends that many perfectly ordinary propositional attitudes can be had only in certain sorts of physical, socio-linguistic, or historical context. To have a belief about Aristotle, for instance, a person must have been causally impacted in the right way by Aristotle himself (e.g., by hearing about him, or reading some of his works). (Newman 2005: 153)

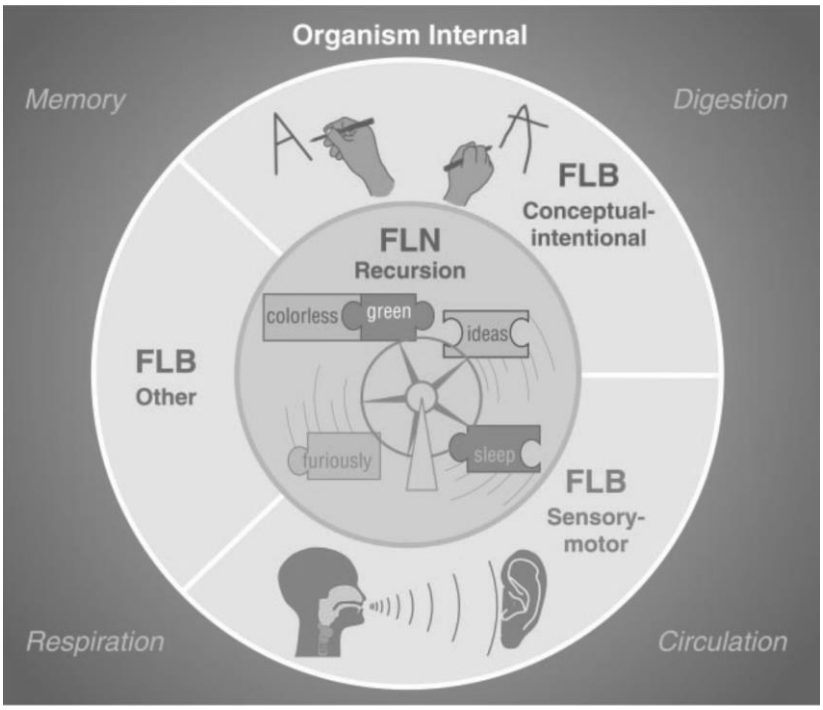
communicative success, and all sorts of other E-language phenomena […] What internalists doubt is rather the feasibility and/or legitimacy of providing a serious science of any mind-external phenomena such as these." (Lohndal and Narita 2009: 326)

Thus the debate between externalists and internalists is a matter of whether it is feasible or legitimate to study mind external phenomena in a scientific way. It is not a debate about whether or not the external environment plays a role in giving meaning to words, as neither externalists nor internalists would deny this. That the external environment is essential, is not debated. But rather, what is at issue is whether the role of the external environment is amenable to a scientific study and should be a part of semantic theory. (Chomskian) Internalists contend that what can be studied in a scientific way is the mind internal mechanism, i.e. the faculty of language, whereas externalists argue that a semantic theory must account for mind external phenomena like truth and reference. Such phenomena can only be understood if we consider mind-external elements like the context of use etc. and internalists claim that such elements cannot be studied in a scientific way and therefore must be left out of an exposition of a theory of meaning. We will briefly outline the Chomskian internalist framework below, before we proceed to our enquiry.

1.4 The Chomskian Internalist Framework

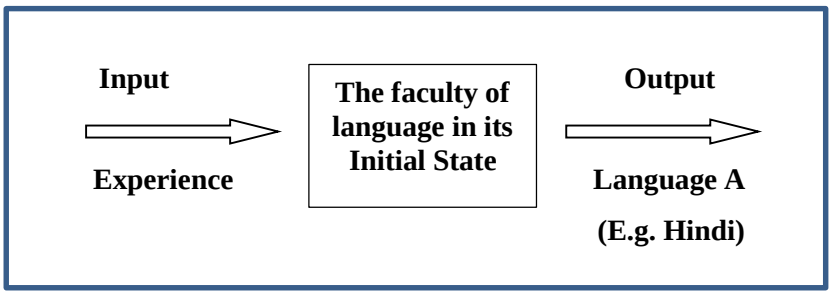
Hinzen (2006) defines internalism as "an explanatory strategy that makes the internal structure and constitution of the organism a basis for the investigation of its external function and the ways in which it is embedded in an environment." (Hinzen 2006: 139) The internal structure relevant to language is, for Chomsky, the language organ or the faculty of language. According to the framework of Hauser et al., 2002 (Shown in Figure 1), the 'Faculty of Language in the broad sense (FLB)' includes an internal computational system (FLN) specific to language and humans; combined with at least two other organism internal systems which Chomsky calls the 'sensory motor' (SM) and the 'conceptual-intentional' (CI) system, it yields the basic architecture of language.

Figure 1: Faculty of language, Reproduced from Hauser, Chomsky, Fitch 2002



What Chomsky calls the 'language organ' or the faculty of language is like other organs in its basic character; the language 'organ' is viewed as an expression of three factors: genes, experience, and 'third factor' conditions, i.e. principles not specific to the faculty of language (Chomsky, 2005). Each language is the result of the environment and the course of experience acting on a genetically determined 'initial state' of the language faculty or organ. See Figure 2 below.

Figure 2: Model of the Initial State of the Faculty of Language



With respect to Figure 2, we can study the input as well as the output, i.e. we can study the course of experience and the properties of languages acquired.[10] The Chomskian internalist approach is essentially concerned with the faculty of language: its initial state and the states it assumes. If, say, Peter's language organ is in state L, then L may be thought of as Peter's language. Here language may be understood as "the way we speak and understand". The theory of the initial state of the Faculty of Language is called 'universal grammar'. "Peter's language determines an infinite array of expressions, each with its sound and meaning... his language "generates" the expressions of his language. The theory of his language is therefore called generative grammar." (Chomsky 2000) Chomsky argues further that a genuine theory of language must satisfy at least two conditions:

1. Descriptive adequacy
2. Explanatory adequacy

The former holds for a grammar of a particular language. A grammar satisfies this condition if it "gives a full and accurate account of the properties of the language" (Chomsky 1997/2004: 383) The latter holds for the general theory of language, universal grammar. To satisfy this condition, "universal grammar must show that each particular language is a specific instantiation of the uniform initial state, derived from it under the 'boundary conditions' set by experience." (Chomsky 1997/2004: 383). There appears to be a serious tension between the requirements of the two conditions. The search for descriptive adequacy leads to complexity and variety of rule systems, while explanatory adequacy requires that language structures be largely simple and invariant to enable a human child to acquire any of the thousands of languages and dialects rapidly in the impoverished conditions of primary linguistic experience.

The condition of explanatory adequacy required that the construction of grammars be based on the impoverished conditions of language acquisition. Further, it is the burden of explanatory adequacy to be able to explain the acquisition of any language by any normal child, as children can in fact acquire any language and are not born with a genetic disposition for specific languages (an Indian born and brought up in China would easily acquire Chinese and may have trouble learning Hindi for instance after the stage of maturation). Thus the core

[10] We may note here that the input to the language organ -- what are sometimes called 'primary linguistic data' -- are themselves linguistic experiences and thereby involve the output itself as an input.

recursive system, FLN, must be available across the species and it must also be able to account for the acquisition of any language. Further, as Chomsky observes, to be usable, the expressions of the language faculty (FLN) have to be legible by the outside systems (the sensory-motor and the conceptual-intentional systems).

So the picture (considerably oversimplified) is as follows: the learning of the lexicon triggers off the internal computational system which consists of 1) the single operation 'Merge' which simply puts together two elements α and β form the lexicon to form $\{\alpha, \beta\}$, and 2) certain principles of computational efficiency (PCEs)[11] , for example, all elements in a structure must be interpretable, the shorter of two converging derivations is valid etc. The output is then interpreted at the two interfaces: Phonetic Form (PF) and Logical Form (LF) or the phonetic and semantic interfaces (corresponding to the 'sound' and the 'meaning' part respectively).

Linguistics for Chomsky, is the study of linguistic competence (Chomsky 1980). The latter is to be distinguished from 'pragmatic competence' (1980: 59) or performance. Linguistic competence refers to the faculty of language described above, whereas performance involves all aspects related to the use of language. This distinction between competence and performance develops (Chomsky 1986) into the distinction between I-language and E-language. Chomsky maintains that linguistic theory is a study of "what we may call 'I-language', where 'I' is to suggest 'intentional' and 'internalized'. The I-language is what the grammar purports to describe: a system represented in the mind/brain… and is in this sense internalized; a system that is intentional in that it may be regarded as a specific function considered in intension – that is a specific characterization of a function – which assigns a status to a vast range of physical events" (Chomsky 1997: 10) E-language on the other hand, is what philosophers have traditionally understood 'language' to refer to: "'E' is intended to suggest extensional' and 'externalized'. The definition is 'extensional' in that it takes language to be a set of objects of some kind, and it is 'externalized' in the sense that language, so defined, is external to the mind/brain" (Chomsky 1997: 7))

Thus we note language is grounded in human biology, it is stimulus independent, and a uniquely human element. With such a general conception of language, we will now begin our

[11] These PCE's belong to the 'third factor' (Chomsky: 2005) mentioned above. These include "(a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including principles of efficient computation, which would be expected to be of particular significance for computational systems such as language." (Chomsky 2005: 6)

inquiry with the question whether it is possible to reconcile internalist and externalist concerns within a theory of meaning. I take up this question with respect to the work of Witgenstein, who most notably changed his own position to incorporate a use theory of meaning.

Chapter 2: Wittgensteinian Motivations

2.1 Introduction

As noted in the introduction, I began this project with an initial (and long felt) intuition: that meaning must already exist within a symbol (with respect to a symbol system) for it to mean anything. Within the domain of language (as a symbol system), one way to understand this could be simply that there must be a meaningful structure for symbols to be considered a language and have meaning. The 'meaningful structure' could in turn be understood as a (Chomskian) Universal Grammar, that is a part of the biological faculty of language, and that which triggered at certain critical periods, enables any child to learn any (and multiple) language(s). This, however, would not account for the aboutness inherent in language. For Chomsky (as noted in the previous chapter) this aboutness is a product of the use of language, as distinguished from the grammar (syntax/semantics, given that he argues against the distinction between the two). Wittgenstein on the other hand captures both these aspects, although they appear disjointed and the received views of his writings suggest that they form parts of different theories[12] which he defended at different phases of his philosophical career. In what follows, we pick some elements of Wittgenstein's writings, with the aim to see whether an interpretation of them would reveal or lead to a possible connection between these aspects: internal significance, and the inherent aboutness of language. Bear in mind that my aim isnot to form a coherent and accurate interpretation of his work, but only to see whether an interpretation of some elements of his work would help us with our original intuitions pointed out in the Introduction.

2.2 Centrality of Internal Structure

The idea which I believe is most relevant to this work, is TLP §3.14: "What *constitutes* a propositional sign is that in it its elements (the words) stand in a determinate relation to one another." (Emphasis added)[13]. So here it is constitutive of the sentence or proposition that the words stand in a determinate relation to one another. In other words what makes a sentence a

[12] If we can call his writings theories at all.
[13] Compare TLP 2.14: "What constitutes a picture is that its elements are related to one another in a determinate way."

proposition, is that the words stand in such a determinate relation. What, however, would characterize something as a sentence or a proposition rather than an arbitrary set of words? We find the following possible answer in the *Tractatus*: "a proposition is not a blend of words" (TLP §3.141) Rather a proposition is articulate. The notion of articulation itself appears to be rather ambiguous. Is it merely the pronunciation – the verbal expression? If that were the case, then we would be forced to include even word salads (e.g.: Home the Ram get ran book), which we can certainly pronounce or verbally *articulate* as propositions. However, what is meant here is that all the elements in a proposition stand in a determinate relation to each other (TLP §3.14). This could be understood simply as that a proposition is necessarily a grammatical construction.

2.2.1 Determinate Relation and Grammar

The following picture emerges from the TLP. A proposition is a *picture* of reality as opposed to an *external description* of it. One need not look outside the picture to see the reality that it depicts. The picture itself depicts the reality.[14] However, he argues that a picture represents reality from *outside* and cannot step beyond representational form.[15] To me, this suggests that through depiction an external relation is internalized. Let's see further how he argues for this position.

A picture, for Wittgenstein, has in common with reality its *logico-pictorial form* (TLP § 2.2[16]), i.e. the possibility that things are related to each other in the same way as the elements of the picture (cf TLP §2.203[17]). In a picture, one can recognize the shapes in themselves, and also their representation, without acquaintance with the objects or situation represented. We can, for instance, recognize the picture of a human face, without actually knowing the referent of the picture. In fact, we do not normally assume that objects in a picture actually represent a *reality*, and we understand what they depict, independently of any claim to correspondence with reality. Pictures are in this sense self-contained. They do not require knowledge of anything external in order for us to understand them. What they depict, is a possible reality that is contained within the structure of the picture itself.

---

[14] "A proposition is a picture of reality: for if I understand a proposition, I know the situation that it represents. And I understand the proposition without having had its sense explained to me."(TLP § 4.021)

[15] "A picture represents its subject from a position outside it. (Its standpoint is its representational form.) That is why a picture represents its subject correctly or incorrectly." (TLP § 2.173)
"A picture cannot, however, place itself outside its representational form." (TLP § 2.174)

[16] "A picture has logico-pictorial form in common with what it depicts" (TLP § 2.2)

[17] " A picture contains the possibility of the situation that it represents." (TLP §2.203)

For the Tractarian Wittgenstein, *thought* was both, "A logical picture of facts" (TLP §3) and "a proposition with a sense" (TLP §4). Thus, for Wittgenstein, a logical picture of facts was a thought, or a proposition with a sense. A proposition, which he argued is a picture, "determines a place in logical space." (TLP §3.4), and a place "is a possibility: something can exist in it." (TLP §3.411).. 'thought', he had argued, was fundamentally logical, "Thought can never be of anything illogical, since, if it were, we should have to think illogically." (TLP §3.03) And the sense of a proposition was its agreement or disagreement with reality. Given all these elements, let's consider the following nature of thought[18]:

Thought is (a proposition) with (a sense). [given TLP §4]
Thought is (that which determines (a place) in logical space) with (agreement or disagreement with reality) [given TLP §3.4 and §4.2[19]]
Thought is (that which determines (a possibility of existence) in logical space with (agreement or disagreement with reality) [given TLP §4, §3.4, §3.411 and §4.2]

Thus it appears that in 'thought' he combined the *logical* possibility with the *real* possibility (of agreement or disagreement with reality). The logically possible was also, for him, the possibly real. If these two possibilities were not interchangeable for him, perhaps we would not have a problem accepting this. To understand this, it would require a much more detailed treatment than what I can currently offer.

However, for our present purpose, I would like to point out that the relationship between the proposition, i.e. the sensuous expression of thought (TLP §3.1), and reality is established through the notion of depiction. This is evident through Wittgenstein's assertion that "The possibility of all imagery, of all our pictorial modes of expression, is contained in the *logic of depiction*." (TLP §4.015, Emphasis added.) A proposition provides this depiction through a definite relation between its various elements. Through this internal relation between the elements of a proposition, it acquires the relation to the world – the sense of the proposition. Here the question arises: what is a definite relation? Does it become a definite relation through it being a picture or is the relation definite and therefore it is a picture.

Is the former even possible? Can a string of words be a picture without having a definite relation between the elements? Now if we understand this question as asking whether the

[18] Here brackets, (), have been added in order to delineate all that is included in 'thought'.
[19] TLP 4.2: "The sense of a proposition is its agreement and disagreement with the possibilities of the existence and non-existence of the atomic facts."

particular words and the word order is important, then perhaps this is not strictly so – as we can have different words and word order bringing forth the same description. For instance 'John gave the pen to Mary' and 'The pen was given to Mary by John' both appear to be interchangeable forms. Interestingly, this is not necessarily the case. It has been argued that there is a slight difference in meaning even between the active and passive form of the same sentence. Consider for instance the following forms:

(1) Precisely two arrows hit the target

(2) The target was hit by precisely two arrows

Apart from the different subjects in the two sentences, the first sentence claims that among many arrows that were shot, only two hit the target. The second, however, suggests that the number of arrows that hit the target was two. Thus, the two sentences, apart from having different subjects, are saying very different things.[20] Similarly, compare 'John rolled Mary the ball' and 'John rolled the ball to Mary': the latter, but not the former, implies that the ball arrived at Mary.

A similar argument can be found in the *Investigations* (PI §531): "We speak of understanding a sentence in the sense in which it can be replaced by another which says the same; but also in the sense in which it cannot be replaced by any other... In the one case the thought in the sentence is something common to different sentences; in the other, something that is expressed only by these words *in these positions*." (emphasis added) But in the second case how would one explain the expression? Here Wittgenstein asks, how does one *lead* someone to the comprehension of a poem or a theme? You can tell someone to look at a painting like this, or to listen to a piece of music in a particular way.

[20] However, we may mention at this point that synonymy is extremely rare in human languages. This, it has been argued (Carstairs-McCarthy, 1999), may be due to the cognitive principle of 'synonymy avoidance', which "demands that all contrasts in form imply a contrast in meaning."(Hinzen 2006: 136). Intuitively we can see the need for such a principle, as synonyms would make half of language redundant, and this would be implausible if language has an optimal (least costly) design. Thus each linguistic device should be viewed as playing a unique role in sound-meaning correlation**.** However, the idea that language has an optimal design is, of course, yet, itself an assumption. Chomsky argues in various places that though language may not be optimally designed for use (given ambiguities etc), it may be optimally designed 'from the perspective of internal structure' (Chomsky 1999), to interact with other mind internal systems, which are external to the language faculty. He argues that 'To be usable at all, a language must be "legible" at the interface: the expressions it generates must consist of properties that can be interpreted by these external systems' (Chomsky 2000.).

But there are languages where the word order isn't important. Mohawk, for instance is a language (still spoken by a few Native Americans in Quebec, Ontario and New York) which allows completely random word order. Consider the following example, pointed out by Baker (2001):

Sak	ranuhwe's	ne	atya'tawi	(Mohawk grammatical sentence)
Sak	likes	the	dress	(literal English translation)

The following however, are also grammatical formulations in Mohawk:

Ranuhwe's ne atya'tawi ne Sak
Ranuhwe's ne Sak ne atya'tawi
Sak atya'tawi Ranuhwe's
Atya'tawi Sak Ranuhwe's
Atya'tawi Ranuhwe's ne Sak

It may therefore be possible to say the same thing in different words[21], or with varying order of words; however, in the case of any of these particular instances there must nonetheless be a definite relation between the different elements of the proposition. Here perhaps our sense of a definite relation is merely that it must in fact be a proposition as opposed to a mere string of words e.g.: 'A number of books are lying on the table' instead of 'books the of lying table a on number are'. But we just saw, in the Mohawk example, that the word order may not be important. However, even in Mohawk, each of these formulations are recognized as grammatical sentences and the words do of course stand in grammatical relations. More importantly, the relation between the particular words in any one of these *particular* instances is, in each case, definite, from a grammatical standpoint.

We have argued that the 'definite relation', or more simply, the syntax, is what makes a sentence a sentence. So even within the picture theory, the notion of 'definite relation' is in a sense prior to that of a picture: the proposition becomes a picture through having a definite relation between its elements. Thus, here the possibility of depicting reality, the connection between the proposition and the world, is itself parasitic on the notion of a definite relation (or grammar).

[21] We argued above that this too isn't obvious, as the active and passive sentences are not identical, by virtue of their structure alone.

In the *Investigations*, having given up the picture theory, Wittgenstein argues accordingly that "When I think in language, there aren't 'meanings' going through my mind in addition to the verbal expressions: the language is itself the vehicle of thought." (PI § 329) And how is it that language becomes the vehicle of thought? It is, he answers, through the grammar of our language: "Like everything metaphysical, the harmony between thought and reality is to be found in the grammar of the language" (Z §55). For instance, he asks us to compare the following:

"This sentence makes sense." – "What sense?"
"This set of words makes a sentence." – "What sentence?" (PI §502)

The argument is the following: when someone utters the statement 'This sentence makes sense' and as a response to the utterance we ask the question 'What sense?', it is comparable to a situation where someone utters the statement 'This set of words makes a sentence', and we ask 'What sentence?'. The answer to the question 'What sentence?' can only be the set of words themselves. If the two situations are comparable, then the answer to the question 'What sense?' should also consist in presenting the particular sentence itself. Consistent with his position in the TLP (§3.14), here Wittgenstein claims that the sentence is the combination of words in that particular order, and that combination of words itself and nothing else, is the sentence. Similarly, the sense of the sentence lies in the sentence itself. We find for instance, the following remark in the Blue and Brown Book "the content of the sentence is in the sentence" (BBB 167).[22] Interpreted in the same vein as above, this statement claims that it is the grammar of the sentence itself which provides content to the sentence. Therefore, the sentence gets its content from within the sentence itself and there is nothing external to the sentence which is content or significance.

Thus it appears that Wittgenstein is arguing that the sense of the sentence is in the sentence itself. So any sentence or grammatical combination of words would be admitted as

[22] This relation of content to the sentence is comparable with the *Vaisesika* category of *Samavaaya*. It is defined as "that relation of things mutually involved or associated in nature and bearing to one another the relation of the contained and the container - which is the source of intuition in the form of "(It is) here""(Nandalal Sinha 1911: 267) e.g. "There is cloth here in the threads". It is argued by the Vaisesikas that "the non-intermixture" of the cloth and the threads "follows from the very rule of the container and the contained" (Nandalal Sinha 1911: 267) This means that the cloth is not the same as any bunch of threads and any bunch of threads do not make a cloth. Rather cloth results from a *particular combination* of threads, and only when the threads combine in such a way can we say that 'there is cloth here in the threads'. Thus cloth and thread are not identical; however, in 'this cloth here' the threads are inseparably related to the cloth. This inseparable relation (*samavaaya*) exists in this cloth and without this particular relation it would not be a cloth or it would be a different cloth.

significant. However, he points out that it appears that language allows for senseless combinations of words and one wants to say that a significant sentence is one which one cannot merely utter, but also think.

There can however be sentences that look like sentences that we understand and yet they don't yield any sense, for instance: "The number of pages in my book is equal to a root of the equation $x^2 + 2x - 3 = 0$." He adds that this "throws light on the concepts 'understanding' and 'meaning'." (PI § 513)The above is important as it seems to suggest that for Wittgenstein the notion of 'understanding' has a wider usage perhaps than meaning (Pietroski 2005). At first sight we may think we understand the sentence, but can we be mistaken? What is for instance the difference between the above sentence and a sentence like 'Green ideas sleeping furiously'? Initially it seems that we can make sense of the former but not of the latter. But on a closer look we may say the former doesn't mean anything because it is illogical while the latter does mean something. The former seems to be devoid of meaning because it is not recognized as possible.

However we do *somehow* also recognize what it would be, for it to be possible. We can imagine an instance where the apparently nonsensical equation is used. Let's consider the following situation: a child, perhaps 12-13 years old, has just learnt this equation. She, like many other children of that age, has a habit of tearing off the pages of her notebook, starting from the middle. After many such occasions, the book is now left with only the front and last pages. To ensure that both pages don't come off, she carefully tears off just a little less than half a page. Now her father hears the noise and starts screaming at her, and asks how many pages are left in her book now – upon which the girl replies with the equation as she doesn't like to lie but wants her father to stop screaming, and knows he'll be embarrassed asking her how to solve it.

The purpose of the example is to point out that even if a particular formulation appears to be impossible (although grammatical), it may well serve a purpose in language. And here this purpose is not limited to the *sound* of the equation: the complicated appearance or sound can serve to confuse anyone, and thus may be accepted on face value. But in that case, any other equation of the same kind would serve the same purpose in the given situation. Our point is however, deeper. The use of this equation was meaningful in this context, not just for its phonetic qualities, but for the structure itself, i.e. it did provide an answer to the question posed. It may further be noted that even contradictions are used as (valid) lines in a formal proof.

So we are here arguing that it is the sentential form, the grammar, that makes something a significant proposition – "what will be called (logically) possible and what not"? (PI §520). But, Wittgenstein importantly argues that, "It is not every sentence like formation that we know how to do something with, not every technique has an application in our life; and when we are tempted in philosophy to count some quite useless thing as a proposition, that is often because we have not considered its application sufficiently." (PI §520) Thus grammar is not the sole determiner in what counts as a significant proposition and what does not. Rather, it is that which has an application in our life, that which we can use, which counts as significant. Thus *use* is, in a sense, the determiner of what is significant.

The above line of thought is also important to keep in mind when we attempt to understand Wittgenstein's use of certain terms such as *grammar* and *determinate relation*. Although Wittgenstein himself argued for the centrality of grammar, the term *grammar*, as I understand it, has been variously used by Wittgenstein[23] to refer to the ways in which we use a word or sentence, the rules of such use, the entirety of our language, or even particular language games. Thus interpreting Wittgenstein's use of *grammar* as simply syntax will not be accurate, and consequently *determinate relation* must also not be confined to syntax. It would perhaps be more accurate to understand the determinate relation of the *Tractatus* as a grammatical relation, in the broader sense of grammar, denoting the various ways in which Wittgenstein uses the term: use, rules of use, language games as well as the entirety of language. This notion of grammar, however, already involves an element of use, as noted above.

However, this may be a misrepresentation. As we saw, for the Tractarian Wittgenstein, a proposition was a possibility in logical space (from TLP §3.4 and TLP §3.4.11). Thus the determinate relation, which Wittgenstein argued was constitutive of a proposition (TLP §3.14), could here refer to a logical relation between the various elements of a proposition. What combined the logical and the real for the Tractarian Wittgenstein, was a thought, i.e. a proposition with a sense (where sense refers to the agreement or disagreement with reality). Thus Tractarian *thoughts* are important for us, as we want to understand what makes significant propositions significant; and a significant proposition, within the Tractarian framework, would be defined as propositions with a sense.

[23] See for instance Baker/Morris (2004:55-59)

While *thought* connected the logical and the real in the Tractatus, in the *Investigations*, this was achieved through the idea of use in accordance with (the broad sense of) grammar. Thought however remained in the domain of possibilities, as it connected the logically possible with the possibly real. But the idea of grammar (in the broad sense) goes further in terms of connecting these possibilities with reality (as opposed to a possible reality). In *Zettle* (Z §55) and then later in the *Investigations*, grammar, in the broader sense pointed out above, became what connected thought and reality. Thus one way to address the disparity between the Early and the Later Wittgenstein opens up here. If we understand the determinate relation of the *Tractatus*, that which constitutes a proposition, as the grammar of the later Wittgenstein, a notion which already involves an element of use, then we could retain this aspect of Wittgenstein while giving up the controversial notion of depiction. In this case we would still have a way to connect what is internal to the mind (thoughts) and what is external to it (reality).

Let's try to articulate this possibility further.

(i) Consider the situation where we understand the Tractarian determinate relation as the broad sense of grammar.

(ii) Then grammar in the broad sense already includes an element of use, and

(iii) The determinate relation is that which constitutes a proposition.

(iv) The above [(i), (ii) and (iii)] implies that grammar (which already involves an element of use) constitutes a proposition.

Now consider the following:

(v) A thought for the Tractarian Wittgenstein is a proposition with a sense.

(vi) And the sense of a proposition is the possibility of agreement or disagreement with reality.

(vii) Thus constitutive and essential elements of thought become grammar (from (iv) above) as well as the possibility of agreement or disagreement with reality.

From the above (re)interpretation of Wittgenstein, thought is no longer a connection simply between the logical and the real, but is rather constituted by that which is beyond the realm of the logical and includes grammar in the broad sense, along with the sense of a proposition, which determines the agreement or disagreement with reality.

It would be useful here to explore the arguments of some continuity theorists regarding the interpretation of Wittgenstein's early and later work, as these views are pertinent to our

current discussion. We will especially note where they differ from our interpretation (or rather re-interpretation) of Wittgenstein.

2.2.2 Arguments in Favour of Continuity

There are various continuity theorists in the interpretation of Wittgenstein. I am neither arguing for continuity, nor discontinuity, as to the interpretation of Wittgenstein's works. I do however think it useful to mention some views regarding the continuity thesis as it is pertinent to my own project, i.e. attempting to capture both the aboutness of language as well as its internal significance.

Koethe (1996) argues that the theme that spanned across Wittgenstein's works was that "the semantic and mentalistic attributes of language and human life are shown or manifested by our verbal and non-verbal conduct, but they resist incorporation into the domain of the straightforwardly factual." (1996: ix) More specifically, he argues that "the principle that language's semantic aspects are shown by its use or application is already present in an incipient way in the *Tractatus* and, far from being repudiated, is developed more fully later, freed from an entanglement with the earlier doctrines that Wittgenstein did come to reject." (1996:2) Among the doctrines that Wittgenstein later came to reject was the picture theory, which according to Koethe "offers a conceptualization of what it is to say that something is the case"(Koethe 1996: 43). But, he goes on to argue "if we ask how we determine just *what* is being said to be the case… Wittgenstein's answer is that this is shown by how the sentence is being used" (1996: 43) He cites for example the following passages from the TLP and the Notebooks:

> "In order to recognize a symbol by its sign we must observe how it is used with a sense." (TLP §3.326)
> "What signs fail to express, their application shows. What signs slur over, their application says clearly. (TLP §3.262)
> "the way in which language signifies is mirrored in its use" (NB 82)

This continuity thesis however is different from our proposed re-interpretation of Wittgenstein in the previous section. We argued for an understanding of the Tractarian determinate relation in line with the grammar of the *Investigations*, so that there is an element of use already present within the concept of a determinate relation, which in turn is what makes a proposition a proposition. The way the sentence is used gives us the definite relation, which

in turn gives us the proposition. The above examples make a claim, on the other hand, regarding a further application of use, secondary to the application we have already mentioned. This comes at the level of comparing reality to a proposition (what a proposition signifies or applies to) as opposed to what constitutes a proposition. If what constitutes a proposition is itself dependent upon elements of use, then the later claim about what is involved in comparing reality to a proposition will be not be primary, and would only take place post the first application of use. If both our claim and Koethe's (1996) continuity claim are correct, then there was already embedded in the *Tractatus*, a double application of (the concept of) use. But even if we just stick to our own interpretation (which may or may not be correct, but is interesting with respect to the kind of theory we are seeking), there is an element of use already inherent in what constitutes a proposition. If we were able to systematically understand the nature and origin of this element, it is plausible that we would then better understand the nature of the aboutness inherent in language. In chapter 3, we will make an attempt at such an understanding.

2.3 Wittgensteinian Analogies

Wittgenstein uses various analogies to illustrate his arguments regarding meaning. A few in particular e.g. his analogies of facial expressions and music, could themselves be looked at as (subsidiary) arguments in favour of his theory of (or rather views on) meaning.

2.3.1 Facial Expressions

Seeing an aspect is, Wittgenstein argues, a voluntary act. He famously gives various examples to bring out this point: the drawing that can be viewed either as a duck or as a rabbit; the triangle that can be viewed as a mountain or a geometrical figure. This line of argument also gave Wittgenstein a way to argue against the picture theory he defended in the TLP. He argued that all we have, for instance in the case of the triangle, is three lines connected in a particular way, and what I see it as depends on the concepts I apply to it, all of which can be understood as coinciding with the triangle. Thus recognizing a tiger in a painting is not really comparing the object with the picture but rather the object coinciding with the picture.

In the *Blue and Brown Books* and later in the *Investigations*, Wittgenstein argues that we generally understand a picture as an image in the mind – a paradigm to which we compare reality. However seeing the drawing  as a face does not involve a comparison between a group of dashes and a particular human face. There is nothing besides the group of dashes to

compare it with, no paradigm for comparison. However, one may say that there is in fact something in our head which we are comparing it with, which makes us say for instance that the dashes in the above picture represent a face and those in another, say □ do not. Wittgenstein argues however that there is only an illusion of comparison. What's involved here, are not two distinct processes, one of seeing it as dashes, and the second of seeing it as a face; i.e. it is not a 'composite' experience in this sense. In seeing a painting three-dimensionally, we must not ask what is added to the experience of seeing the painting to make it a painting experienced three dimensionally. For instance any written word, e.g. 'read', is experienced as a word and not just a bunch of scribbles, though it would only be a word to those who recognize and understand English.

Facial expressions are a useful analogy for Wittgenstein, not only with respect to word meanings, "to study what may be involved in 'semantic interpretation'" (Mukherji 2010: 206), but also to think against the picture theory. Because facial expressions occur naturally, hence their significance lies in the expression itself, in some sense (we shall consider this in greater detail in the next chapter). For instance, interpreting facial expressions need not be taught, and the interpretations are, to a large extent, universally uniform. Further, although a facial expression may be viewed as expressing a particular state of mind (e.g. the feeling of joy), the attachment itself (of the expression to the state of mind) is *external to the expression*. In that sense, a facial expression is what it is – i.e. it expresses itself. To put it differently, even if some other state of mind could be attached to an expression in a different context or culture, the expression itself would be invariant with respect to these differential attachments. Therefore, shifting to the phenomena of facial expressions provides Wittgenstein with a handy analogy, where one can talk of significance of an expression without necessarily bringing in something like a picture theory.

How does Wittgenstein, then extend the effects of this analogy to putatively symbol systems such as language (and music, as we shall see in the next section)? This question arises because facial expressions are not symbolic at least in the way in which language is.[24] In order for the preceding non-denotative aspect of facial expressions to extend to symbol systems, we ought to be able to say that the configuration of signs in language are also internally significant

[24] The analogy is therefore limited, given that, although the significance of a facial expression can be grasped without a specific reading of the state of mind involved, there is really nothing else to read in a facial expression except someone's state of mind.

in the sense that their significance can be grasped independently of *external attachments*. We can, therefore, see the need for a notion of grammar.

Before we proceed, it would be useful to briefly note another Wittgensteinian argument against the picture theory, which is the received view from the *Investigations,* i.e.: to understand the meaning of a word one has to look at the various uses of that word. Words function in various ways, we can do various things with words, and to know how a particular word functions, we must look at the various things for which it is in fact used. Meanings of words are conventional. Conventions are, simply speaking, arbitrary actions, reproduced through weight of precedence (Millikan 1998). They do not *require* that people must follow them, but people do so simply by weight of precedence. Now what would this signify for language? I cannot, for instance, decide that I will henceforward use the word 'cold' to mean what we normally use the word 'chair' to mean, and expect that others will understand what I say. So there appears to be a certain sense of normativity involved in our use of words. But this sense of normativity is not prescriptive but rather involves meeting a standard (arbitrarily set) in order to achieve, for instance, successful communication (Hattiangadi 2006). Association of a sound with its meaning is arbitrary just like the association of agreement with a certain neck movement or association of wink with mischief (I could have used a wink to mean dinner is ready) But the arbitrariness does not make it prescriptive or social in the sense of taking an oath to use the word 'chair' to mean what it is normally used to mean, or something of the like, as perhaps part of an initiation into the English linguistic community.

2.3.2 Music and Language

Another recurring analogy, along with facial expressions, which we find in Wittgenstein is the analogy of language and music. Wittgenstein argued in the *Tractatus* (Wittgenstein 1922: 4.014) that: "A gramophone record, the musical idea, the written notes, and the sound-waves, all stand to one another in the same *internal relation of depicting* that holds between language and the world. They are all constructed according to a common logical pattern. (…They are all in a certain sense one.)" (Emphasis added)

It is interesting to note that in the instance of comparing language and music it is easier to understand the relation between the various elements at the musical end of the comparison. The musical idea is captured by the written notes, and it is on the pattern of the written notes that the gramophone record is constructed (though the medium and the language differ). The sound waves again will follow the same pattern (though in a different dimension).

We may note here that each of these elements has a certain pattern, a symbolic quality attached to it. There appears to be a mapping in place: systems (of patterns or symbols) appear to be mapped on to one another, translated in a different medium. "There is a general rule by means of which the musician can obtain the symphony from the score… It is the rule for translating this language into the language of gramophone records." And "That is what constitutes the inner similarity between these things which seem to be constructed in such entirely different ways." (Wittgenstein 1922: 4.0141) He thus recognizes that each of the elements in the comparison are themselves different kinds of languages (or perhaps different forms of symbol systems – in different mediums). There appears to be nothing problematic about the translation of a sentence in one language to another language. The problem rather arises with the notion of depiction: language depicts the world, an idea that Wittgenstein himself gave up later in the *Investigations*.

Let's now look at the language end of this comparison. The elements include – a linguistic thought (as opposed to a musical idea), which is expressed by written sentences in a language. The spoken sentences and the sound waves share the linguistic thought in some physical sense. The spoken sentence expresses the thought and is in a sense contained in the sound waves. It is interesting that here the world and depiction have not entered our description yet at all. In the TLP the depiction relation was captured at the first of these stages: the thought. However, the depiction relation was not central to the analogy noted above in the first place (and was given up by the later Wittgenstein). This is because there is, as we noted, no depiction involved in the music case at all. There is, in the case of music, only translation from one language to another. Translation, we may note, is a lexical and grammatical notion, and involves no non-linguistic component. To put it differently: the factuality of the world does not enter into translation; it involves only linguistic items (words) and grammar. Thus the internal relation that allows for the translation is a grammatical rather than a depiction relation. The internal grammatical relation retained its position as the central point of the analogy. The centrality of the notion in both language and music appears to be fully appreciated in the remark in the Investigations "understanding a sentence is much more akin to understanding a theme in music than one may think." (PI § 527)

The direction of the analogy suggests that if we understand what 'understanding music' is all about, then we would be able to understand language along the same lines. Towards this end of trying to understand music, Wittgenstein points out that "it has sometimes been said that what music conveys to us are feelings of joyfulness, melancholy, triumph etc." This seems to

suggest that music is an instrument for producing these feelings and that any other way of producing these feelings would also do instead. However, he argues, this appears to be a misconception and rather "music conveys to us itself" (BBB 178).

"There is a tendency to talk about the 'effect of a work of art' – feelings, images etc." (LA 29) If these effects could be produced by something else, could that replace the work of art?[25] "doesn't the minuet itself matter? – hearing *this*: would another have done as well?" (LA 29) As Wittgenstein argues, "reporting a tune to ourselves and letting it make its full impression on us, we say 'This tune says *something*' and it is as though I had to find *what* it says. And yet I know that it doesn't say anything such that I might express in words or pictures what it says."[26] Now, to say that it expresses a musical thought would be saying just that – it expresses itself. The following points can be noticed here:

(1) We understand a tune as significant

(2) This significance, however, cannot be expressed in words or pictures.

So how do we understand it as significant? As a possible reply, Wittgenstein argues further that you can't just play a piece of music anyhow – it must be played in a particular way. We may say: this isn't how it is supposed to be played; and here it appears that there "must be a paradigm somewhere in our mind,... But in most cases if someone asked me "How do you think this melody should be played?", I will, as an answer, just whistle it in a particular way, and nothing will have been present to my mind but the tune actually whistled (not an image of that)" (PI §163).

This suggests firstly, that there is a notion of 'correctness' involved, in the performance of a piece of music. Secondly, however, the correctness involved cannot be defined or explained in any other way except by demonstration. If someone sings a particular piece of music, we sometimes have the inclination to say, for instance, the pitch is not right, or this

[25] We may note that though he uses the generic term 'art' Wittgenstein focuses his discussion on music. The reason for this can also be seen immediately, as other forms of art, such as painting, may have representational features that music lacks. In any case, a general conception of the arts is not the topic here.

[26] Compare Wittgenstein's discussion of the natural color pattern of the pansies in the Brown Book. Wittgenstein points out that one can be so impressed by the patterns so as to say for instance that 'each pattern says something'. He argues that though we are 'trying to distinguish between meaningless patterns and patterns which have meaning;' however 'there is no such expression in our game as "This pattern has the meaning so and so"' (Brown Book, page 179) In the case of the pansies, however, we can also replace the statement with say 'each pattern impresses me'. This is because, there is not here, a recognized need to seek 'the something that they express'. Thus the intransitive form would also work.

theme must be repeated etc. Wittgenstein observes that "Yet there just is no paradigm apart from the theme itself. And yet again there *is* a paradigm apart from the theme: namely, the rhythm of our language, of our thinking and feeling. And the theme, moreover, is a *new* part of our language; it becomes incorporated into it; we learn a new *gesture*."

Comparing music with facial expressions, Wittgenstein remarks that when we say "This face says something" and follow this remark with say "The whole face expresses bewilderment", then, "here we have used 'expressing' transitively." Similarly, in the case of music, we do feel the need to find what it expresses. But, as the above discussion suggests, for Wittgenstein, music expresses itself. This however does not mean that there is nothing that it expresses, but rather there is nothing *apart from itself* that a piece of music expresses, thus expresses is used reflexively.[27] Therefore, what happens in the case of music is that

1. it is similar to facial expressions in the sense that the states of mind, like the feelings, emotions etc. that can be read in a musical expression are also external to it in the same way as in the case of facial expressions, but
2. there is a stronger notion of internal significance to musical expressions even when we set aside all possible emotional content from it. This is because in the case of music the internal significance of musical expressions can be traced exclusively to the configuration of the signs.

In a sense, therefore, the 'itself' in Wittgenstein's claim that music expresses itself is profoundly non-trivial in the case of music. Now if language is like music then it too must convey itself in a(n) (internally) significant sense. It was perhaps this realization that led Wittgenstein to abandon the picture theoretic notion of depiction and thereby the picture theory itself. Language was to be understood on the model of music, and not vice-versa. So if we cannot make sense of depiction in the musical case, then even in language, the notion of depiction must not be theoretically salient.

[27] Roger Scruton (2004) argues that Wittgenstein's comparison of musical understanding and the understanding of facial expressions does not have the implication that he seems to draw from it, that musical expression is confined to the intransitive idea. We argue however, that Wittgenstein was not claiming that musical expression is intransitive, but rather that it is a reflexive relation. Here we may define the terms in use as follows. A verb is intransitive if it has a subject but not an object. In the sentence 'Music conveys to us itself' the verb 'convey' is used, and it is a transitive verb. A transitive verb requires both a subject and one or more objects; the meaning of a transitive verb is incomplete without a direct object. Here the direct object, however, is the subject itself. Thus in this use, the transitive verb is used reflexively. A reflexive verb is one whose semantic subject and direct object are the same. In a wider sense, it refers to any verb form whose grammatical object is a reflexive pronoun (himself, itself etc). Thus Wittgenstein, as is evident from his use of terms, was not confining music to the intransitive idea, but pointed out rather, that the musical expression has an object, and this object is, non-trivially, *itself.*

2.3.3 Recent Accounts of Musical meaning

Levinson (2003) argues that music is a form of thought: musical passages bear the appearance of thoughtful acts. He supports the thoughtful nature of music with a range of musical examples. Thus, we can think of music as drawing a conclusion (Beethoven's Piano Sonata op. 110, Dvorak's 7th Symphony), coming to a close (minuet movements from Mozart's 40th or 41st), asserting (the opening of Schubert's Piano Trio No. 2, Op. 100), questioning (the opening phrases of Beethoven's Piano Sonata No. 18, op. 31 no. 3), imploring (the flute introduction to Bellini's aria 'Casta Diva'), defying (the opening of Beethoven's Fifth Symphony), disapproving (the orchestral interjections in the first part of the finale of Beethoven's Ninth Symphony), etc.

Further, he argues that musical passages give evidence of thought processes in their creator, e.g. "the assessment we infer Bach must have made in devising a fugue theme combinable with itself in counterpoint, or the judgment we suppose Mozart to have exercised, in composing a piano sonata, in designing a second theme whose character would contrast suitably with that of the first theme"(Levinson 2003: 63) and so on. He argues further that the "mere succession from chord to chord, motive to motive, or phrase to phrase at every point in any intelligible piece of music" gives evidence of "*inherent musical thinking*"(italics in the original) as "it is a purposive-seeming, goal-directed temporal process, an intelligent form of continuation in time" (Levinson 2003: 64)

Thus it appears that Levinson does emphasize the structural element. However, he proceeds to claim that understanding music, like understanding language is a matter of use, "of knowing how to operate with the medium in question… in particular contexts." (2003:68) We come to understand music as well as language "by learning how to respond to it appropriately, how to connect various bits of it to one another, and how to ground it in our lives."(Levinson 2003:68). I think the way we have interpreted Wittgenstein in this chapter may even be useful for buttressing Levinson's discussion. The idea is that if use becomes constitutive of propositions then understanding the way language is used, the way it is grounded in our lives, is also involved in, and involves, understanding language as such.

Kaduri, however, accounts for the idea of structural relation in Wittgenstein's remarks on music, through his study of the general pause – silent figure (aposiopesis) – in Haydn's instrumental music. He notes that Haydn's music 'contains so many general pauses that it seems they form an intrinsic component of his musical language.' (Kaduri 2006: section 3) The pauses seem to fall under definite 'style'-specific categories, differing as styles differ: string

quartets, symphonies, movements in sonata form, rondo movements and minuets. According to Kaduri, thus, the ‘different ways in which Haydn employs the general pause and the logical links between them constitute a "grammar" of the general pause, which provide its meaning.’ (2006: section 3) This suggests how musical understanding can be explained in part from structural conditions alone. Thus Kaduri’s analysis comes closest to that which we have so far attributed to Wittgenstein.

Jakob Christensen-Dalsgaard (2004) argues that “the true musical universals – that slow, low-pitched sounds are ‘sad’ and fast, higher pitched sounds joyful (or aggressive) – would also be prosodic universals.” (2004: section 2.6)

As Mukherji (2010) points out, three main points emerge: Firstly, the same piece of music is not always viewed to have the same affect: the first movement of Mozart's G minor symphony, now considered tragic, was viewed by 19th century critics as cheerful. Secondly, “two very different pieces of music, from the same or different genres can both be viewed as cheerful: Mozart’s *Ein Kleine Nachtmusik* and the medium-tempo parts of performances of, say, *raaga*s *Jayjawanti* or *Bilaval*. If the ‘cheerfulness’ is traced to the tempi, then the tonal structure of music becomes irrelevant, apart from the implausible consequence that all music in a certain tempo would have to be viewed as cheerful.” (Mukherji 2010: 204) Suppose, for instance, we think of the Beatle’s song *Michelle* as generally sad. Ignoring the words, could the ‘sadness’ be traced to the fact that much of the melody in *Michelle* ‘moves in a relatively small range in the mid-to-low vocal range, with a generally descending contour,’ as Jackendoff and Lerdahl (2005: 63) seem to suggest? Unless ‘sadness’ is associated with this range by definition, one can easily cite sad music that ‘wails’ as the notes ascend in the mid-to-high range. On the other hand, (the early part of) the slow movement ‘in a relatively small range in the mid-to-low vocal range’ of the so-called ‘morning’ *raaga aahir bhairav* is meant to evoke not sadness, but spirituality tinged with joy as the ‘dawn breaks.’

Thirdly, Mukherji raises the question whether this [that slow, low-pitched sounds are ‘sad’ and fast, higher pitched sounds joyful (or aggressive)] is a *musical* universal. It may be argued that this universal applies to our *appreciation of* music, rather than music itself. Can we say, for instance, that in our appreciation of music, we tend to superimpose on the formal features of music, the emotional significance we attach to the prosodic features of speech? If this is the case, then it would also counter the idea that music imitates certain (prosodic) features of speech. If the features are attached or superimposed on the musical composition,

then they are not a part of the structure of the music itself. Thus, if this is the case, then it makes no sense to argue that music imitates the features of speech; these are, rather, externally attached in the appreciation of music.

Here it is interesting to note that Langer (1942) argues that "there are certain aspects of the so-called "inner life" – physical or mental – which have formal properties similar to those of music – patterns of motion and rest, of tension and release, of agreement and disagreement, preparation, fulfillment, excitation, sudden change etc." (Langer 1942: 46) To say this however doesn't amount to an explanation, as it is possible to find similarity between most things. Each of these, we may note, however, are structural qualities. Thus perhaps, due to the structural similarities between aspects of the inner life and music, people tend to attach an affective element to music. Along similar lines, one may argue that Mukherji (2010) is missing the point entirely, i.e. perhaps the point is rather that our *understanding* of music imitates our understanding of certain features of speech/language. For instance, we find the following passage in the *Zettle*: "Doesn't the theme point to anything outside itself? Yes it does!" (Z §175) Wittgenstein goes on to explain what he means as follows: "it (the theme) makes an impression on me which is connected with things in its surroundings – e.g. with our language and its intonations; and hence with the whole field of our language games." (Z §175) As was pointed out earlier, there is perhaps an element of prosodic universals in the experience of a piece of music. And this is again due to the *fact* (as opposed to the necessity) that we do have such a language – with such prosodic features.[28] Lewis (1977) in fact mentions the above passage as falling in line with the rest of Wittgenstein's writings: "The sense in *Zettle* could be captured, I think, by saying that understanding a musical theme overlaps with or is not isolated from other aspects of human life" (Lewis 1977: 120).

Given these arguments, it appears that the focus for Wittgenstein, is not language as such or music as such but rather how these are incorporated in human life, how they are used and understood. The above discussion suggests that though music indeed does express itself, it also points to something else outside it, i.e. the contours of our language and its use. To put it differently and perhaps more appropriately, the only external significance that can be attributed to music is that which it achieves simply by being a part of human life, a life which is distinct from other life forms mainly by being linguistic. That is, being a language user influences all

[28] For instance – there appear to be fine distinctions in the music and the interpretation of music, corresponding in some way to the particular language.

aspects of human life and music is no exception. In understanding and appreciating music, we do so importantly as a language user, i.e. language use serves as the background of our understanding music (and as the background of our understanding as such).

Almost by contrast, understanding language is for Wittgenstein (PI §527) like understanding music (and not the other way around as suggested above). And, as we just saw, language use serves as the background of understanding music. Therefore, language *use* must also serve as the background of understanding language. So, although language is internally significant and "the content of the sentence is in the sentence" (BBB167), this internal significance can only be understood within the confines of language use. In other words, the way we use language and all the various language games, form the background for understanding. This is because (a) language must be understood the way we understand music, and (b) music in turn is understood within the background of language. Thus language must 'convey itself' the way that music does, but music conveys itself so within the context of language use. And language is used essentially within a community of language users. Therefore, language must be internally significant but only as a part of the life of language users.

Let us compare this understanding of a background (of language use) with Searle's (technical) usage of Background. Searle talks about "a network of representations" and a "Background of nonrepresentational mental capacities" and insists that "in general, representations only function, they only have the conditions of satisfaction that they do, against this nonrepresentational Background" (Searle 1983: 21) Regarding the former, he explains that to have any single intentional state, involves having a series of other intentional states (comparable to Quine's network of beliefs). For instance, if I intend to go to a convention: "One finds that any such intention is embedded in a network of intentional states and that it can have the conditions of satisfaction that it has only in virtue of its position in the network. In order that I should have specifically that intention, with the content that it has, I must believe and desire a whole lot of other things which go far beyond the logical consequences of the corresponding statement. I have for example a whole series of beliefs about airlines, and geographical locations, and hotels, and professional associations. And these beliefs in turn are connected to a series of desires about tickets, reservations and travel arrangements." (Searle 981: 731) To put it differently, "in any real life situation, the beliefs and desires are only part of a larger complex of still other psychological states; there will be subsidiary intentions as well as hopes and fears, anxieties and anticipations, feelings of frustration and satisfaction. For

short, I have been calling this entire holistic network, simply, the 'Network'" (Searle 1983:141).

This network of beliefs and desires functions within a "background of mental capacities that are not and could not be representations." (Searle 1981: 731) Searle argues that "intentional contents are not self-interpreting. One has to know how to apply them, but the know-how in question cannot consist entirely in further intentional contents, without an infinite regress… Suppose, for example, I now intend to drink a glass of water. That very token intention will not be satisfied if I 'drink' the water through my ear, if the 'water' is in gaseous or frozen form, or if the 'glass' is a ten-thousand gallon glass. But none of these conditions are represented in the original intention; it contains no representation of any of these requirements." (1981: 732) In other words, these conditions are not a part of the network of intentions, but are rather nonrepresentational mental capacities or know-how that are needed to interpret the intentional content. "The Background, therefore, is not a set of things nor a set of mysterious relations between ourselves and things, rather it is simply a set of skills, stances, preintentional assumptions and presuppositions, practices and habits" (Searle 1983: 154)

Now let's try to use these tools towards our interpretation of Wittgenstein. We argued earlier that for Wittgenstein language use must serve as a background for understanding language. Searle's account of Background as applied here would mean that language use involves nonrepresentational mental capacities or know-how needed to interpret intentional content of language. In other words, it is through knowing the rules of language use (although one may not be able to specify or identify them per say) that we are able to interpret the language used by others, in terms of its intentional content, and thereby communicate successfully. For instance, taking the above example from Searle, to interpret correctly a statement of the form: 'I intend to drink a glass of water' the interpreter needs to be aware of the Background within which this statement can be satisfied. This Background, as Searle points out, involves the knowledge of how these terms are used: i.e. one cannot *drink* through the ear but only through the mouth; to drink *water*, the water must be in liquid and not gaseous form; and to be able to drink a *glass* of water, the glass must not be so huge as say containing ten-thoudsand gallons of water. There could however be other contexts of use, for example a story (like say Alice in Wonderland) within which these very forms of use, which we suggested fall beyond the scope of normal language use (drinking through the ears, drinking gas, and drinking ten-thousand gallons in a glass) would make sense. But, given the rules of use of this new context, it would only involve understanding these rules, for us to interpret these statements

correctly. Thus the nonrepresentational mental capacities (or Searle's Background) involved in interpretation could be the capacity to understand the rules of language use, given a context. Alternatively, the mental capacities could be understood as the know-how needed to interpret intentional content. This know-how would involve the capacity to comprehend (or even internalize) the rules of use in a context (without necessarily being able to identify which rules are being used), and interpret content accordingly.[29]

2.4 Implications in Terms of Our Original Intuitions

We saw in this chapter that for the early Wittgenstein of the *Tractatus*, thought was a proposition with a sense. A proposition was the picture of reality, and a picture shares with reality a logico-pictorial form: the possibility that things in the world are connected with each other the same way that elements of a picture are connected. The domains of language and reality are thus distinct and self-contained, but they mirror each other. Language or pictures have their own internal connections, following an internal logic. What makes a sentence a sentence is that the words stand in a determinate relation to one another (TLP §3.14) [30]. In the Philosophical investigations (PI § 502)[31] and the Blue and Brown Books (BBB 167)[32] we find a similar line of thought, but here, as in *Zettle* (Z §55), the determinate relation of the *Tractatus* has been replaced by a broad sense of grammar as explained in section 2.2.

Returning to the *Tractatus*, a proposition with a sense constituted a thought, and the sense of a proposition was its agreement or disagreement with reality. Unless you believe in the possibility of perfect knowledge, there is always the possibility of error in assessing the said agreement or disagreement. In order to reduce the chances of error, we need, at least another mind. But thoughts in other minds are not available to us, we cannot access the propositions entertained by other minds: all we can observe, what is available to us is behavior,

[29] We will attempt in chapter 5 to use a Davidsonian decision theoretic framework to individuate this intentional content, given the *use of language*.

[30] "What *constitutes* a propositional sign is that in its elements (the words) stand in a determinate relation to one another." (TLP § 3.14, Emphasis added). So here it is constitutive of the sentence or proposition that the words stand in a determinate relation to one another.

[31] Here Wittgenstein argues that the following are comparable:
"This sentence makes sense." – "What sense?"
"This set of words makes a sentence." – "What sentence?" (PI § 502)
The claim is that a sentence is a combination of words in that particular order, where the words have specific grammatical relations between them, and that combination of words itself and nothing else, is the sentence. Similarly, the sense of the sentence lies in the sentence itself. See section for details.

[32] "[T]he content of the sentence is in the sentence" (BBB167)

the actual uses of words. This was perhaps among the reasons (along with others mentioned in this chapter), that drove the later Wittgenstein to claim that meaning is use - as all we have by way of evidence for meaning is the use of words in contexts. From the above[33] it appears that the notion of *Thought* involved a connection of internal content (the proposition with its internal logic or determinate relation) and external significance (agreement or disagreement with reality in the TLP and later the use theory in the *Investigations*). The later Wittgenstein also retained both these elements: internal content (PI § 502) and external significance (use theory). For the later Wittgenstein, the connection between language (or thought) and reality was achieved through the notion of grammar (where grammar was to be understood in the broad sense). While the sense of a sentence is to be found in the sentence itself, the meaning of a sentence is nothing but the use of that sentence. We noted a possible way to connect these disparate claims through re-interpreting the determinate relation of the *Tractatus* as the grammar of the later Wittgenstein. This would entail that an element of use is already a part of what constitutes a proposition. This may or may not be Wittgenstein's intention (most likely, it was not), but it does provide a useful way in which we could possibly connect the aboutness of language to the idea of an internal significance (that meaning must already exist within a symbol, for it to have meaning). In what follows, we shall explore this connection further, to see whether there is a systematic way in which we can understand the idea of use being constitutive of a sentence. Thus our study of Wittgenstein has led to a possible way in which we could satisfy our intuition that meaning must already exist within a symbol for it to have meaning, while also provide a way to account for the inherent aboutness of language. Whether this alternative is feasible or not forms the matter of the rest of this thesis.

[33] See section 2.2.1 for details.

Chapter 3: Systematizing Use: Towards a Use-Theoretic Account of Meaning

3.1 Introduction

In the previous chapter, we identified a possible way to reinterpret Wittgenstein in order to retain both an internal significance and the aboutness of language. The way forward, we found, was to reinterpret the determinate relation of the *Tractatus* as the grammar of the later Wittgenstein. Among the various uses of the term *grammar* in the later Wittgenstein, we also find the use of a word/sentence as well as the rules of use. Moreover, determinate relation was that which constituted a proposition. Therefore, the use of a word/sentence could potentially already be an element of that which constitutes a proposition. If use is constitutive of a proposition in this way, then use and meaning become intertwined. In other words, it would not be possible to know the complete meaning of a word/ sentence without knowing its use. Consequently, *use* which accounts for aboutness is incorporated within internal significance or meaning. However, given the various uses of *grammar* in Wittgenstein, it remains to be seen how we may best interpret the notion of *use*, to serve our purposes.

Shifting from our interpretation of Wittgenstein in the previous chapter, to the accepted interpretation, one that maintains a fundamental discontinuity between Wittgenstein's earlier and later work, we find what has traditionally been associated to the later Wittgenstein: the use theory of meaning. The use theory of meaning, as it appears in Wittgenstein's later writings, leaves us with a notion of use that is both obvious and vague at the same time. The legacy tells us that the meaning of a word is its use in language, and you can't really *say* anything about the meaning of a word except to *show* it through the use of that word in a context. So what is 'use'? We are told that it has to do with following rules, however, rules which are not set in stone. Meanings, thus, form a complicated network of what Wittgenstein terms 'family resemblances', as opposed to the direct word-object relationship of simpler times. The picture practically precluded a deeper understanding of what 'use' is by virtue of the saying/showing

distinction.[34] To explain what the meaning of a word is, the best you can do is to in fact just use it in a context.[35]

Even at first glance, the idea that meaning is use inherently involves a tension: given that use refers to particular instantiations, and meaning requires us to abstract from particular instantiations and arrive at what is common among them. Further, we need to distinguish between the meaning/use of a word as opposed to that of a sentence, to see whether both can be arrived at in a similar way. The use or instantiation of a word involves the utterance of the word within a sentence, in a particular context. Here the sentence within which the word occurs may differ, along with the context of use, while the meaning of the word may remain the same. Thus for instance two instantiations of 'dog' would be 'Reena took her dog for a walk' and 'Diya's dog was barking loudly in the park' Here the sentence and the context is different, though the meaning of the word 'dog' remains the same in both instances. In the case of the meaning of a sentence however, the use or instantiation of a sentence involves simply the utterance of the sentence, and the variable is the context of utterance (as opposed to the use of a word where both the sentence and context are variable).

Getting back to the original problem (that use requires specifying particular instantiations, while meaning requires abstracting from particularities), Gupta (2003) mentions that a "natural way of dealing with the problem is to appeal only to structural aspects of use in the account of meaning"[36] However, if we appeal only to structural aspects in the account of meaning, what would this involve for both word and sentence meaning? In the case of word meaning, the occurrence of the word within a sentence could give us various structural properties of the word itself (e.g. that it is a plural, noun etc), through the position it occupies within the sentence. The structural aspects of the use of a sentence, however, would involve structurally analyzing the sentence itself, essentially devoid of the context of occurrence, and therefore would amount to the same as what would be involved in understanding the meaning of the sentence. It appears that words and sentences require different treatment. This is

[34] In this chapter (as well as in this thesis) we will be sticking to our interpretation of Wittgenstein, and therefore, we will not be elaborating on the received view of the use theory of meaning, as it appears in his later writings.
[35] Here the question arises: is it in the nature of the concept 'use' that there is no more systematic way of analyzing it than just 'showing' what the use is *through* a particular use. See section 5.4.1 for further discussion.
[36] Gupta continues, however, that "it is completely mysterious how the relevant structural aspects can be specified without any appeal to meanings." (2003: 663)

precisely what we get in Horwich (1997, 1998, 2004), who has differing accounts of word and of sentence meaning. Horwich's theory is particularly interesting with respect to the direction of our own inquiry, as we shall see below, given that within Horwich's approach, use becomes constitutive of propositions though it does not constitute the entirety of meaning. Among other theories stemming from the Wittgensteinian maxim 'meaning is use' (e.g. Dummett's justificationist semantics 1959,1991, 2002, 2006 and Brandom's inferential role semantics 1994, 2000, 2007), Horwich is unique in that he attempts to systematize use (whether he succeeds in his attempt, and whether this is at all possible or not, are points we will be considering later). Though both Dummett and Brandom subscribe to the above maxim, they do so in different ways, understanding use in terms of other concepts.

Following elements of the early Wittgenstein and logical positivists, Dummett argues that the meaning of a statement should be specified in terms of the grounds of asserting the statement. Thus in his early paper 'Truth' he argued that we should "no longer explain the sense of a statement by stipulating its truth-value in terms of the truth-values of its constituents, but by stipulating when it may be asserted in terms of the conditions under which its constituents may be asserted' (Dummett 1959:17-18). For a justificationist, understanding the content of a statement consists in "an ability to recognize evidence for [the statement] when presented with it, and to judge correctly whether or not it is outweighed by any given piece of counter-evidence' (Dummett 2006: 59) Central to his position is his argument for Anti-realism, through a rejection of the principle of bivalence, along the lines of the mathematical intuitionists. "If a mathematical statement is true only if we are able to prove it, then there is no ground to assume every statement to be either true or false… once we have lost any reason to assume every statement to be either true or false, we have no reason, either to maintain the law of excluded middle" (Dummett 1991: 9). It follows, for example, he argues, that "no one taking an anti-realist view of the past could suppose every past-tense statement to be true or false, for there might exist no evidence either for its truth or for its falsity" (1991:10). Drawing inferences, giving arguments in support of some conclusion, is, he claims, an essential part of using language. "Hence a theory of meaning ought, as one of its tasks, to give an account of the procedure of deductive inference, and supply a criterion for the validity of such an inference… a semantic theory… explains how the fundamental semantic property of any statement is determined in accordance with its composition in order to characterize valid deductive inferences by reference to this." (Dummett 2004:31-32) Further, "What a deductive

argument, to be valid, must preserve from premises to conclusion is justifiability, where a statement is justifiable if it is possible to justify it." (Dummett 2006 :64)

Brandom in turn contends that "concepts are essentially inferentially articulated. Grasping them in practice is knowing one's way around the properties of inference and incompatibility they are caught up in." (Brandom 1994: 89) He argues further that "Inferring is a kind of doing" (1994: 91), however we do not need to be able to explicitly formulate the rules of inference that we are following. "The idea is to understand propositional contents as what can both serve as and stand in need of reasons, where the notion of a reason is understood in terms of inference. So propositional contentfulness is taken to be a matter of being able to play the role both of premise and of conclusion in inferences." (Brandom 2007: 654)[37]

Thus, while for Brandom, it is through a network of inferential connections that we understand concepts (resonating both Quine and the later Wittgenstein), Dummett maintains the salience of justifiability (and even truth) in significant uses of expressions. Both, however, subscribe to the idea that the meaning of an expression is its use in language, though use in one case is determined through inferential connections and in the other through justifiability. Here 'use' is not analyzed in and of itself, but is rather understood in terms of other, perhaps even more complex concepts, such as inference and justification. There are indeed many arguments for and against these theories, but we need not get into these details. In an attempt to see whether it is possible to systematize use, it would prove unhelpful to focus on theories which do not deal directly, as it were, with the concept of use.

It may be argued however that these theories do go some way towards systematizing use through considering a single property of uses, for instance, that they are justified under certain conditions. So we have a certain sentence x containing the word U. An utterance of Ux is an instance of use (of the word U). There are various ways to systematically understand Ux.

A1 We could analyse and understand Ux in itself: the particular instance of use.

A2 We could understand Ux in terms of other instances of use: Uy, Uz etc.

[37] Clearly connected to this is Quine's web of belief argument. "The totality of our so called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or to change the figure, total science is like a field of force whose boundary conditions are experience." (Quine 1951: 42) Horwich wants to accept this picture, while also advocating non-holistic meanings of words which are specified through specific regularities of use.

A3 We could also understand Ux in terms of the conditions that justify this use: C1, C2 etc.; inferential relations with other sentences and so on.

In order to fully understand the meaning of U, it appears, we need each of A1-A3, and in that order. The claim here is that understanding the meaning of the word is dependent upon a) understanding the relation of the word to others in the use-sentence; b) understanding the relation of the particular use sentence, Ux, with other use-sentences, Uy, Uz etc.; c) understanding the relation of Ux to external justifiability conditions or inferential relations etc. As we move from a to b to c, we are increasing the level of complexity and dependence. So, in claiming that the only way to specify meaning is in terms of conditions of justification, is also to claim that the simpler elements are not analyzable in themselves or that the meaning we get from them is incomplete. So before one goes on to A3, we should have at least considered the possibility of A1 and A2. Horwich appears to be dealing with use itself, in advocating a use theory of meaning (though we shall see that he is dealing more with A2 than A1, or rather merging A1 and A2). We shall therefore proceed to elaborate Horwich's picture given its priority (over explanations dealing with say A3 above) with respect to our present endeavor: systematizing *use*.

For a use theory to qualify as systematic, it should at least *explain* what the use of a word is, and tell us *how* to arrive at meaning through use. A highly systematic use theory would also be able to systematically derive the various uses of a word. Horwich tries to systematize a use theory of meaning through a set of claims regarding what the use of a word amounts to and how this contributes to the meaning of a sentence. In this chapter, we will see whether it is possible to systematize a use theory of meaning, focusing on Horwich's attempt, and how much of this approach we can use for our own enquiry.

Horwich argues for the following:

B1 Word meanings are use properties (or 'implicit definitions').

B2 Sentence meanings are derived from combining word meanings and assuming compositionality.

B3 Compositionality is not a substantive concept, but has rather a deflationary interpretation: compositionality does not constrain the content of word meanings in any way; meanings of words could be use properties, truth conditions or whatever else.

We shall consider each of these steps to see whether they are internally consistent as well as coherent as a whole. In case the theory does not stand against criticism, we shall try to

come up with an alternative, or an addition that could potentially strengthen the theory. Now we will proceed to elaborate on B1-B3 above, as well as note some objections to these points available in literature.

In Section 3.2 we will discuss Horwich's position that meanings of words are their use properties and point out various objections to this position. Then in section 3.3 we consider his claim that sentence meanings are arrived at trivially from word meanings when we assume compositionality, and objections to this. In section 3.4, we shall try to build a defense for Horwich by first taking a closer look at what he means by use-properties, and revisiting the objections raised in section 3.2.1 Then we will consider his claims regarding sentence meaning and see whether we can strengthen the response to objections raised in section 3.3.2. Then in 3.5 we proceed to point out our final objection. We shall conclude by relating what we learn from this chapter to the goals of our thesis: tracking our Wittgensteinian intuitions.

3.2 Horwich's Use Theory of Meaning

For Horwich: "The meaning of a word, w, is engendered by the non-semantic property of w that explains w's overall use.-And this will be a property of the following form: that such-and-such sentences containing w are the ones that are accepted underived (by the relevant 'experts', given certain meanings attached to various other words)" (2004: 352)

Since what is important here is what is *considered* to be true, and not 'truth' itself, how is it that Horwich can talk of *regularities* of use at all, as what one person considers to be true may differ greatly from another. In this regard Horwich resorts to 'the relevant experts'. This does not pose a major hurdle to his story, as he claims that whatever the implicit definition, it is a matter of empirical inquiry and may be changed over time.

But here we may ask, what experts are we to consult in determining these implicit definitions? The definitions are vague enough for ordinary speech, so perhaps the 'relevant' expert here are any competent speakers of English. In case of technical terms in any field, we may need to consult an 'expert' in that field, or someone who understands that language. As I understand his argument, the inclusion of the term 'expert' does not add much, as an 'expert' is just someone who understands the language (English, mathematics, science, etc).

Michael Devitt (2011) argues, however, that Horwich's appeal to experts ("deference" for Devitt) is in fact what proves irreconcilable within his use theory of meaning. The main

reason he gives for this is that an expert's use/acceptance property is greatly different from a non-expert's. So deference may explain why we change our use property in cases of ignorance or error, but not the meaning it now has. As far as I understand it, once we change our use property, the new use-sentence is again basic and in fact constitutes the meaning, and so the question of being able to explain the new meaning does not arise. However, Devitt may have a point when it comes to the difference in the acceptance properties of experts and non-experts. Experts can change the use properties based on empirical evidence, which may be causal in nature, and thus we need, he claims, a causal theory to explain expert meaning, and a use theory to explain non-expert meaning. However, though experts may change meaning based on empirical observations, which Horwich would admit to, their understanding of the meaning of the word may still be its use property (which is empirically determined). To put it differently, there may in fact be a difference in the way the expert and the non-expert arrives at the particular use property: the expert's beliefs are empirically determined while the non-expert can simply defer. But Horwich may well accept that and still claim that the two share the same use-property: the expert determines which is to be the basic use property, and the non-expert accepts such a determination, thereby sharing the belief. Once a particular use-property is accepted as basic, how it came to be regarded as such is no longer the issue.

Devitt could be making a different point: an expert may accept many sentences regarding A, most of which a non-expert may be unaware of; therefore, the expert and non-expert meaning are different. However, somewhere within the claims the expert accepts will also be the basic use property, which he will also accept as such (perhaps with the qualification of acceptance within say non-technical usage). And in all such usage, the expert may also share the use property of a non-expert. Thus deference does not prove fatal to Horwich's account. Devitt also argues, however, that Horwich fails to explain what he means by deference and does not give any theory as to how it works,[38] whereas Devitt himself can offer such an alternative (his causal theory of reference). I agree that Horwich has not attempted to explain or give a theory of the phenomenon of deference. Rather he claims that for most common words, almost everyone will count as an expert, thereby making the underived acceptance of a sentence by a community of speakers *including the speaker himself* relevant. I tend to

[38] He gives five reasons why Horwich's account isn't sufficient to explain deference: "(i) People will often not defer where they should. (ii) They will often try to defer but fail to identify an expert. (iii) They will often defer to a nonexpert. (iv) When the bearer of a name has been long-dead – for example, Aristotle - there will be no experts around to defer to. (v) Where there are surviving experts about a dead person, there seems to be no change that a deferrer could be disposed to make to conform to the experts' basic acceptance properties." (Devitt 2011: 209)

sympathise with Horwich here. Though he has not explained the phenomena, deference isn't at issue in most cases of use. It comes into play when we need to change our basic acceptance property, in case of ignorance or error. Horwich notes that there may be whole communities who live in ignorance of the correct meaning of a term (based on basic acceptance property of an expert).

Let's consider further the nature of these statements which Horwich calls use properties. Meaning-entities are, for Horwich, use properties, and are universals; and to have a certain meaning is to exemplify a certain use-property. These meaning entities, or use properties are exemplified as follows:

> 'true' means what it does to us in virtue of our provisional underived acceptance of instances of the schema, '<p> is true ↔ p'; 'bache1or''s meaning is engendered by our underived acceptance of the sentence, 'The bachelors are the unmarried men'.
>
> 'red''s meaning stems from our underived propensity to accept 'That is red' in response to the sort of visual experience normally provoked by observing a clearly red surface.
>
> The meaning of 'water' is constituted by the fact that we accept, as basic, 'x is water ↔ x has the underlying nature of the stuff in ***our*** seas, rivers, lakes and rain'"
>
> (Horwich 2004:351)

It may be argued, however, following Hinzen (2007), that the meaning of the Equivalence Schema certainly does not arise simply from regarding its instances as true. On the contrary, it is our knowledge of grammar that entails that we will regard these instances as true. This seems to be a problem for the use theory, since it shows that there is an explanation that goes deeper than simply appealing to a brute decision to regard something as true. Horwich may reply that in stating that 'regarding the truth schema as true constitutes the meaning of truth', he is not claiming that truth itself is so constituted. It is only the meaning of the term that is constituted by the decision. So he may be willing to accept that our knowledge of grammar entails that, for instance, we can only apply truth to sentences at a specific level of grammatical complexity, but that is a grammatical issue. Horwich agrees that syntax or the combinatorial principles are a constraint on the meaning, but he may still argue that the meaning of individual words, or what notions we attach to which words, is a matter of what we are disposed to accept. To elaborate, we may distinguish between logical notions and the terms used for them, and the assignment of those particular logical relations to that particular term would still be a matter of

a brute decision. We could, for instance, have associated the equivalence schema ("<p> is true ↔ p"), which we now associate with 'truth', with say 'table' instead. But the brute decision to attach the equivalence schema to 'truth' (as opposed to 'table') gives it the meaning it now has. What is *not* a matter of a brute decision but rather a disposition is which particular meaning, from all the various uses of the term, we consider to be constitutive.

Further Schiffer (2000) notes that while it may be easier to see how Horwich's thesis regarding meaning-constituting properties applies to 'truth' or 'red', Horwich does not explain how it applies to words like 'dog': "'Dog' may mean DOG for someone who is blind or who does not know what a dog looks like, so it cannot be required that anyone who understands 'dog' must be disposed to accept 'That's a dog' when confronted with a paradigm dog" (Schiffer 2000: 534). For Gupta (2003) "an identification of principles that are fundamental to an explanation of the acceptance of sentences with principles that are fundamental to meaning", which forms the basis of Horwich's use theory of meaning, is "an unacceptable identification" (2003:654). He argues that the two are entirely different things, and that which is fundamental to meaning need not be fundamental to the acceptance of a particular sentence (2003: 664-65). "For, plainly, the acceptance of sentences depends not just on the meanings of words but also on the methods of obtaining information (and misinformation) about the world. Explanatorily basic patterns [of sentence acceptance] will, thus, typically reflect not just meanings but also methods of obtaining information." (Gupta 2003: 666)

Fodor and Lepore (2001)[39] point out that it is not the implicit nature of these use-properties, but rather their dispositional character, which precludes questions regarding how they are derived. They argue that whether a definition is accepted implicitly or explicitly, it would still be open to the same objections. What is important is that these definitions are the ones which we are *disposed* to accept. To claim that we have such a disposition automatically precludes questions regarding how these statements are *derived.* But this leaves out any interesting discussion regarding why it is that we are disposed to accepting these as opposed to any other statement containing the word.

[39] Fodor and Lepore (2001) argue that, although both distinctions apply to mental states, there is a difference between the explicit/implicit distinction and the distinction between occurrent/dispositional. "The former is epistemological; it's a matter of whether the creature that's in a state has (non-inferential) access to it's being there. In the simplest examples, implicit/explicit is about whether a creature is able to report being in the state that it's in. Whereas the second distinction is *ontological*; we suppose that *occurrents*, but not dispositions, are species of events."(2001: 357-358)

By way of a possible defense, Horwich (1998) argues that an implicit definition is far from an analytic truth. So saying that we are disposed to accept certain statements doesn't imply that these are necessarily true. It is consistent with a sentence being taken to be true, that it is in fact false, and that it is so regarded for empirical reasons. One may for example be wrong or ignorant about the meaning of a word, and what we consider to be true may thus be revealed as false at a later time. Further, not only are implicit definitions not all apriori, but it is likely that none are a priori, if we consider for instance Quine's 'web of belief'. (Horwich 1998: 143)

Another important element in the above description, which Horwich gives of his use theory (2004: 352), is the idea that these use-properties are 'non-semantic'. This term opens up the theory to the charge of behaviorism. We shall take this up again in section 3.4.4, where we will be in a better position to deal with this aspect of his theory. For now, let us point out objections to Horwich's use theory, based simply on his notion of word meanings noted above. Later, particularly in section 3.4, we will consider whether any of these objections can be neutralized through a consideration of other aspects of Horwich's theory. Horwich is a deflationist concerning both compositionality and truth. It should be useful to discuss deflationism here in order to better understand Horwich's use theory.

3.2.1 Objections to Horwich's Theory of Word Meanings as Use Properties

This account faces various problems:

> F1. It is not clear how one could systematically arrive at such paradigm uses; i.e. the non-semantic property mentioned in the beginning of section 3, which are accepted underived. It is not clear, in particular, whether these acceptance regularities indicate merely regularities in use or also concept regularities, i.e. whether the 'acceptance' indicates or is due to adhering to conventions or usage within shared conceptual specifications. For example, when a person uses a word, she may do so according to conventional usage, but she may have a completely different concept in mind, showing that merely accepting a certain usage shows little about what one understands a word to mean, and what the basis of the use is. This is particularly important when it comes to language acquisition in children. A child may for instance acquire the concept: APPLE = all red, round objects; though he continues to use it correctly in all instances (in this case, the only red and round objects he encounters are in fact only apples). The point is, in other words, Quinian: from the mere use of a word ('rabbit', for example)

we cannot know what concept is in the mind of the speaker (it could be a part of a rabbit, the whole rabbit, a state of the rabbit, etc.). Thus it needs to be made clear what these acceptance regularities refer to.

F2. What does the expression 'underived propensity to accept' really signify? Does 'underived' here mean innate?[40] Perhaps what is rather meant is that all our uses of the term are derived from this particular use, and thus it has the character of a definition. So we accept these sentences as implicit definitions of the terms. But why do we do so? Such a theory would inherit all the problems of a theory of description.[41]

F3. How are we to know that the meaning refers in one case to a law of logic (e.g. in the case of truth, the logical predicate is literally defined by the given sentence, or sentence-schema), and in another to a visual property (red), and in yet another to a constitutive natural property of an element (water). It seems that we need an independent account of how to sort which word into which kind of meaning property. Alternatively, it appears that in order to be able to formulate such meaning properties (as the paradigm use properties) we already need to know what the words mean, at least in a primitive sense.

F4. How could one account for non-paradigmatic uses of words at all within this theory? A use may differ from the paradigmatic one. So would a case of non-conformity to established paradigms of use simply be rejected? If that were the case, we would never be able to change or add to existing meanings of words at all. Further, even in cases where a person uses the term to mean what it is normally used to mean, this usage may not be captured by what Horwich defines as the use/meaning property of the term. The knowledge that the speaker possesses while using the term may not be of the basic use

[40] Horwich suggests in personal communication that there may be at least some element of innateness in meaning (of certain kinds of terms). However, for most part, at the time of learning a language, a child is exposed to various different uses of a term, and then at a certain later stage, decides or comes to recognize one of these several uses as basic, or the implicit definition of the word. So perhaps an interesting question to ask, if we are to accept this claim, would be: what is it that makes the child accept that particular use-sentence as the basic underived use-property of the term as opposed to other use-sentences. Horwich claims that it is empirically determined. And in making this decision, the speaker may rely on the relevant experts as well, as discussed in the previous section (see section 3.1 above).

[41] Russell (1918/1956) argued that our ordinary proper names are not really names, but rather abbreviated descriptions. To this Searle(1958) argued that a name, under Russell's theory, must be equivalent to a particular description; but a speaker does not have any single determinate description in mind while using a name. Searle argues rather for a cluster (vague and unspecified number) of descriptions. Kripke (1972/1980) argues that it is possible that all the descriptions associated with a name may be false, but the name doesn't lose its reference, even if all the descriptions turn out to be false.

property, and remain more ‘superficial’, while at the same time using the term correctly in every case. For instance, a speaker may not know that water, which he drinks from taps, comes from and has the ‘underlying nature’ of the stuff in seas, rivers etc. (for he doesn’t know that there are rivers etc. for instance). Moreover, Collins (2001) raises the concern that “while the meaning of nurse might well be some set of paradigm nurse features (including, presumably, being female), the meaning of male nurse need not be a paradigm at all.” What do paradigm nurses have to do with the meaning of male nurse? “It seems that Horwich is saying that one could understand ‘male nurse’ without having a clue about paradigm nurses, for what constitutes the meaning of the complex does not express what constitutes the meaning of its parts, the parts just need to have some meaning or other.” (Collins 2001:58-59) Collins goes on to argue that it is precisely compositionality which we want to explain, as we want to know what a sentence means in terms of its parts. We shall see in the next section whether this objection holds, given the different accounts of word and sentence meaning in Horwich’s theory. At this point we may note, however, that Horwich himself distinguishes his use theory from, and points out its advantages over, ‘The Prototype Theory’: “that the meaning property of a word consists in there being certain paradigm cases of its applicability and there being a ‘similarity metric’ which determines how close other things are to these exemplars.” (Horwich 1998: 52) The two main advantages of the use theory, he notes, are its generality and its explanatory power. Prototype theory cannot deal with meanings of all words, “Exemplars and similarity metrics seem out of place in connection with the theoretical vocabulary of science.” (Horwich 1998: 53)

We have seen in this section that Horwich has to deal with various issues within his conception of word meanings. We shall return to these problems in section 3.4, after considering other aspects of his theory in the next section, in particular his account of sentence meaning, with which his account of word meanings is combined. Our contention, thus far, is that his notion of word meanings cannot withstand criticism on its own. It needs bolstering from external sources, which should differ, at least, in some manner of explanation, from that accounting for *word* use. In the next section we will see whether his explanation for sentence meaning proves to be a solution.

3.3 Deflationism and Horwich

Deflationism about truth refers to "a family of theories about the meaning or expressive role of the locution 'is true', often attended by effacing metaphysical claims about the nature of truth itself" (Burgess 2012: 397). Deflationism contends in general that there is nothing mysterious about the notion of truth; it is much simpler than most philosophers believe. Within this purview there are those that regard this simplicity as a semantic fact: that, for example, the meaning of 'truth' can be captured by for example, Tarski's T-schema. And there are those that go further to claim that the meaning of truth so defined captures all that there is to capture about truth; there is nothing more to the notion of truth than what can be so defined.

Burgess (2012), for instance, notes that there are two main strands of Deflationism, semantic and metaphysical. The former "begins with an account of the meaning of 'true' and cognate expressions – and then goes on to explain why we would have such a 'transparent' word in our vocabulary." And "The secondary metaphysical strand of deflationism is roughly that, once we have explained what 'true' means, there is little more to say about what truth is. Unlike *being a sample of water* or *having commercial value*, truth is not a 'substantive' property susceptible to illuminating, naturalistic reduction. Differences in detail between competing versions of the view are overshadowed by a common commitment to the idea that the metaphysics and semantics of truth are exceedingly simple." (Burgess 2012: 398)

Gupta however focuses on the latter of these two strands and argues that deflationism "maintains that truth is a simple and clear concept and has no substantive role to play in philosophy." (Gupta 1993; 58) He contends that deflationism "holds that once we understand the meaning and function of 'true' – and this understanding, according to deflationism, is not hard to achieve – we shall see that truth has no substantial role to play in philosophy." (Gupta 1993: 59)

This is in direct contrast to the philosophical tradition that contends that truth-conditions play a central role in the theory of meaning. In the next chapter, we will take a look at a recent advocate of this tradition to see whether it can provide answers to the questions we began with. For now, we will try to see what deflationism means for Horwich. His deflationism is not limited to truth; he also deflates compositionality.

The Compositionality Constraint, Horwich points out, may be formulated as follows. "An adequate theory of meaning must enable us to see how the meanings of complex terms

may be determined, and thereby explained, by the meanings of their parts." (Horwich 1998:33) Horwich however provides a deflationary account of the above general constraint, which amounts to the following: "Understanding one of one's own complex expressions (non-idiomatically) is, by definition, nothing over and above understanding its parts and knowing how they are combined. In other words, once one has worked out how a certain sentence is constructed from primitive syntactic elements, and provided one knows the meanings of those elements, then, automatically and without further ado, one qualifies as understanding the sentence. No further work is required; no further process needs to be involved, leading from those initial conditions to the state of understanding the sentence." (1997: 504)

Thus the meaning of the sentence does not have to be *worked out* on the basis of what is known about how it is constructed; for that knowledge by itself constitutes the sentence's meaning. If this is so, then compositionality is a trivial consequence of what we mean by "understanding" in connection with complex expressions. It can put no constraint whatsoever upon how our *words* come to mean what they do—allowing, in particular, that the meaning of a word derives from its use. (Horwich 1997: 504)

The following claims are argued for here, beyond the compositionality constraint itself:

D1. Knowledge about how a sentence is constructed itself constitutes knowledge of the sentence's meaning. Compositionality (or how the sentence is constructed in terms of its parts) is a trivial consequence of what we mean by 'understanding' sentences (or knowing the meaning of a sentence). This is so because Horwich claims that knowing the meaning of a sentence (as opposed to that of a word) is nothing over and above knowing how the sentence is constructed and what the individual words mean. In other words, the construction property itself constitutes the meaning of a sentence, along with word meanings.

D2. Compositionality, or how a sentence is constructed, does not put any constraint on what constitutes word meanings, or how these primitives come to mean what they do.

It appears from this description that what is central to his project is an account of how a sentence is constructed. However, it is precisely regarding this that Horwich has no significant explanations. This gap not only prevents the theory from being explanatorily complete, but also from being anything close to a systematic scientific analysis, as opposed to philosophical speculation. But this is not necessarily a drawback, considering the philosophical

aims of the theory. As we will notice later, Horwich is not aiming for a rigorous scientific theory at all, but the ordinary non-scientific understanding: "the aim of the present use theory is to give an account of meaning in the ordinary, non-scientific sense of the word… The purpose of the account is not scientific explanation, but rather a demystification of the ordinary concept of meaning, followed by the philosophically beneficial consequences of that demystification: namely, solutions to the numerous problems… that are produced, …, by confusion about that concept. (Horwich 1998: 87) Horwich also points out that although his theory may not be rigorous enough to be termed scientific, it may nonetheless serve to "provide a helpful clue" regarding the kind of meaning needed for a scientific theory. Such a scientific theory, for him, is Chomskian linguistics, and he believes that his regularities-of-use conception of meaning can inform this theory by essentially replacing I-meanings. We will return to this point in section 3.4.3. To elaborate on points D1 and D2 above, we will first consider the nature of sentence meaning and then go on to his deflation of compositionality, before we proceed to discuss further objections and counters.

Using his convention of capitalising words when using them to refer to meanings, he argues: the meaning property 'x means DOGS BARK' consists in the construction property: 'x results from putting terms whose meanings are DOG and BARK, in that order, into a schema whose meaning is Ns V'. So a complex like 'dogs bark' gets its meaning from the following:

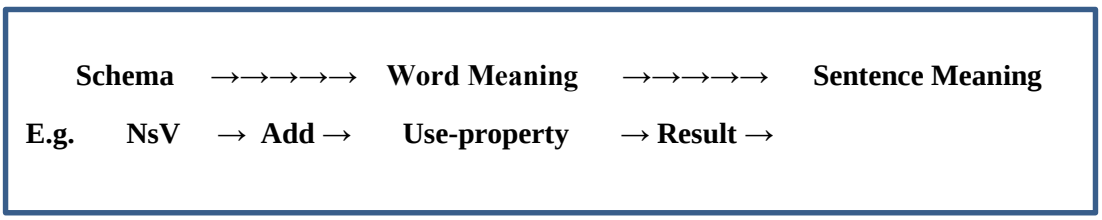
E1. The meanings of the two words 'dog' and 'bark'

E2. The meaning of the generalization schema 'ns v'

E3. The fact that the sentence results from placing those words in that schema in a certain order.

Thus 'dogs bark' would result from putting words meaning what 'dog' and 'bark' mean into a schema meaning what 'ns v' means. See Figure 3 below.

Figure 3: Horwich's Use theory



We may here immediately notice that he draws a parallel between the distinction of words and word-meanings on the one hand and schemas and schema-meaning on the other. It is however not immediately obvious how the similarity is to be understood. What is the meaning of a schema? How are we to distinguish it from the schema itself? If the parallel with word meanings follows then the meaning-property of a schema must be arrived at in a similar way as the meaning property of words. I don't think this can be, since the construction accounts for the meaning in the one case, but not the other. But how are we to come up with acceptance conditions or paradigmatic uses in the case of schemas? Perhaps it is not his intention to draw such a parallel at all. Thus alternatively, he could be interpreted as simply suggesting that schema meanings are the way that words are put together in a particular schema. The schema could be determining what is the function and what is the argument, and that the former is to be applied to the latter. So in the case of 'dogs bark', for instance, a schema meaning Ns V, when applied to an ordered pair of word meanings (<DOG, BARK>) gives the meaning of the complex. However, Horwich does not provide any principle for the generation of schemata, nor a way of identifying the primitives which may constitute them. The only cue we have is the 'ns v' structure he provides. We may note here nonetheless that this need not adversely affect his position, as he does not claim to be providing details in any case. His position is a philosophical one, and he is happy to leave the scientific particulars for the linguists for instance.

We can see how generic his theory is through the following.. He argues that: for any complex expression e constructed by combining certain primitive terms <w1,…,wn>, e is the result of applying combinatorial procedure P to the primitives <w1,…,wn>, the meaning property of e: 'x means E' is constituted by the construction property: 'x results from applying combinatorial procedure P to primitives whose meanings are <W1,…,Wn> (Horwich 1997: 507)

We can notice that he does not specify any of the constituents here: for instance, what kinds of things constitute the primitives, or what is the combinatorial procedure. This strategy, he points out, is deflationary: "since our explanation did not involve any assumptions about how the meaning properties of the primitives are constituted, it would seem that compositionality per se provides absolutely no constraint upon, or insight into, the underlying nature of meaning." (Horwich 1997:507)

Fodor and Lepore (2001) argue that although it may be possible to abstract from a thick substantive account of compositionality, which comes with further assumptions about reverse compositionality, systematicity etc., in a way that puts no constraint on the nature of word meanings, doing so would only leave a trivial notion of compositionality. "But it's also true that it doesn't matter much that compositionality doesn't matter much lacking such further assumptions. What matters is that there appears to be a plethora of truths about the semantics of complex expressions that the assumption of compositionality, together with a good theory of the lexicon, explains; and that, as far as anybody knows, can't be explained if the assumption of compositionality is left out." (Fodor and Lepore, 2001: 367) they argue that compositionality comes with assumptions like for instance that language is systematic, and if we can get a meaningful sentence from combining words together in accordance with some principle, we must also be able to systematically and meaningfully break apart a sentence into its constituents. So, "The meanings of 'dogs' and 'bark' must be contained in the meaning of 'dogs bark' because people who understand the sentence likewise understand the words." (Fodor and Lepore 2001: 366). Reverse compositionality, or the theory that the meaning of a complex expression determines the structure of the expression and the meaning of its constituents, may in itself be too strong, but intuitively there must be at least some meaning that is preserved. For instance, we may not know the meaning of a particular word in a sentence, but given that we know the language and the other words, we would be able to say something about the meaning of that word. Take the sentence: 'She sat on the togo'. Now we can tell, given the structure, that 'togo' should be a noun and given the meanings of the other words, that it is something that it is possible for someone to sit on. Even if we don't know the meanings of all the other words, the structure itself is capable of providing us with at least some formal elements of the meaning of the word. The idea is essentially that, although it is possible to deflate compositionality, in doing so we lose a lot more, including explanatory capacity, systematicity and richness of language.

It may be useful here, in order to understand Horwich's deflationary stance, to contrast it with truth conditional theories such as Davidson's, for which compositionality is not non-substantive in the above sense.

3.3.1 Davidson on Deflating Compositionality

Davidson's programme appears to be in direct conflict with what we have discussed so far regarding Horwich's theory. For Davidson, "knowledge of the structural characteristics that make for meaningfulness (=syntax) in a sentence, plus knowledge of the meanings of the ultimate parts, does not add up to knowledge of what a sentence means" (Davidson 1967: 307) This, he argues, is because such knowledge cannot account for propositional attitudes, e.g. belief sentences. This leads Davidson to conceive a theory of meaning as a theory of truth: "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… To know the semantic concept of truth for a language is to know what it is for a sentence - any sentence - to be true, and this amounts, in one good sense we can give to the phrase, to understanding the language." (1967: 310) Thus he imposes the Tarskian "s is true iff p" (where s is the sentence and p the proposition, for instance, '"Snow is white" is true iff snow is white') as a fundamental constraint on a systematic theory of meaning.

Horwich considers two problems faced by Davidson's approach. The first of these relates to how his account could be applied to all sentences (for example those containing adverbs, that-clauses, attributive adjectives). The other problem relates to whether the truth condition of a sentence does in fact suffice to determine its meaning. He certainly does see other problems, e.g. that a naturalistic semantics cannot be based on truth, and that meaning is prior to truth. Horwich argues that these problems all have a trivial solution based on the principle of compositionality (that the meaning of a complex is determined by the meaning of its elements and by its syntactic structure) but the solution involves no explication of meaning (e.g. in terms of truth conditions, or use properties, etc.) and hence offers no explanation of *why* the principle of compositionality holds. He argues that we reach an interpretation of 'Gira marte' through the following steps:

'Marte' in Italian means the same as *our* 'Mars'
'gira_' in Italian means the same as *our* '_rotates'

Then, from the principle of compositionality, we can infer:
The result of applying 'gira_' to 'Marte' in Italian means the same as
The result of *our* applying '_rotates' to 'Mars'

And finally, given the syntactic facts,
The result of applying ‘gira_’ to ‘Marte’ = ‘Gira marte’
The result of applying ‘_rotates’ to ‘Mars’ = ‘Mars rotates’

we can deduce the interpretation:
‘Gira marte’ in Italian means the same as *our* ‘Mars rotates’
(Horwich 2001: 370-71)

Davidson had argued against such a deflationary account on the grounds that such a manual of translation does not convey meanings, and thus cannot amount to an interpretation. One can be told that two expressions have the same meaning, and thus are inter-translatable, without having any understanding of the sentences. Davidson was seeking to address the problems faced by a radical interpreter, one without any prior knowledge of a speakers belief or meanings of his utterances. Thus for him, interpretation requires an account of both belief and meaning, and he argues that the way to achieve this is through the principle of charity, by interpreting the speakers, wherever possible, as holding true beliefs. So the elements to be accounted for within this wholistic interpretation, as consistent with one another, include at least the speaker’s beliefs and meanings, his behavior and his environment.

Horwich, however, sets himself a much more restricted task. He is not seeking to give a holistic account of meaning faced with the problem of the radical interpreter. Rather, he admits that “a translation manual can supply interpretations only relative to our understanding of our own language”, but he argues further that “the same can be said of any alternative view one might have of which explicit assumptions would work.” (Horwich 2001:374) So he is not trying to account for interpretation, but rather confines himself to explaining the meaning of a sentence given our shared beliefs, knowledge and environment.

Within this restricted context, Horwich argues that sentence meaning *is* compositional: so the meanings of constituents along with the combinatorial principle do add up to the meaning of the whole sentence. But the reason why this avoids the problems raised by Davidson is that on Horwich’s account compositionality is deflated, i.e. it puts no constraint on the nature of meanings of constituents; and meaning is not truth-theoretic. Given his account of sentence meaning and compositionality considered above, Horwich argues that the problem that led Davidson into characterizing meaning in terms of truth would not arise. However, as we just noted, Davidson and Horwich are trying to address different problems to begin with. Thus

Davidson's objections to a Horwich-like account may not directly address the context of his theory. We will now go on to consider other objections to this view.

3.3.2 Objections to Deflating Compositionality

Pietroski (2005), on the other hand, argues that we should place the following substantive constraints on a theory of meaning: (i) a theory of meaning is a theory of understanding, (ii) which may be of the form of an algorithm associating signals with interpretations, (iii) and which should explain how a speaker of a language L associates signals of L with interpretations. He claims that if phenomena like ambiguity, entailment, indexicality etc. bear on (iii), then they also bear on theories of meaning. (2005:255-256) Such phenomena, however, remain unexplained within Horwich's proposal. Pietroski points out that Horwich cannot explain the various entailment relationships that exist between natural language sentences. For instance, 'If every dog barked, and Fido is a dog, then Fido barked' or 'If Pat boiled the soup, the soup boiled'. He also cannot explain why, for instance, we can infer (ii) from (i),

i. John ate an apple
ii. John ate

Unless the meaning is compositionally determined, and compositionality is non-trivial (i.e. compositionality is understood in a non-deflationary sense), how can the truth of (i) guarantee the truth of (ii)? Further, if speakers don't understand 1 and 2 compositionally, how do they know that we can infer from (i) to (ii)? (Pietroski 2000)

Collins (2003:415) equally points out that even if a Horwich-style schema does capture the meaning relevant structures of a sentence, it does not tell us how the sentence relates to another that means the same. To take a simple example, the relation between the active/passive forms in (iii) and (iv) remains unexplained.

iii. Bob kicked the ball
iv. The ball was kicked by Bob

It can of course be argued that the two expressions here are not necessarily synonymous, as may be understood by the different questions that they may answer to. For instance one may utter (iii) or (iv) in response to 'Who kicked the ball?', but in response to for instance, 'What happened to the ball?' it seems that (iv) would be more appropriate than (iii). This is because the subject position is occupied by 'Bob' in (iii) but in (iv) 'the ball' is the subject. Other cases of synonymy also appear to be subject to a similar argument. Synonymy appears to be rare in human languages. This, it has been argued (Carstairs-McCarthy, 1999), may be due to the

cognitive principle of "synonymy avoidance", which "demands that all contrasts in form imply a contrast in meaning."(Hinzen 2006: 136). Thus we can argue in favour of Horwich that it is difficult to find sentences that in fact mean the same, and the meanings of the sentences (iii) and (iv) can be explained by their respective construction properties. It is indeed true that anyone who understands the two sentences would in fact also recognize that these are closely related, and there is a thematic core of meaning of these two sentences which is uniform. However, the way that Horwich can account for this is to say that it is a matter of how the schemata works in each case.

Pietroski further points out that it also needs to be explained, for example, why we cannot infer from (v) to (vi):

v. John is too clever to catch a fish

vi. John is too clever to catch

although we can infer from (vii) to (viii)

vii. John ate an apple.

viii. John ate.

We know that (v) means that John is too clever for him to catch a fish, while (vi) means that John is too clever for anyone to catch him. Why is it that we don't interpret (vi) on analogy with (viii): assuming that since (viii) means that John ate something, (vi) could mean that John is too clever to catch something? He points out that non-ambiguity also needs to be explained; we need to explain the negative fact of why, for instance, 'John is easy to please' can't mean that 'It is easy for John to please us' or 'The millionaire called the senator from Texas' mean that 'The millionaire called the senator, and the millionaire is from Texas'.

We may again note here, however, that Horwich's position is not necessarily affected by examples such as (iii)-(vi). The reason for this is essentially the lack of detail provided in Horwich's argument. There is nothing in his writings that *precludes* a more detailed account of sentence construction. In fact he leaves the details to be worked out by linguists. His point is a philosophical one: that we get sentence meaning through word meanings (whatever we take word meanings to be) and syntax (or schemata, or combinatorial procedure P: whatever such a procedure happens to be). As we saw earlier, he does specify what he takes to be word meanings (though his theory of sentence meaning is consistent with anything else we may understand by word meanings) but he doesn't specify the kind of schemata or the details of the combinatorial procedure.

Thus, consider the contrast between (ix) and (x). Pietroski points out that ‘eager’ and ‘easy’ differ semantically in an important sense, which cannot be explained simply by saying that ‘eager’ means EAGER and easy means EASY. This, however, may be a criticism too harsh for the view Horwich is advocating. For Horwich, the meaning properties of ‘eager’ and ‘easy’, i.e. EAGER and EASY, would be the use properties of these terms. This plus the construction properties of the sentences into which these words enter would perhaps explain the semantic difference between the two, notwithstanding all other problems with this theory.

(ix) John is eager to please.

(x) John is easy to please.

We will consider other criticisms of Horwich again in the next section. At this point we may conclude provisionally that Horwich’s account may not be affected by the arguments levelled against it, simply by virtue of the lack of details provided in it and his actual goals. His position can be strengthened by just filling in the required details and there is nothing in his account that prevents such strengthening. His focus has been mainly to make a philosophical point and not to provide detailed grammatical explanations.

3.4 A Defence of Horwich

So far, we have acquired the general picture of Horwich’s theory. In Section 3.2 we discussed Horwich’s position that meanings of words are their use properties. We noted the kind of entities Horwich calls use properties, and pointed out various objections to his position. Then in section 3.3 we considered his claim that sentence meanings are arrived at trivially from word meanings when we assume compositionality. Again we pointed out objections to this claim and tried to respond on Horwich’s behalf. In this section, we shall try to build a defense for Horwich by first taking a closer look at what he means by use-properties, and revisiting the objections raised in section 3.2.1 Then we will consider his claims regarding sentence meaning and see whether we can strengthen the response to objections raised in section 3.3.2.

3.4.1 Use-Properties: Implicit Definition and Rule following

As we noted in section 3.2, Horwich argues that word meanings are use properties. By use properties he means law-like regularities of use. These law-like regularities are not explicit rules we are aware of, but rather implicit ones, as noted in section 3.2. To call them implicit doesn’t amount to saying that we unconsciously follow them either (as we could be unconsciously following explicit rules): rather “no formulation of the rule is directly

operative." (Horwich 2010b: 115) So here the following problem arises: since there is no formulation of the rule, it is difficult to see what the facts might be in virtue of which someone would be implicitly following a given rule. We don't see what would count as a violation of the rule. This problem is not specific to language: there are rules that have nothing to do with word meanings (e.g. the amount of distance to maintain while talking to someone). Even within language – there are various kinds of rules: meaning determining rules of use, those instructing us to conform our meanings with the community, rules of epistemic justification, rules dictating that we try to accept only those sentences that are true. (Horwich 2010b)

Horwich suggests that a person S implicitly follows a rule R if and only if

G1. S's activity is governed by the ideal law R

G2. There is some tendency for S to correct instances of non-conformity (i.e. to react against his initial inclinations)

He notes (Horwich 2010a) that it's far from obvious that each of us knows, independently of any sophisticated empirical research, which rules we are implicitly following. The most we can be expected to detect through introspection are occasional inclinations and disinclinations to do this or that particular thing. The underlying rule is available to us only via objective methods of inquiry that are equally open to everyone else. Thus the isolation of the explanatorily basic rule of use is a matter of empirical enquiry. So objection F1 and F4 above are taken care of: as he argues that the meaning constituting property is one which explains these acceptance phenomena – the words' basic rule of use. But it is an empirical discovery, often quite hard to make, which particular rule is the explanatorily basic one for the use of any given word. The second part of objection 1 is directly answered as he claims that 'meanings=concepts=properties' (Horwich 1998:4).

The question raised in F2 (what does 'underived propensity to accept' mean, or what does Horwich mean when he says that these use-properties are accepted underived) is diluted when we take into account his distinction between explicit and implicit definitions as noted above. So these use-properties do have the nature of a definition, but they are implicit definitions rather than explicit ones. In case of explicit definitions, we can raise the question, how do we come to acquire or accept that particular definition, but in the case of implicit definitions, this question does not arise, as it is a matter of empirical inquiry which sentence (or rule) we accept as the implicit definition. It may be argued that empirical inquiry will then ask how we acquire this sort of behaviour, so that the question seems to arise all over again. In other words, we may ask: how is it that we come to accept an implicit definition? How far can

this question be answered by claiming that it is a matter of empirical inquiry? It may be a matter of empirical inquiry which particular sentence we come to accept as the implicit definition, but saying that still does not explain why we come to accept it at all. As we will note later, saying that we do in fact accept certain sentences underived does not explain why we do so; and it is such behavior that we need to explain in case we want an explanatorily adequate theory. Further, he notes that if rule following is involved in the constitution of meaning, then it is plausibly always a matter of implicit rule following. This avoids the regress that would result if we accept that some terms are a case of explicit rule following: we would not know how to account for which are basic. Regarding objection F3, he argues that "one should not expect to be able to derive a word's extension directly from its meaning constituting property… Because of the plausibility of the deflationary view of the truth theoretic notions, 'true', 'refers' and 'true of'." (Horwich 2010: 131) Horwich does not deny that a word's meaning constituting property fixes its extension but argues that it does so only because it first fixes the word's meaning. In F3, we had raised the point that we need an account of how to sort which word into which *kind* of meaning property, in order to fix its reference or extension. Here, however, we are pointing out with Horwich, that the way the use property fixes a word's extension is by first fixing its meaning. Whatever the *kind* of extension a word has, whether it is a law of logic, a visual property etc., we obtain such information through the word's meaning. Meaning, thus, comes before reference. It may be argued however, that for Horwich meaning is inherently dependent upon what uses of expressions people accept and what they regard as true. So if we do not have a notion of meaning independently of this, then isn't truth or extension fixing the meaning? But as Horwich points out, the difference between what people accept as true and truth itself is important to note here. We need to distinguish a disposition (to accept a certain statement) from an occurrence (Fodor and Lepore 2001) or acceptance from truth. To say for instance that we have a disposition to accept the equivalence schema is not to say that the equivalence schema is true. However, simply to stop at the claim that we have a certain disposition is to ignore interesting and perhaps significant questions as to why that is so.

Another objection we mentioned in F4 raised by Collins (2001), regarding examples such as *male nurse*, can now be analyzed in light of Horwich's whole theory. If we keep in mind the difference in the way in which we arrive at word meanings and sentence meanings, examples like this would not be a problem. This is because male nurse is a compositional structure, the meaning of which would be determined according to the particular schemata as well as the meaning of the individual words. The problem with male nurse was that it does not

fall within the paradigm associated with either ‘nurse’ or ‘male’, but as we already noted, Horwich distances himself from the prototype theory. Further, in light of how Horwich explains sentence meaning, it is clear that the meaning of the compound would be determined in accordance with the relevant schema, and the meanings of the words (whether use properties, paradigms, referents or whatever else) can then be inserted. There would, in this case, be no problem as ‘male nurse’ need not have a paradigm, and the meaning of the compound would simply be compositionally determined. But the problem for Collins was that here (and in Horwich’s theory in general) what constitutes the meaning of the complex does not express what constitutes the meaning of its parts, and thus we do not get the meaning of the sentence in terms of its parts. For Collins, it is compositionality that needs explanation. In the next sub-section, 5.2, we will consider in greater detail Horwich’s stance on compositionality.

3.4.2 Sentences: Entailment, Compositionality and Truth Theoretic Accounts

Now we may look at the objections mentioned in section 3.3.2. Horwich (2008) argues against the idea that semantics needs to account for the following:

> H1. Conceptual relations among sentences (such as entailment. For instance that we need to know how examples (iii) and (iv) above are related, and also how (v) and (vi) are not.)
>
> H2. That the meanings of sentences (as characterized theoretically) is explained on the basis of the theoretically-characterized meanings of their parts and how those parts are combined (i.e. Collins and Pietroski’s claim that compositionality needs to be explained.)
>
> H3. The fact that a sentence means what it does is identified with a truth-theoretic fact of some kind.

Regarding (H1) he argues that “we cannot account for facts of entailment by reference to truth-theoretic properties of the expressions involved. For the explanatory level of these semantically characterized properties is no deeper (but is in fact somewhat more superficial) than the level of the phenomena they are being invoked to explain.” (Horwich 2010: 147) Secondly, “it’s highly questionable whether the determinants of what our logical terms mean coincide (even remotely) with the logical facts that they enable us to articulate. Our meaning what we do by ‘‘neutrino’’ is no guarantee of the truth of neutrino theory...And similarly our meaning what we do by ‘and’, ‘or’, ‘not’, etc. does not fix what the logical facts are” (Horwich

2010: 148). So it is also questionable whether and how the formalization in truth-theoretic terms is related to our natural language sentences.

Regarding (H2) he argues that compositionality is not so difficult to accommodate: by identifying the meaning-constitutive property of a compositional sentence with its construction property. The consequence of this view is that: (a) compositionality would not require a substantive explanation as it will be trivially easy to deduce sentence meaning from its structure and the particular word meanings; (b) compositionality won't constrain the nature of word meanings. Further, Horwich's account can be strengthened by adding a detailed account of how the syntax determines meaning, when we already possess a lexicon (word meaning). We shall consider such a position in the next section.

Regarding (H3) above: that semantics must be truth theoretic, Horwich argues, in line with other deflationary accounts, that meaning is more fundamental than truth: "the basic rule we follow in our use of "true" is to apply it to a statement, s, when we take s to have the same content (i.e. meaning) as something we are already disposed to assert… But it presupposes that we be able to recognize 'sameness of meaning'. Thus the concept, MEANING, is more fundamental that the concept, TRUTH. We must therefore renounce the orthodox assumption that sentence meanings are to be analyzed in terms of truth conditions." (Horwich 2010: 17)

Rather for Horwich, the facts that semantics is meant to explain are facts concerning circumstances in which sentences are accepted. Facts concerning inferences are included in such acceptance facts. And it is not clear how truth theoretic semantics could account for them: how do you move from "the fact that S1's truth conditions are amongst S2's to the fact that someone is disposed to infer S2 from S1?"(Horwich 2010: 153) unless you add the thesis that any speaker of the language knows these facts. But this would involve infinite regress as the explanatory inferences will have to be explained in terms of knowledge of the meanings of the language in which they are conducted (Mentalese[42]).

A move to avoid the regress is to claim that our knowledge of referents and truth conditions is implicit. But then how would we account for such knowledge? He argues that the central truth theoretic properties and relations - truth, being true of, reference – are reducible to acceptance phenomena at the non-semantic level. We could for instance suppose, he argues, "roughly speaking—that the correlation between English and Mentalese terms becomes

[42] The term Mentalese refers to the mental language in which we think, or in which thought is represented, and processed, in the brain, i.e. Fodor's Language of Thought (1975).

hardwired and feeds into an automatic translation procedure—a mechanism whereby, when a Mentalese sentence is in the belief box, then an English sentence (one possessing the same structure, but imposed instead on the correlated English terms) is caused to go into the ‘‘To say’’ box?”(Horwich 2010: 157)

In fact he explicitly mentions (1998: 86-87) that he regards Chomsky’s contention that meaning may have to do with internal states of the brain as ‘quite plausible’.[43] However, the internal states he suggests meanings are related to, are dispositions. He argues that given that what are to be explained are phenomena of sentence-acceptance, the primary explaining facts (namely, the meanings of words) should also be facts of acceptance. For we have a model—namely, inference—of how a person’s acceptance of some sentences can cause his acceptance of others. Thus what needs to be explained on this view is not that certain inferences are valid or rational, but rather that certain inferences are made (or would be made).

We may here briefly consider Horwich’s disagreement with Chomsky’s picture. If Horwich’s account can be made to reconcile with Chomsky, it would certainly gain explanatory power, but would it still remain a use theory of meaning in any technical sense?

3.4.3 Horwich and the Chomskian Programme

Horwich makes a claim that sets his enterprise apart from Chomsky’s. To understand this, we need to elaborate on the outline of Chomsky’s picture provided in section 1.4. Horwich (2003) argues for what he calls a ‘very simple picture’ of the language faculty’, one which he claims is a simpler, more minimalist account than Chomsky’s picture. He claims that he shares the following assumptions with Chomsky: (a) that each human being has a faculty of language FL as a component of his mind/brain; (b) the possible states of FL, L1, L2, L3… are possible I-languages; (c) each state L is a computational procedure that generates infinitely many expressions E1, E2…; (d) each expression E is a pairing of a phonetic and a semantic object <PHON(E), SEM(E)> which through their respective interaction with the perceptual/articulatory (P/A) and the conceptual intentional system (C/I) determine an association of a sound with a thought; (e) these PHON-SEM pairs are constructed from lexical items, LI1, LI2,.., and (f) these lexical items are stored in a lexicon which is accessed by the computational procedures that form I-expressions. (Horwich 2003: 165)

[43] In the next section we consider Horwich’s disagreement with Chomsky.

He suggests that lexical items should be understood as pairings of I-sounds and formal properties which indicate the 'functional type' to which the lexical item belongs. So in his picture, there are such lexical items,<I-sound, formal properties>. These are combined according to certain schemata, combinatorial procedures that specify the type and order of constructions. The expressions so computed is a pairing of a phonetic and a semantic object <PHON(E), SEM(E)> , which through their respective interaction with the perceptual/articulatory (P/A) and the conceptual intentional system (C/I) determine an association of a sound with a thought.

This is his major distinction from Chomsky's picture, as for Chomsky, "lexical items consist not merely of I-sounds (and their formal properties), but of I-sounds paired with I-meanings (and their formal properties); and (b) that the input to and output of the C/I system – namely SEM(E) is constructed not from I-sounds but from I-meanings" (*Ibid,* 170)

So what Horwich appears to be changing about Chomsky's picture is the I-meanings, which he claims are inessential, 'an extra layer of complexity' (*Ibid*, 172). He understands I-meanings to be 'conceptual roles' which are explicitly represented in FL. He argues however, that all that is required is that "when the sound appears C/I is already in the state required to understand it."(*Ibid,* 173) We may ask: what would facilitate C/I's being in such a state? Horwich suggests that C/I would be in such a state if the I-meanings were *composed* of their (implicit) definitions: where "I-meanings take the relatively liberal, flexible form of the specification of a regularity of use, rather than the relatively constrained form of an orthodox explicit definition." (*Ibid,* 171), for instance, if "the conceptual role of a mental tem "f" consists in the underived acceptance within C/I of certain sentences, "#f" containing it." (*Ibid*, 172). As he clarifies, "the idea is that the meanings of I-sounds are constituted by such regularities, which are exemplified in C/I." (*Ibid,* 177)

Here Horwich appears to be modifying the nature of I-meanings, in line with his understanding of word meanings which we discussed earlier, rather than getting rid of I-meanings completely. As Chomsky points out in his reply to Horwich (2003), and I tend to agree, it is a matter of choosing 'regularity of use' as opposed to semantically interpretable formal features, as constituting I-meanings. Chomsky points out two problems with the idea of regularity of use: firstly that I-sound to which regularities are paired, and PHON(E) are distinct, with a one-many or many to many relation. "Thus, semantic properties of *persuade* appear in "persuasive" and "persuasion," though I-sound has different images in PHON(E)… In so far

as we speak of "regularity of use," it has to be assigned to the underlying element that has these manifestations in PHON(E): its lexical item." (Chomsky, 2003: 300)

A further and greater problem is that "regularity" cannot explain or contribute to "acquisition of semantic or other aspects of language". As poverty of stimulus arguments suggest, the acquisition of word meanings is very rapid and proceeds under highly ambiguous conditions. In other words, children are able to develop their language skills almost automatically, with very little exposure to language. In fact, children exposed solely to pidgin or deficient grammatical forms of language are able to develop a full grammatically rich language, creole, within a single generation (Pinker 1994). Further, experience is shaped by internal aspects specific to the organism. To put it differently, our biology constrains our various capacities: human beings normally can, for instance, hear sounds only between 16 Hertz to 20,000 Hertz, though dogs would have different experience of sounds as their range of hearing is normally 40 hertz to 60,000 hertz. Similarly, in the case of other capacities, our experience is constrained by our specific organic makeup as human beings. It would thus follow that the regularities we perceive are also determined by our inner aspects. In fact, often, we choose properties that lack regularities, even with overwhelming regularities in experience, such as in the case of children exposed to pidgin developing creole for instance. So a single experience of a bicycle being ridden, for example, may lead us to conclude that bicycles are for riding, though we may have mostly seen the bicycle being used as an umbrella holder on multiple prior occasions. And we choose to ignore all other factors common to that experience: the person, the place, the weather etc. (Chomsky 2003: 300-301)

Aside from problems with the notion of regularities of use, I-meanings, or rather semantically interpretable formal features, are salient. It is these features which determine the syntactic frames in which a lexical item can appear, and help distinguish between lexical items that appear in similar frames, for example that 'persuade' has a causative feature . It is these features that help us for instance to distinguish between 'persuade' and 'expect'.

Thus Horwich's argument, that we can give up I-meanings in favour of regularities of use, does not seem to hold its ground. Horwich accepts the rest of the Chomskian picture, apart from the introduction of regularities of use. 'Regularities of use' is a notion central to his use theory. If we give it up, then what is left of his picture can easily be reconciled with Chomsky, and in fact would not be very different from Chomsky's position. As the above discussion suggests, Horwich only added regularities of use to the existing Chomskian framework of I-

sounds and (semantically interpretable) formal features (or I-meanings). So, if we abandon the addition of regularities of use, the theory that remains is identical to that of Chomsky's and is no longer a 'use theory of meaning'. Chomsky, however, regards himself as a general use theorist, so what is left of Horwich's theory may be a use theory in the general sense in which biolinguistics is also a use theory of meaning: through shedding light on what the ability to use language is. But what remains is no longer a 'use theory' in the sense of explaining meaning through systematizing use. Here the distinction to note is that between the *ability* to use language, i.e. our language faculty, and the use of that faculty, i.e. language use itself. This is on the lines of the competence-performance distinction that Chomsky stresses. And there is a further distinction between the general use of language and the specific use itself which goes towards explaining the meaning, within a use theory of meaning. This distinction is one akin to that between performance and what the performance signifies. Here following Wittgenstein, we may claim that you cannot say anything about the meaning of a word except to show it through use: in other words, Wittgenstein went back to performance in order to explain its significance. For Chomsky the answers lie in the ability to use language, in the study of the faculty of language itself. We thought, to begin with, that Horwich was adding to this picture by systematizing use, or the actual performance itself. However, we now find that what can be saved from his theory, while avoiding contradictions, is Chomskian and not use-theoretic. This suggests that Horwich's additions to the standard Minimalist picture are vacuous, and the use theory doesn't add anything. We shall revisit this problem in section 6 below.[44]

3.4.4 Behaviorism

Let us now come back to the charge of behaviorism, an additional objection to his theory (mentioned earlier at the end of section 3.2): whether considering the meaning/use property as a 'non-semantic property', commits him to a kind of behaviourism. Horwich has been accused, for instance by Boghossian (1993), of behaviorism, by being attributed with the following premise: 'the meanings attributed in folk psychology are purely behaviorally constituted properties, properties whose attribution presupposes nothing about the character and nature of the mind/brain that underlies linguistic activity.' (1993:140)

[44] We may note however, that the problem of systematizing use or focusing on the actual performance in the above competence/performance/significance distinction may lie within this characterization of use as performance, i.e. within the fleeting nature of use itself. The use aspect only lasts as long as utterance, the moment we start to analyze the sentence or how we understand it, we move towards either significance or competence. We shall take this argument up again in a different chapter.

Boghossian gives two main arguments against behaviorism:

> I1. We could construe meaning properties mentalistically and still give a satisfactory account of the epistemology of other minds.
>
> I2. there is no way to reconstruct what I now mean by an expression from my actual behavior with it right now – that behavior is consistent with my following an infinite number of rules.

However, the attribution of behaviourism to Horwich seems rather inaccurate. As we noted briefly above, Horwich would not object to a Chomsky-like internalism, and the only reason why he does not talk in these terms is that he wants to confine himself to 'ordinary' speech. Infact, Horwich sees a way of simply introducing or adding use properties to the Chomskian picture, notwithstanding the various problems with such a move, also noted above. But, given the various problems with combining Chomskian internalism with Horwich, we need not even invoke Chomskian internalism here. Meaning properties are for Horwich, accepted underived, although they are in no way a priori. As noted in section 3.2, Horwich (1998) gives a dispositional account of these properties. Thus the acceptance is not simply a behavior, but is rooted in a *disposition* for such acceptance. Regarding objection (I2), as we noted above, he argues that isolating the meaning constituting property of the expression is a matter of empirical enquiry. In any case, within Horwich's theory, it is not the actual behavior with an expression at any point in time, but rather an 'underived acceptance' of a certain use property which constitutes the meaning of the term. So while the use properties are not simply a matter of behavior and are rooted in deeper dispositions, they are at the same time empirically determined.

We saw in this section that once we consider all the aspects of Horwich's theory, we can overcome many of the objections raised in previous sections. Moreover, supplementing Horwich's account with biolinguistics can help his theory sidestep objections as well as make it stronger and more explanatory. However, it compromises his use-theoretic stance. Thus far, we have tried to present the strongest version of Horwich's theory, which can withstand many objections. In the next section we will attempt to put forth an objection to his theory.

3.5 Formulating a Stronger Objection to Horwich

We have noticed that Horwich overcomes many of the objections to his theory through his account of the distinct ways in which we arrive at word meanings and sentence meanings. So for him, word meanings are arrived at through use conditions, and to arrive at sentence

meanings all we need are word meanings and the principle of compositionality. He argues, as we saw, that this is compatible with any particular kinds of meanings we give to words: whether we take word meanings to be concepts, referents, use properties, or whatever else. Thus from word meanings, assuming a combinatorial principle, we can trivially deduce sentence meanings.

Further the kind of meanings that words have amounts to use properties of the following kind:

The word ""bache1or"'s meaning is engendered by our underived acceptance of the sentence: "The bachelors are the unmarried men"" (Horwich 2004:351). Now let's try and understand this a bit further. The meaning of "bachelor" comes from our acceptance of a particular sentence. The definition is special for the following reasons:

J1. It is in some sense generally accepted.

J2. It is accepted underived.

So this is an improvement from the traditional definitional accounts as it deals with a phenomena of acceptance, thereby averting many traditional problems. This is what Horwich calls an implicit definition. It avoids the flaws of any explicit definitional account of meaning, as we are not equating the meaning of x with the definition y; rather we claim that it is through accepting #x (where #x stands for the implicit definition or sentence containing x, as was our convention earlier) that we get the meaning of x.

This seems perfectly acceptable so far. However, if we add to this the way in which we get the meaning of a sentence, we again run into inconsistency. To get the meaning of the sentence we need:

K1. the meanings of words and,

K2. the combinatorial principle that combines them into a sentence.

So the meaning of #x would depend on the principle P1 and the meanings of terms including x. Now perhaps we can see that the acceptance fact of accepting #x is not that simple. The problem isn't simply that we need to accept some sentences underived, or that we already need to know the meaning of x in order to accept #x for instance. These are problems he deals with and accepts to some extent (see for instance Horwich 1998: 61-62). He argues that the use theory requires a certain degree of meaning interdependence, but this is to be distinguished from Quinian indeterminacy: that there can be no objective fact of the matter whether a given property is or is not what constitutes the meaning of the word. I believe the problem is rather deeper.

In order to accept the sentence #x, which gives us the implicit definition or the use property, we need

L1. to at least understand the meaning of #x; and

L2. in order to achieve (L1) according to Horwich's theory, we need to already know the meaning of the terms including x, as well as the combinatorial principle.

This is fine so far, but then what use is the use property #x in providing us the meaning of the term x? We could have just said for instance that in order to use the term x, in any particular instance, we already need to (implicitly) know the meaning of the term. So Horwich's use properties would not have a superior status over any other use we make of the term.

Alternatively, if we argue, as Horwich does, that the use properties are accepted underived, and must not be analyzed further, how is it that we get the meaning of a word from the underived sentence? In order to get the meaning of the word, the sentence must be understood compositionally, in a non-trivial sense. With a deflationary account, we would not be able to arrive at the meaning of the term at all. And if the use-property sentence is understood compositionally, then there appears to be no reason for claiming that in the case of all other sentences, compositionality is trivial.

Thus we can either give a substantive account of compositionality and explain both the 'derived' and the 'underived' sentences, thereby removing such a classification. Use properties would then be redundant. Alternatively, the underived sentences would not be able to give us the meaning of the term they are supposed to implicitly define, and if they are 'accepted underived' then why that is so or what use they have in a theory remains to be explained.

To say that certain use properties can be raised to the level of implicit definitions simply because they are for instance 'generally accepted underived' doesn't do any more than name a phenomenon that seeks explanation. One way to explain this 'acceptance phenomena' would be to say that in these particular cases, we seem to be *deriving*[45] the meaning of the word from its position in the sentence. So here compositionality is no longer trivial. But, if this is the case, then there seems to be no reason for assuming that compositionality is sometimes a trivial principle and at other times it is not. In other words, we should admit that from any and all uses

[45] Horwich (1998) argues rather that the particular use property sentence constitutes the meaning of the word, and no further step of derivation is required. But this, as we just pointed out, is precisely what we need to explain.

of the term x, it may be possible for us to derive the meaning of the term. Therefore, if compositionality is not trivial, then we need not have special use properties at all in order to derive the meaning of terms.

Thus, the problem arises in the following steps:

I. For Horwich the meaning of the word x is constituted by the implicit definition #x (a particular sentence which contains x) which he claims is accepted underived.
II. The meaning of a sentence, for Horwich is derived from the meanings of the words along with the combinatorial principle P.
III. So to understand the meaning of the sentence #x, we must also understand the meanings of the words in the sentence as well as the combinatorial principle P.
IV. This would amount to saying that we already need to know the meaning of a word in order to understand the implicit definition (or any sentence containing that word).
V. But Horwich would argue that the implicit definition is accepted underived and cannot be analyzed further.
VI. If this is so, the question arises: how do we get the meaning of the word x from the sentence #x?
VII. To get the meaning of the word from the sentence, the sentence #x (or the implicit definition) must be understood compositionally in a non-trivial sense, which makes it possible for us to derive the meaning of the word from its position in the sentence.
VIII. But Horwich gives a deflationary account of compositionality, where there is no connection between sentence construction (dependent on the construction property or combinatorial principle) and word meanings (derived from use properties).
IX. Thus, from (VII) and (VIII) above, it appears that in the case of implicit definitions we understand the sentence compositionally in a non-trivial sense, but in all other cases compositionality is trivial.

To avoid falling into such inconsistency, we could say either (I) that compositionality is always trivial or (II) that it is not in any case.

M1. If compositionality is in every case non-trivial, then we can in general derive the meanings of words (at least partially in every case) from their position in any sentence. So in this case implicit definitions are no longer special.

> M2. If compositionality is always trivial, including the case of implicit definitions, then we have no way of getting the meaning of the words from the implicit definition sentences.

Horwich appears to be defending his theory from a similar objection raised by Fodor and Lepore (in Horwich 2001: 379):

> (1) A complex's meaning what it does is determined by its structure and the meanings of its words.
>
> (2) A complex's G-property[46] is *not* determined by its structure and the G-properties of its words. (This is supported by examples: – e.g. the stereotype associated with 'pet fish' is
>
> not determined by the stereotypical pet and the stereotypical fish).
>
> (3) Therefore, the meaning of an expression is not engendered by its G-property.

But there is a hole in this line of thought. No matter what is substituted for 'G', the argument is valid only in the presence of a further premise: the following Uniformity Assumption

> If the meanings of words are engendered by their G-properties, then so are the meanings of complexes. (Horwich 2001: 379)

Without that assumption the most one can conclude from (1) and (2) is that *either* the meanings of words aren't engendered by their G-properties, *or* the meanings of complexes aren't. Thus it would be perfectly coherent for someone to deny the Uniformity Assumption and maintain that whereas the meanings of *words* are engendered by their inferential roles (say), the meanings of complexes are constituted in some other way. (Horwich 2001:379-80)

But we can see that this response does not apply to our current argument. We do not assume any uniformity thesis. Our claim is rather that a substantive account of compositionality is needed in order to derive meanings of words from use properties, or rather for explaining acceptance properties. Either we need such an account, or use properties are rendered useless within the theory as all we need to say is that in using terms, we already need to have an implicit knowledge of their meaning (as we noted earlier in this section, Horwich does admit to a degree of meaning interdependence, and so he may not see this as an objection: but again this points to the need of a substantive account for his theory). In fact, even in case we admit a substantive

[46] Here a G-property refers to that property which engenders the meaning of a word: it could be its inferential role, associated stereotype, recognitional capacity, etc.

account of compositionality, such an account could be given in each particular use of the term, again rendering use-properties useless. There would, as noted above, be no difference between a 'derived' and 'underived' sentence in this case. In other words, we would be able to arrive at the meaning of the term compositionally in any instance of use. Thus it appears that the only claim that Horwich should be making, even from his modest non-scientific, philosophical position is that we already (implicitly) need to know the meaning (which for Horwich consists in the acceptance of the relevant use property sentence) of a word in order to understand a sentence, and the meaning of a sentence is derived compositionally from the meaning of words and the combinatorial principle.

3.6 Conclusion

We started out hoping that Horwich's account would provide us with a way of systematizing a use theory. However, now that we have analysed his account, in its weaknesses and strengths, we note that to strengthen the account in the ways suggested, would render the use-theoretic aspect useless. Given the objection noted in the previous section, even at the end of our analyses of his theory, we seem to be in need of a way of systematically identifying the semantically/syntactically relevant use properties. If this is the case, then why do we need to invoke use properties at all if everything that the use theory actually does can be done without such properties? We seem to need, minimally, the semantically/syntactically relevant features, which may be directly available from the lexicon. In this case, we could simply abandon Horwich's account to adopt a more explanatory account, such as Chomsky's for instance, and still not lose the philosophical position that Horwich wants to advocate (though all that would be left of this position would be that which was already a part of the Chomskian enterprise to begin with, as suggested in section 3.4.3). In the next chapter, we will look at Borg's approach towards exploring this route further.

Thus it appears that we have the option of adopting a stronger theory, with the same philosophical import, but here 'use' remains as much in need of explanation as when we started our analysis. Is this a problem with the concept itself, or only with the way we have so far approached it? Before we return to this problem, in the next chapter, we will carry forward the above argument, abandoning use theory in favour of a more Chomskian approach.

3.8 Conclusions Towards Our Original Intuitions

Before we move ahead in our work, we need to consider here how the above discussion on Horwich contributes towards our original intuitions. We set out on this project in search of a theory that is both internally significant and can account for the aboutness inherent in language. In other words, we started out seeking a theory that would answer to both internalist as well as externalist intuitions. We then took up a study of Wittgenstein's writings towards this end, and a reinterpretation of a crucial element of his work (the concept of determinate relation) led us to a possible way in which these two aspects can be connected within a theory: if there is an element of use already present in what constitutes a proposition.[47] From our discussion in this chapter, we can see that Horwich informs this reinterpretation: use-properties as word meanings provide a possible way for use to be incorporated into what constitutes a proposition. As the thesis progresses, we will see whether we can build a consistent position which incorporates use properties as a 'content internal' understanding of meaning. In the next chapter, we pursue an opposite, externalist position: we try to see whether such a position, which argues against contextual access, is consistent and explanatory and whether we can gain something from it.

Again, to review the insights we have so far: with Wittgenstein, we saw that there is a way to understand use as constitutive of meaning. The determinate relation or grammar which constitutes a proposition can (as with later Wittgenstein) include use of a word, or rules of use, thereby making use constitutive of meaning, at least in terms of serving as a background for understanding language[48]. With Horwich's claims of systematizing use, we started out expecting a systematic way of arriving at and explaining use-properties, which are constitutive of the meaning of a word. However, by the end of our discussion of Horwich, we could see that Horwich was not able to fulfil our aim within the confines of his theory. We noted various problems with his theory, especially problems related to consistency of the theory as a whole.

Although use-properties can't be consistently explained within the parameters of Horwich's theory, I would like to retain this element of his work to be revisited later. This is because Horwich's use-properties could potentially be useful for us in our quest to see how use can be constitutive of meaning. It was important for us to consider not just Horwich's use properties in isolation, but within the context of the rest of his theory, including deflated

[47] See section 2.2 for details.
[48] See discussion of background in section 2.3 above

compositionality – to see whether they form a consistent whole. Having now considered the theory in all its aspects, we can just pick what we wish to retain (i.e. use-properties) while discarding the rest.

Chapter 4: Reference and Context Dependence in a Minimal Semantic Theory

4.1 Introduction

We claimed with Wittgenstein that elements of use (use in a context, rules of use etc.) creep in to what constitutes a proposition. Now we try to see whether there is a minimal level of significance devoid of intrusions from use. Semantic Minimalism (Borg 2004, 2012; Cappelen and Lepore 2005a, b) claims to offer such a theory. Borg (2004, 2012) in particular, defends an externalist referential account of minimal propositions against both internalism and contextualism.

Why this is interesting for us at this point in the development of our thesis is because it will enable us to see whether an externalist referential account, involving minimal propositions, could answer to our original intuitions requiring both an internalist and externalist explanation. In other words, if, apart from experiencing Reality, all that we have to access Reality is language, then how is it that we are able to reach out to the world, while being confined to the domain of language. A referential account of minimal propositions might serve our purpose here. Contextualism poses a special challenge to such an account arguing for instance that there is no minimal semantic level devoid of contextual input. Borg's argument against contextualism is also her argument in defence of her minimalist position, and for propositionalism. She understands the latter as posing the biggest challenge for her position. In order for all well-formed declarative sentences to express a complete (although minimal) proposition, there must be no such sentences that fail to express complete propositions prior to contextual input. To understand her position better, we need to understand what she is pitting herself against, and therefore we discuss Contextualism as well as arguments against Contextualism.

As noted above, in chapter 2 we saw, with Wittgenstein, or rather with our interpretation of Wittgenstein, that elements of use already creep into what constitutes a proposition. Our task in what follows is therefore to see whether there is, consequently, no minimal level of significance devoid of intrusions from use. In other words, if we were to understand a sentence

(or a proposition) as constituted by a structural element (syntax), and an element of meaning (semantics), could these stand alone or constitute a complete proposition (one that can then be compared with reality to determine its truth value), without accounting for any elements of use (pragmatics) or the context of utterance (contextualists)? We also saw, particularly with Wittgenstein's analogy with music (which is that we must understand language the way we understand music and not the other way round), that there must be significant sense in which language (or a sentence) 'conveys itself'. We saw that there is a way of understanding language which is internal to the domain of language. We considered use theory in chapter 3. The position we now consider, in order to test the validity of our intuitions, is an externalist theory that argues against contextualist intrusions to a minimal semantic level of interpretation. A good candidate for such a theory is Borg's Semantic Minimalism (2004,2012).

When I utter the sentence, 'I am extremely hungry' to a friend, I expect her to understand that here 'I' refers to *me,* and by this utterance I mean that I am hungry *right now*, and even perhaps intend for her to infer that we should go get something to eat. So the meaning of the utterance (in a broad sense) would be different depending on the person who uttered the sentence and the time of the utterance and the occasion of utterance. In fact radical contextualists[49] (e.g. Travis 1975, 2008), argue that there is no (semantic) content available prior to contextual access.[50]

In this chapter we will, firstly, attempt to see whether there is a minimal semantic level of explanation possible in such cases, or even at all. It will be our task here to see how much of the context can or even should be removed from an understanding of minimal *semantic* content, and what features a minimal semantic theory, if possible at all, should have. Secondly, we will try to see whether minimal semantics can (or even should) include a referential

[49] Radical contextualism, as the formulation here suggests, is the theory that there is no context free level of understanding an utterance. To claim this however is not to say for instance that for Radical Contextualists, statements with indexicals and without are treated at par (as Bezuidenhout (2006) clarifies, contrary to what Cappelan and Lepore 2005a seem to suggest)

[50] Consider a classic example from Travis (1975): "Fred is walking with his young nephew beside a pond where a decoy duck is floating. Pointing to the decoy, he says, 'That's a duck'. Again we might ask whether what he said is true or false. But again, the above description is not enough for us to tell. If Fred has just finished laughing at a sportsman who blasted a decoy out of the pond, and if he has been trying to show his nephew how to avoid similar mistakes, then what he said is false. But suppose that Fred and his nephew are attending the annual national decoy exhibition, and the boy has been having trouble distinguishing ducks from geese. Then what Fred said may well be true. It would also be true had Fred said what he did in pointing out the fact that all the other ducks were poor copies (perhaps on the order of Donald Duck)." (Travis 1975: 51) The same sentence is judged to be true in one case and false in another, depending on the context (or rather for Travis, the *occasion*) of utterance. We will consider the various implications of radical contextualism in section 2.1 and then in section 5.

component, which is suggested by Borg (2004), and which gives the theory its externalist character.

To begin, we first need to clarify the terms under discussion. What is 'Minimal Semantics'? Intuitively, it must be a theory which specifies what is minimally required to have (semantic) meaning. Borg (2004) interprets the minimal requirements to be that of a formal semantic theory, claiming specifically that "it is possible to deliver an account of the propositional or truth conditional content of a sentence in natural language simply via formal operations over the syntactic features of that sentence, that is, over the lexical items it contains and their mode of composition." (2004: 3) She adds to this a modularity of the mind claim: "the mind is composed (at least in part) of discrete, relatively autonomous modules" such that these modules "have access to something less than the full range of information which the cogniser as a whole possesses." (2004:7) Given these two claims, she maintains that there is a distinct module for semantic processing, governed purely by formally triggered syntactic operations, which does not require access to all the (contextual) information regarding the utterance which the speaker possesses, and which is able to deliver the literal meaning of (the utterance of) a sentence. In other words, all we require is the formal features of a sentence, and then the modular mind is able to compute the literal meaning of an utterance, hence giving us minimal semantic content. For Borg, even overtly context sensitive terms, like the indexical in the example we began with, can be understood as relativized to a context of utterance, but we do not need to know the context in order to understand them. For this reason, they can be included within the scope of a minimal semantics. We will discuss this further in the next section.

Related to this point, Cappelen and Lepore, henceforth C&L (2005a, b) recognize that some terms are inherently context sensitive. For C&L semantic minimalism has the following features:

> a) The most salient feature of Semantic Minimalism is that it recognizes few context sensitive expressions, and hence, acknowledges a very limited effect of the context of utterance on the semantic content of an utterance. The only context sensitive expressions are the completely obvious ones ('I', 'here', 'now', 'that', etc., (essentially those Kaplan lists in 'Demonstratives', (1989, p. 489). These are not only obvious, they also pass certain tests for context sensitivity we spell out below.

> b) It follows that all semantic context sensitivity, insofar as it exists, is grammatically (i.e., syntactically or morphemically) triggered. c) Beyond fixing the semantic value of these obviously context sensitive expressions, the context of utterance has no effect on the proposition semantically expressed. In this sense, the semantic content of a sentence S is that proposition that all utterances of S express (when we adjust for or keep stable the semantic values of the obvious context sensitive expressions in S). (C&L 2005b: 197-98)

C&L add to this a Speech Act Pluralism to explain the relationship between semantic content and speech act content (Cappelen 2005a: 15). Speech Act Pluralism is the belief that "No one thing is said (or asserted, or claimed or…) by any utterance: rather, indefinitely many propositions are said, asserted, claimed, or stated. What is said (asserted, claimed, etc.) depends on a wide range of factors other than the proposition semantically expressed. It depends on a potential infinitude of features of the context of utterance and of the context of those who report on (or think about) what was said by the utterance." (C&L 2005b: 199)

Thus for C&L it is possible to hold both a story about minimal content (which includes *some* context sensitive terms) and also grant that that content is not the same as what is said, rather what is said by an utterance depends hugely on the context of the utterance. This claim however, does not undermine their claim that there is minimal semantic content - which is the proposition that all utterances of a sentence express.

In order to hold either of the above claims about minimal semantics (Borg 2004, C&L 2005a,b) the advocates of these theories must at least do the following:

A. Defend their minimalist claims against contextualist arguments.
B. Defend a propositional account of semantic content.
C. Specify why their minimal semantic content is referential, i.e., why it involves a referential component.

We shall discuss how far these approaches succeed in these tasks. Beyond the statement of the minimal semantic position, Borg (2004, 2007, 2009) goes further to distance herself from internalism, which, she claims, advocates that, following Chomsky, "semantic content should be specified in entirely internal non-referential terms." (Borg 2012:148)

Thus in addition to 1 and 2, Borg must also undertake to:

D. Defend an externalist, referential account of semantic content from internalist criticisms.

We will now begin with (A), the defence of formalism over contextualism as offered by these versions of minimal semantics.

4.2 Favouring Formalism over Contextualism

Formalism concentrates on the formal features of words and sentences whereas Contextualism places greater value on the input from context. We can not wish away either element if the proposition uttered is to have a truth value. For instance, to know whether "It is raining" is true we need to know the place where the utterance takes place. Thus formalism must also allow for contextual influences. But the question that arises for formal semanticists is: how much intrusion from context should be permissible? A formalist would reply that only those contextual intrusions which are triggered by the formal elements of a sentence are semantically relevant. Again, however, various formalists differ regarding the kind of syntactically triggered contextual elements. C&L for instance allow for a small list of contextually sensitive expressions, as noted above. Borg on the other hand claims that even these obviously context dependent terms can be interpreted syntactically, in a context neutral way (i.e., they have a kind of meaning that is context-independent), and insists that the process of saturation from context need not be a part of semantics as such. In making these exclusions from semantics, Borg severely restricts its scope. We will, in what follows, try to understand whether Borg's formalism (semantic minimalism) is warranted, even within its modest claims. For this, we first note the traditional case for contextualism, then we look at Borg's arguments against it. We then move on to evidence from linguistic processing in favour of contextualism and see whether this is able to counter Borg's argument.

4.2.1 The Case for Contextualism

Traditionally, the case for contextualism involves arguments from context shifting, incompleteness and inappropriateness. The problem of incompleteness comes across in examples such as: 'Steel isn't strong enough' or 'Jill is ready'. Here a complete proposition emerges only when we find out from the context what it is that steel isn't strong enough for or what Jill is ready for. Thus incompleteness arguments "draw on the intuitions that context must contribute something to what is said, since the sentence itself does not have a content that is truth evaluable." (Pagin and Pietelli 2005: 34) Relatedly, the problem of inappropriateness

emerges in examples like 'It is raining'. When someone in London utters 'It's raining' it means that 'It's raining in London'[51]. When you say 'There is nothing to eat' you mean there is nothing to eat in the house. Thus, "sentences express complete propositions without contextual enrichment, but these propositions are not the ones speakers express by uttering those sentences." (Borg 2006: 4). Hence the inappropriateness. The problem of context shifting arises when for instance a relative term is used without specifying the relation to a class. When you say 'Murali is tall' it may mean for instance that Murali is tall for a five year old, though he may be short for a basketball player. As Pagin and Pietelli paraphrase: "If a semantic feature of utterances of sentences containing *e*, and associated with *e* changes from context to context, then *e* is context sensitive." (Pagin and Pietelli 2005:31, here '*e*' stands for an expression).

C&L argue that context shifting arguments can be provided for any sentence. For instance in the sentence 'That is a dangerous dog' the relation of the particular animal to the class of dangerous dogs would depend on the context, i.e. the dog could be aggressive, or he could carry a viral disease etc. Similarly, C&L argue regarding incompleteness arguments that in case we allow for incompleteness in one sentence, we must admit it for any sentence, and you can always add factors that require explanation: e.g. for the (more complete) sentence: "Steel isn't strong enough to support the roof" you can always ask for how long the support would last, etc.

The problems become clearer if we look at another way of classifying them. There is traditionally thought to be a distinction between literal meaning of an utterance, and the speaker meaning or what the speaker wishes to convey. Within the literal meaning, there is, it has been argued, a distinction between the unchanging sentence meaning and what is said. There may be a syntactically triggered need for contextual interpretation, for instance in case of indexicals and demonstratives. As Recanati points out, according to Minimalism, "'what is said' departs from the meaning of the sentence (and incorporates contextual elements) *only when the sentence itself sets up a slot to be contextually filled.*" (Recanati 2001: 77) For instance sentences with indexicals such as "He is angry" need contextual fillers for the pronoun "he" before the sentence becomes a complete proposition. This, for Recanati, involves *saturation*: "the contextual process whereby the meaning of such a sentence is completed and made

[51] A Radical Contextualist would deny such claims, and argue rather that 'it's raining' could mean a lot of different things depending on the context: e.g. when your task is to watch a lamp that flashes whenever it rains in Dubai, and to shout 'it's raining' when you see the light, then 'it's raining' is not short for 'it's raining in London'. We will deal consider this view later in this section.

propositional." (2001: 77) There are however other contextual processes which have to do with speaker meaning. These involve for instance what Recanati calls modulation (which are also primary, though optional, processes) along with saturation (which both Recanati and C&L agree is a mandatory process, though Borg disagrees, as we shall see later). Modulation includes *free enrichment*, *loosening* and *predicate transfer* or *semantic transfer* as shown respectively in the following examples:

(1) The table (of our living room) is covered with books.
(2) The ATM swallowed my credit card.
(3) The ham sandwich left without paying.

In (1) the bracketed part is not mandatory for a complete proposition, and is not articulated, but is tacitly added through free enrichment (Recanati 2004: 10). This makes the truth condition more precise, i.e. free enrichment is "the process responsible for making the interpretation of an utterance more specific than its literal interpretation" (Recanati 2005: 177). The word 'swallowed' in (2) is used loosely with extended meaning, to include application over credit cards and being undertaken by machines. The phrase 'The ham sandwich' in (3) is used to refer to the person who ordered it instead of reference to the sandwich. Thus predicate transfer "takes us from a certain property, conventionally expressed by some predicative expression, to a distinct property bearing a systematic relation to it." (Recanati 2005: 177) Apart from primary pragmatic processes, Recanati argues for secondary pragmatic processes involving Gricean implicature, as for instance in our initial example 'I am extremely hungry' as said to a friend. What is implied is that they should get something to eat, though this is not what is expressed by the proposition. Let's consider the following example from Recanati (2004: 8) uttered in reply to the question: 'Do you want something to eat?'

(4) I've had breakfast.

The indexical here requires saturation from context. Further, the sentence has the added sense (through free enrichment) that 'I've had breakfast *today*', though it would be strictly speaking true even if the speaker had breakfast a week or even years ago. Also, as uttered in reply to the above question, (4) may imply that the speaker is not hungry at the time of utterance, since she had breakfast earlier, although it need not imply that.[52]

[52] The utterance of 4 may imply that the speaker does want something to eat, though not breakfast, or she wants to eat despite having had breakfast.

Thus the argument for contextualism is that many sentences need saturation, modulation, as well as involve implicatures, in order to be truth evaluable and have meaning. These processes necessarily involve inputs from context. Therefore context is an essential (minimal) element of meaning. In other words, there is no minimal truth conditional content which is devoid of contextual inputs. In the next subsection we will take a closer look at the minimalist retort to this argument. Radical contextualists, however, do not resort to either incompleteness or context shifting arguments explained above. The Travis (1975) example stated in the introduction is a case in point. It has the following implications: firstly, as we already noted, the same sentence, and this can apply to any (closed) sentence, can be judged to be true in one case and false in another, depending on the context (or rather for Travis, the *occasion*) of utterance. Second, this severely undermines the idea that sentences have straightforward disquotational truth-conditions, since here, to say that 'That is a duck' means *that that is a duck*[53] does not serve us in any way towards determining the truth of the sentence, or providing a substantial condition for truth. Thirdly and derivatively, it suggests that truth conditions are not a property that utterances of sentences inherit from the sentences instanced[54]. In other words, there are no determinate truth-conditions which we can recover from the utterance of a sentence, Consider the utterance 'Pia dances'. Travis (2006) argues that understanding the truth conditions of this sentence requires much more information than is provided by the proposition *that Pia dances*. "If Pia is a trained, skilled, dancer but refuses now to do it, is that proposition true then? If she eagerly takes the floor at every opportunity, but is so clumsy that one might refuse to call it dancing, is it true then? If she dances, but only when you heat the floor enough, is it true then? And so on." The argument is that if the truth or falsehood of 'Pia dances' depends on any of these stated scenarios, or if the minimalist answers yes to the first scenario, "then that proposition cannot be the one the English sentence 'Pia dances' expresses. For it does not, as such, speak of dancing on that understanding of engaging in it. If they systematically say 'Don't know', or steadfastly maintain silence, then (unless they are hiding something) there simply is no answer to the question when that would-be proposition would be true in which case it is no proposition at all."[55]

[53] We may note here that this is not a typical disquotational statement of truth conditions: indeed, it is a statement of meaning, not of truth conditions.

[54] Occasion sensitivity is for Travis (2008) distinguished from context sensitivity in that it is not just the idea that every sentence can depending on the context be understood as either true or false, as in the case of the example above. It is rather the idea that this is so because truth-conditions are not a property that instances of sentences inherit from the sentences instanced.

[55] A possible answer could be to state this proposition in a deflationist or disquotational way,i.e. not informatively.

So the Radical Contextualist's argument is that it is not the case that truth conditions of various sentences require inputs (through modulation, saturation and implication) from context, but rather that there is no determinate proposition which can be recovered from the utterance at all, even with appeal to contextual features of the sentence (for instance in the 'That is a duck' example.). However, there are many assumptions which are a part of this view: for instance a proposition is something which can be true or false, and knowing a proposition involves knowing what makes it true or false, not simply the disquotational truth condition. One can argue that the kind of answers on which the truth of 'Pia dances' in the above, is dependent, are themselves a matter a judgement and not of meaning. For Travis however, that is not a criticism. His argument is rather that these matters of judgement affect our ability to pick out the proposition expressed by the sentence. In this section, we also noted the contextualist arguments from incompleteness and context shifting. The challenge that contextualism offers to minimalism is the claim that either there is no determinate proposition or a proposition, in order to be complete, requires appeal to contextual elements. Thus to meet the challenge, it would appear, Minimalism needs to show that there is in fact a determinate proposition (the minimal proposition) which does not require appeal to contextual elements. In section 2.2 and 2.3 we will see how far minimalism is able to meet this challenge.

4.2.2 The Minimalist Argument Against Contextualism

We have seen that there are many elements of meaning[56] which are in fact contextually determined. In an attempt to exclude these elements from 'meaning', a minimalist could, in scenario one, limit the scope of what we can strictly call (semantic) meaning (Borg 2004; C&L 2005a,b). This would require defending a sharp distinction between meaning and use, semantics and pragmatics. In scenario two, the minimalist could claim that there is an explanation for these elements from context which lies within syntax. So the two strategies involve either excluding or subsuming context. The latter (as for instance in Stanley 2000), according to Recanati, requires one to "re-analyse the example so as to show that the pragmatic

[56] Though it remains to be seen whether this 'meaning' is to be understood as semantic meaning or deferred to pragmatics. Thus perhaps 'understanding' would be the appropriate usage here, though this would in a sense preclude semantic meaning, since understanding, with all its psychological implications steps beyond pure semantics.

process at issue… is a bottom–up process triggered by some expression in the sentence"[57] (Recanati 2005: 176), for instance a free location variable.

Stanley (2000) argues that context dependence of expressions is restricted to the occurrence of indexicals (I, here, you, now), demonstratives (this, that) and pronouns (he, she) within the expression, and "any contextual effect on truth-conditions that is not traceable to an indexical, pronoun, or demonstrative in the narrow sense must be traceable to a structural position occupied by a variable." (Stanley 2000: 401) So for Stanley the contextual dependence of "It is raining" is due to the hidden variables related to time and place, and that of "Sherman is small" is due to the hidden variable related to the comparison class of 'small'. Stanley importantly distinguishes his account from an unarticulated constituent account (Perry 1986). The latter allows the context to directly supply the location in, for instance,

(5) 'It is raining'.

So the sentence is rather to be understood as: It is raining (here)here the bracketed term is the unarticulated constituent supplied directly from the context. Other examples would be:

(6) You are not going to die (from that cut),

(7) She has had enough (of the lunch)[58]

Stanley argues instead that rain in (5) "occurs with two open positions. Each open position is filled by a term of the form '*f (x)*'.[59] The first function maps entities to times, and the second function maps entities to locations. Context supplies the value of the function variables '*f*' and '*g*'." (Stanley 2000:416) Stanley (2000) argues that there are no unarticulated constituents, and all truth conditional effects of context can be traced to logical form. The problem for hidden variable accounts is determining which of the unarticulated constituent cases (for instance (5)-(8)) require variables within the syntax or the logical form of the sentence, and which of these are not cases where the alleged unarticulated constituent is a part of the truth condition of the sentence. In other words, the question is how much of the context should be formalized, or rather, formally represented in the syntax, and what should not be. For Stanley, the answer becomes: that part of the context which affects the truth condition of

[57] Bottom-up processes, as mentioned in the quote, are those that are triggered from within the (syntax of the) sentence. By contrast top-down processes are those that have their trigger outside the (syntax of the) sentence, for instance elements from the context of utterance or discourse.
[58] See Stanley (2005: 223-224) for more such examples.
[59] Omitted footnote from the original quote.

the sentence, can be traced back to logical form, or in our terms, should be formally represented in syntax. So Stanley (2000) argues that there are no unarticulated constituents, and that any contextual elements which affect the truth condition of the sentence can be traced back to syntax, as a hidden variable. Stanley (2005) argues that the bracketed element in (8) is no part of its truth condition.

(8) John had breakfast this morning (in the normal way, through his mouth).

So it is no part of the truth conditions of an utterance of this sentence that John ingested his breakfast through his mouth.

For our purposes, the more pertinent minimalist view would involve excluding the context, since the exclusion would make the account more minimal. So going back to scenario one, that of excluding the context: Borg (2004) agrees that overtly context sensitive terms like indexicals and demonstratives are special, but she argues that "what is special about demonstrative and indexical content resides not with the kind of epistemic contact between speaker and referent, but simply with the mode of expression of our thoughts. Fundamentally demonstratives and indexicals differ in meaning from quantified noun phrases because they express a different kind of concept – singular concepts in the former case, general concepts in the latter." (Borg 2004: 186-87) Thus in the case of indexicals, in for example an utterance of

(9) You were born in the capital of Nebraska.

a competent hearer is able to pick out that 'you' can be understood as 'the addressee of (9)', even if she is not able to pick out the object of reference. This is what Borg means by picking out a general concept. The meaning of the term 'you' itself leads to this interpretation (or picks out this general concept 'the addressee of (9)'), without any knowledge of the context of utterance. Thus the pragmatic processes of saturation and modulation discussed above are not *primary* as Recanati suggested, as they are not involved in this literal meaning.

In the case of demonstratives, for example in (10):

(10) That is red

Borg (2004) argues that, from an utterance of (10) a competent language user is "capable of recovering the semantic content that *α is red* (via the conditionalized T-sentence: [If the speaker of 'this is red' refers with the utterance of 'this' therein to α and to nothing else, then that utterance is true if and only if α is red]). Here α is a syntactically generated singular concept,

the content of which is exhausted by the object referred to by the speaker in this context" (Borg 2004:192) However, "Possessing a singular concept does not," for Borg, "entail that an agent is able to (non-descriptively) identify the object being thought about from all other things" (2004:193).

There are a couple of things we may note from the above exegesis. Firstly, what one refers to with indexicals are general concepts, which refer (to an object) through a general description (of the form "the addresse of an utterance u") while demonstratives are special in that they are singular terms, which involves, Borg maintains, entertaining a singular concept, which is a 'genuinely referential thought'. Secondly, though both the descriptive and the singular terms serve to pick out objects of reference (the latter more directly than the former), the competent hearer needs only to recognize this fact; she need not be able to identify which particular object is referred to. In other words in order to understand a sentence with indexicals or demonstratives, one need only recognize *that* there is (potentially, that is without commitment to it's existence) an object of reference, and need not be able to identify what (or where) the object is.

Identifying the object of reference or the content of these singular or general concepts is, according to Borg, not a job for semantics. But even her minimal account of 'you' as 'the addressee' and 'that' as a singular concept, involves, it may be argued, more than what can be recovered by mere linguistic meaning. And here comes in the notion of character: "It is character which is responsible for connecting linguistic content with non-linguistic information, combining purely linguistic understanding with the rich vein of information we have from perceptual, memorial and other processes." (Borg 2004: 197) Borg's use of the term character is different from Kaplan's. Borg sees this difference as that of domains of application: according to Borg, Kaplan's character links language and the world, while her use of the term provides a "bridge between the language faculty and the rest of the cognitive domain" (2004:197)

The 'character' Kaplan argued for is a function from the context to the content; "I call that component of the sense of an expression which determines how the content is determined by the context, the 'character' of an expression." (Kaplan 1978: 84) The content on the other hand is always taken with respect to a given context of use. So the utterance of 'I am here now' would have different content depending on the person, place and time of the utterance. Kaplan argued that "demonstrative free expressions have a constant character, i.e. they express the

same content in every context." (Kaplan 1978: 85) In changing the domain of application of 'character' from linking language to the world, to linking language to other cognitive domains, Borg appears to be changing the application of 'context' from the world to mind-internal cognitive domains – a prima facie internalist move. However, it may be noted that the content of a concept is, even for Borg, exhausted by the object referred to (whether or not it is possible to identify the object), as suggested by her analysis of (6) above. Thereby arises Borg's externalism and the need to defend it against internalism.

A further question arises here: how does this process, of linking the various cognitive domains, take place within Borg's modular theory of mind? The following provides an answer: "A concept generated on the basis of a token linguistic expression of 'you' will *only* merge with a perceptual concept of an addressee in the context in question, 'yesterday' requires the concept brought out from the agent's memory store to be that of the day before the day of utterance, while 'that' requires a perceptual concept of an object at some proximal distance from the speaker, and so on and so forth." So for linguistically generated concepts, their syntax "dictates the kinds of mergers they will allow with concepts emerging from other modules. (2004: 199)

Here we may note that Borg follows Fodor (1983) in the claim that the mind is at least in part modular. For Borg, a module is "a probably innate, encapsulated body of information, together with processes operating only over that information, which is responsible for realizing a given cognitive function." (Borg 2004: 80)[60] Fodor identified modular systems as, "domain specific, innately specified, hardwired, autonomous, and not assembled" with "informational encapsulation" being "perhaps the most important aspect of modularity." (Fodor 1983: 37). Importantly, Borg follows Fodor in his claim that "the notion of modularity ought to admit of degrees… When I speak of a cognitive system as modular, I shall therefore always mean, "to some interesting extent". (Fodor 1983: 37). Thus the kind of modularity Borg supports is not a strong or massive modularity[61] but rather a weak form of modularity, which can admit of relatively non modular systems (e.g 'character' for Borg is either a- modular or is capable of

[60] As Prinz (2006) points out, modularity needs to be distinguished from functional decomposition, where the mind/brain is a network of interconnected systems which are functionally distinct.
[61] The thesis that the mind is massively or entirely modular, as the term itself suggests. See for instance Sperber (2001) for a defense of this thesis.

application across modules), along with (more) modular ones.[62] In section 3 we consider Borg's argument from modularity and her commitment to Fodor in greater detail.

Thus, as we noted in this section, Borg (2004) solves the problems raised by Contextualists by shifting the burden towards the other side of the semantics/pragmatics divide. The purpose of a semantic theory is restricted to the 'literal' as opposed to the 'communicated' meaning:

> "To understand what is literally meant by 'It's raining' all we need to know, the (minimal) formal theorist claims, is the meaning of the syntactic parts of the sentence and their mode of combination, however to know what is communicated by an utterance of this sentence we need to know so much more." (Borg 2004: 257)

Contextualists thus address questions of communication while formal semantic theory addresses questions of meaning. The two apply to different domains. Further, she gives three reasons for preferring a formal theory over contextualism or dual pragmatics. First, there is the fact that we can create and understand new sentences. Language is productive and systematic. Compositionality – that the meaning of a complex whole is determined by the meanings of its parts and their mode of composition – appears to be the best explanation for the generative nature of language and the fact that it exhibits these three properties: creativity, productivity and systematicity. Only a formal theory can accommodate compositionality, as according to dual pragmatists / contextualists, a part of the literal meaning is determined by the context and is not contained in the formal constituents. Secondly, she claims, it is always possible to distinguish between the literal and the contextually implied meaning. Literal sentence meaning acts as a 'non-cancellable level of content in a linguistic exchange.' Finally, modularity is, according to Borg, independently desirable. A formal theory allows linguistic meaning to count as part of a modular language faculty, and therefore it should be preferred to theories that do not.

These claims leave room open for many kinds of theories, which admit of a distinguishable level of literal meaning, and accept a modular theory of mind. The key to evaluating whether Borg's arguments for formalism, and against contextualism work, lies,

[62] Robbins(2007) notes that Borg's argument against contextualism, from modularity is as follows: the mind is modular, and in particular, the semantic component of the language faculty is modular. But Contextualism is inconsistent with modularity. Therefore, contextualism is false. Robbins (2007) claims further that Borg can use same argument for weak minimalism of the kind advocated by C&L. However Robbins claims that semantics is not modular, and therefore the argument does not hold.

therefore in understanding the particular kind of modularity and of minimalism she advocates. Even Contextualists such as Sperber advocate a (massively) modular theory of mind, and though C&L admit of a distinguishable level of literal meaning, they do allow for a restricted set of inherently context sensitive terms

To summarise, minimalists argue that the truth-conditionally relevant contextual elements find their source in syntax (hidden variables for Stanley 2000 or the class of demonstratives for C&L 2005a,b) or can be interpreted in syntactic terms (Borg 2004). For instance, for Borg, 'you' is understood as the contextually invariant form: 'the addressee of the sentence'. So even if context is involved in the minimal content of the sentence, as it is for C&L, it is limited to syntactically triggered elements, and even these syntactically triggered contextual elements must, for Borg, find a context invariant interpretation. Borg argues that there is a minimal semantic level of interpretation prior to contextual interference. At this minimal level, we may not be able to identify the object of reference, though through the notion of character, we have a bridge between the linguistic and other cognitive domains which help attach a concept to the descriptive understanding of for instance 'addressee of the sentence'. And the content of a concept is, for Borg, the object itself. Hence, through the interaction of cognitive domains we are able to pick out the object of reference, and the theory retains its externalist character. In the next subsection we shall see whether the claim that there is a minimal semantic level of interpretation, holds in the face of neurological evidence (presented by Hagoort et al (2004) and Nieuwland & Van Berkum (2006)).

4.2.3 Empirical Evidence: Questioning the Possibility of Non-Contextual Processing

Hagoort et al. (2004) argue that processing the expressions (7) and (8) yields very similar brain signatures in a context where it is well known to the experimental subjects that Dutch trains are in fact yellow:

(11) Dutch trains are sour.
(12) Dutch trains are white.

They argue that (11) is a prima facie 'semantic violation', much like the sentence 'the favorite palace of the present queen of England is divorced' which "has no coherent semantic interpretation, because the predicate *is-divorced* requires an animate argument." (Hagoort et al 2004: 438) (12), in contrast, is a prima facie 'world-knowledge' violation, since it "has a coherent semantic interpretation, but it contains a proposition that is false in the

light of our knowledge". Now, if checking against world knowledge comes after a stage of processing based on syntactic information and its compositional semantic effects only, then effects of semantic violations should precede violations of world knowledge and should invoke other brain areas. But the results of a combined EEG/MRI experiment show that the effects obtained for the world-knowledge violation are essentially identical in onset and peak latency to the 'semantic violation', and they are very similar in amplitude and topographic distribution to the semantic N400-effect as well. Thus Hagoort et. al (2004) claim to "provide evidence against a nonoverlapping two-step interpretation procedure in which first the meaning of a sentence is determined, and only then is its meaning verified in relation to our knowledge of the world.". They conclude rather that "the brain retrieves and integrates word meanings and world knowledge at the same time." (2004:440)

However, (11) and (12) are not that different after all. There is no syntactic violation at play, and compositionally, or in terms of their minimal semantic content, both sentences make perfect sense. This is clear from the earlier sentences used to explain the difference between semantic violation and world knowledge violation as well. That a place cannot be divorced requires world knowledge (about what a place is) and is not specified by syntax, just like it requires world knowledge to know whether or not the present queen is divorced. The difference between these two cases is presumably that it is part of our knowledge of what a place (at least as this term is normally understood) is that it cannot be divorced. But it is not part of our knowledge of *who* the present Queen is that she is divorced or not. The latter case requires world knowledge about the marital status of the present Queen, and not just about who the Queen is. However, even the former case, where it appears to be a part of knowing what a place is that it cannot be divorced, there may be a context in which even this sentence has a coherent semantic interpretation: for instance in a story where places were animate and had marriages and divorces (however farfetched that sounds). Thus neither case is truly free of a contextual interpretation; i.e. the context can override what appears to be part of the meaning of the words themselves.

It therefore appears against the authors that the fact that the two cases (of putative 'semantic violation' and 'world knowledge violation') do not have distinct brain signatures does not in fact provide evidence against a theory that advocates distinctive semantic and contextual levels of interpretation, which Borg appears to support. Since there is no difference in the relevant regard between the two cases in the first place, the experimental demonstration that they do not have different brain signatures does not prove or disprove anything, leaving

Borg's argument regarding a semantic core intact. This data does not justify merging two possible levels of processing: semantic and contextual, thereby not making an argument against an approach such as Borg's.

There is, however, other research which suggests that there is no level of processing which is non-contextual. We will now outline this work and discuss its implications for Borg's work. Nieuwland & Van Berkum (2006) concede that what was claimed above to be a semantic violation is rather a world knowledge violation: "One type of world knowledge that is occasionally 'violated' in a discourse such as Alice's Adventures in Wonderland is knowledge of the things that nonliving, inanimate entities can and cannot do in the real world." (Nieuwland & Van Berkum 2006) They go on to question the assumption that the distinction between semantic violation and world knowledge violation was based on: namely that certain semantic features, like for instance the animacy-constraint (i.e. sentences such as "the favourite place of the present queen of England is divorced" do not have a semantic interpretation because 'is-divorced' requires an animate argument) is primitive to word meanings and thus is processed prior to and independently of any computations that establish meaning in a wider context. Against this assumption they show rather that "contextual appropriateness not only can ''neutralize'' animacy violations, but can even render the animacy-violating predicate more preferred than an animacy-obeying canonical predicate." In particular, the following story was given to subjects:

> A woman saw a dancing peanut who had a big smile on his face. The peanut was singing about a girl he had just met. And judging from the song, the peanut was totally crazy about her. The woman thought it was really cute to see the peanut singing and dancing like that. The peanut was salted/in love, and by the sound of it, this was definitely mutual. He was seeing a little almond.

In the context of this story, the predicate 'salted' as used for the peanut, which is otherwise a common predicate, turned out to be unexpected by the participants of the experiment, as shown by the N400 effect. Here again, however, it may be argued, as above, that whether in the context of a story or without that particular context, 'The peanut was in love' still has its compositional meaning.

Borg may reply that the animacy feature or other semantic features are not primitive or meaning constitutive, even though some of them may or may not be lexically marked. What one means, *minimally*, when uttering the sentence 'The peanut was in love' is nothing but that

- the peanut was in love. A minimalist in general, whether a Chomskian internalist or a minimal semanticist, would agree that the compositional meaning of the sentence is what is minimally required, with or without the context. Nonetheless, this example is particularly relevant in our present context, as we are trying to see whether or not contextual elements enter into semantic processing: this experiment appears to support the view that there is no minimal level of semantic processing devoid of the effects of the context of discourse. What the neuroscience data suggest is that there is no level of processing devoid of contextual input. This may however still not affect Borg's position with respect to a minimal compositional core of meaning: no matter what the context is, the truth conditions of 'The peanut was in love' are that some contextually specified peanut was in love. On the other hand, this very core is what she shares with internalism.

4.2.4 Interim Conclusions

We thus arrive at the following conclusion: Borg's refutation of the 'dual pragmatic' version of contextualism, even if successful, can be viewed as consistent with internalism. Although we did find evidence for contextualism in so far as there may be no level of processing devoid of context, the fact that context enters into semantic processing does not affect the compositional core of meaning, the existence of which Borg adopts along with internalism. The evidence does not seem to support any separate level of semantic processing (though neither does it refute the possibility). What is left intact after considering the evidence from neuroscience is just that part of Borg's view that she shares with Chomskian minimalism. Hence, the aspects that Borg's minimalist theory does not share with internalism still require extra support. We therefore now need to consider the divergence of the two: namely the fact that despite those of her commitments that are consistent with internalism, she is in fact an externalist. We tackle this issue in section 5. Before, we will re-consider the modularist argument against contextualism, examining the Fodorian assumptions that lie behind this argument, as well as the externalist commitments that arise from these assumptions.

4.3 Fodorian Motivations: Modularity, Atomism and Externalism

4.3.1 Modularity

One main argument against contextualism is, for Borg, that it is not consistent with a modularity of mind claim, and the latter is, she insists, independently desirable. To counter this claim, we could argue either that the mind is not modular, or that the mind is not modular in

the sense in which Fodor and Borg maintain, or that the modularity (either of Fodor's kind or a modified version) is consistent with contextualism. An important part of accepting or rejecting modularity thus boils down to somewhat of a terminological issue: what we mean by modularity. To accept one kind of modularity as opposed to another may affect whether or not modularity is consistent with contextualism. To attempt to address this issue, we need to be clear regarding the various possible meanings of modularity.

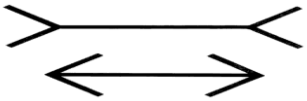
There is a relatively uncontroversial way of understanding modularity, but this is not the one that Borg or Fodor defend: the concept of a functional module.[63] A computer is modular if the different parts (the desktop, keyboard, speakers, cpu etc.) can all be purchased separately and then put together, and applying this to the mind, functional decomposition is "the simple view that the mind is a machine with component parts" (Prinz 2006: 34) each of which performs a specific function. Under this conception of modularity, "the thesis of massive modularity would claim that the mind consists entirely of distinct components, each of which has some specific job to do in the functioning of the whole." (Carruthers 2006:4) This however is not the modularity that Fodor had in mind. Fodor (1998) argued for a modularity of input systems, so importantly computational systems are excluded from the modular picture. Fodor (1983) listed the following as characteristics of input system modules:

i) Domain-specific; i.e. "there are distinct psychological mechanisms – *vertical* faculties – corresponding to distinct stimulus domains. " (Fodor 1983: 48) e.g. vision operates over visual stimuli.
ii) Mandatory operation; i.e. input systems are "constrained to apply whenever they can apply" (Fodor 1998: 53). For example "You can't help hearing the utterance of a sentence (in a language you know) as an utterance of a sentence" (1983:52-53)
iii) Limited central access: "Not only must you hear an utterance of a sentence as such, but, to a first approximation, you can hear it *only* that way." (1983: 56) Intermediate levels of processing are relatively inaccessible
iv) Fast processing
v) Informationally encapsulated: information from other modules or cognitive domains is inaccessible to the individual modules and also other domains/modules cannot access the information from any individual module; in other words modules are immune to feedback either way, from other modules or to other modules: e.g. the Muller-Lyre

[63] Or rather, as Carruthers 2006 puts it: "a dissociable functional component" (Carruthers 2006: 1)

illusion (Figure 4 below) in the case of vision, where we perceive the two lines as unequal, even though we know that that is not the case.

Figure 4: Muller-Lyre Illusion



vi) Shallow output; e.g. "the language input system specifies, for any utterance in its domain, its linguistic and maybe its logical form. It is implicit in this proposal that it does no more than that – e.g., that it doesn't recover speech act potential (except, perhaps, insofar as speech-act potential may be correlated with properties of form, as in English interrogative word order)." (1983: 90)

vii) Associated with fixed neural architecture

viii) Exhibit characteristic and specific breakdown patterns, e.g. agnosias and aphasias.

ix) Ontogenetically exhibit characteristic pace and sequencing; i.e. to put it differently, modules are innate and develop in a predictable way.

Borg argued that the semantic component of the language faculty is a module and exhibits each of the above features. So her argument from modularity, against contextualism, is that: the mind is modular, and in particular, the semantic component of the language faculty is modular. But Contextualism is inconsistent with modularity. Therefore, contextualism is false.[64] Robins (2007) argues against this by attacking the premise that semantics is modular. Another way to counter this argument would be to argue that the mind is not really modular (see for instance Prinz 2006) or that the mind is modular but modularity is different from Fodor's conception of it (e.g. Carruthers 2006). Instead of restricting ourselves to a limited input systems modularity, the mind is rather massively modular, with modularity extending to central systems as well. This could be a functional modularity of the kind noted above. Within this massively modular mind, we could admit of a modular semantics consistent with a form of contextualism. We will discuss this further later in this section, in the context of Carruthers

[64] . Robbins (2007) claims further that Borg can use same argument for weak minimalism of the kind advocated by C&L.

(2006) and Sperber (1994). Robbins (2007) argues that semantics is not an encapsulated process, as to know the truth condition for "I'm going to the bank" we need to know both the identity of the speaker and whether the destination is a financial institution or a topographical feature (Robbins 2007: 308). However Borg's minimal semantics can account for this, as we shall see in section 4.5. Robbins also claims that the breakdown patterns Borg refers to, such as Aspergers syndrome and Schizophrenia, which are supposed to leave syntax and semantics intact, with deficits in mind reading and linguistic pragmatics, do not infact provide evidence for the dissociation of semantics and pragmatics. This is because individuals with Aspergers perform well (or like people without Aspergers) on a range of standard mindreading tests (Bowler 1992) as well as tests for causal mechanical beliefs (Baron-Cohen et. Al. 1999, 2001); and people with Williams syndrome appear to have little or no difficulty understanding stories involving sarcasm and metaphor (Karmiloff-Smith et.al. 1995). Again however, Borg presents evidence in support of her claims, eg. Bellungi et.al. (1993). What is clear is that there are many inconsistencies across various studies and laboratories and "it is difficult to find cases where specific brain regions have truly specific functions." (Prinz 2006: 24), which would cast in doubt upon i, vii and viii above. Prinz systematically casts doubt on each of the nine criteria listed above.

For Borg's argument against contextualism, from modularity, to work, semantics must be modular in this sense i.e. displaying criteria 1-9. If semantics is modular in this sense, then it is possible to have a *minimal* semantics which is context free, as the module would be encapsulated. And if this is the case then contextualism is false. So it appears that information encapsulation is the key feature in the argument against contextualism. There is however one way to understand encapsulation which would be amenable to a contextualist account. Carruthers (2006) distinguishes between a wide-scope and narrow-scope encapsulation. The former is the case when "the system is such that it can't be affected by most of the information held in the mind in the course of its processing", and the latter is the case when "concerning most of the information held in the mind, the system in question *can't* be affected by *that* information in the course of its processing." (Carruthers 2006: 16) Carruthers accepts the former case, where the system is such that it can only access limited information, as opposed to the narrow-scope encapsulation which divides all information available into information that can affect the system and information that can't. He claims, and I tend to agree, that the system would be a lot more frugal with the former case. There appears to be no reason to deny that any item of information could be such that the system could be affected by it vision couldn't

be affected by acoustical information, say. And if all the information available to the mind was sent simultaneously as input to any system, the system would be overwhelmed. (Prinz 2006) But as Prinz notes, this terminological distinction doesn't save encapsulation but rather "In recognizing the power of heuristic search, he tacitly concedes that the primary argument for encapsulation is unsuccessful." (2006: 33)

Borg's argument against contextualism falls with the rejection of narrow-scope encapsulation as a premise in favour of wide-scope encapsulation, since acceptance of wide-scope encapsulation is consistent with a functional-module conception of modularity (Carruthers 2006) which in turn is consistent with contextualism (Sperber 1994). In fact, "the idea that the mind consists of a large collection of functionally specialized mechanisms, or evolved modules (Barrett & Kurzban, 2006; Tooby & Cosmides, 1992)" which is "sometimes called the massive modularity hypothesis (Carruthers, 2005; Samuels, 1998; Sperber, 1994)" (Frankenhuis and Ploeger 2007: 688) is an important assumption that underlies most research in evolutionary psychology[65]. Evolved modules are generally understood as: "neurocognitive mechanisms specialized for solving particular adaptive problems that recurrently faced our hominid ancestors over evolutionary time." (Frankenhuis and Ploeger 2007: 688)

Sperber (1994) argues that these adaptive modules are informationally encapsulated, but the output of one module can serve as input of another. Sperber modules have "structural relations to other mental devices" as well as relations to the environment. "The content of a concept is not an intrinsic but a relational property of the neural realizer of that concept, and is contingent upon the environment and the history (including the phylogenetic prehistory of that neural object. This extends straightforwardly to the case of domain-specific modules. A domain is semantically defined, that is, by a concept under which objects in the domain are supposed to fall. The domain of a module is therefore not a property of its internal structure (whether described in neuorological or in computational terms)" (Sperber 1994: 51) So Sperber modules are domain specific but the domain is defined by the concept which in turn is environmentally determined. Further, these modules can access information which has been processed by other modules, or the output of other modules. These modules can answer to what Sperber believes are the main arguments against massive modularity: (a) Integration of information: i.e. we can have many different kinds of percepts related to a single concept (e.g. see, hear, touch and smell

[65] An approach ''informed by the additional knowledge that evolutionary biology has to offer, in the expectation that understanding the process that designed the human mind will advance the discovery of its architecture'' (Cosmides, Tooby, & Barkow 1992: 3)

a dog); (b) cultural diversity: our concepts range over a multitude of domains whose content varies from culture to culture. Within Sperber's theory, there is scope for integration of information, as modules can access information already processed by other modules. But more importantly integration takes place through the metarepresentational module, whose domain includes all representations available to any module[66]. This module can therefore combine various modes of information to yield a single concept. Cultural diversity is not a problem for Sperber as his modules are related to concepts which are environmentally determined. Thus Sperber, through the inclusion of a metarepresentational module, and by making outputs of one module available as input for another, is able to argue for (massive) modularity within a contextualist framework.

In this section we saw that the argument from modularity levelled against contextualism has its drawbacks: one can either argue against the Fodorian modularity that is assumed in the modularity argument against contextualism (Prinz 2006, Robbins 2007[67]), or argue for a different conception of modularity which is consistent with contextualism (Sperber 1994). For a defender of the argument against contextualism (Borg 2004,2012), one way to go would be to argue for the need for narrow-scope encapsulation or the idea that most of the available information is such that the system cannot be affected by it, and not just a wide-scope encapsulation which is acceptable to defenders of contextualism. Borg achieves this through dividing the information available into that which can affect the meaning of a word and that which cannot. Meanings of words are atomic and cannot be affected by most of the available information, although there is non-meaning constitutive information which is lexically marked. To see how she achieves this, we will first consider an underlying motivation for this view: Fodor's representational theory of mind.

[66] One could ask here whether the very postulation of this metarepresentational module amounts to a rejection of massive modularity. The answer depends upon the concept of massive modularity under consideration. If we are to follow Fodor's understanding of modularity, then the features (iii)-(vi) from section 3.1 would directly conflict with the postulation of a metarepresentational module. However, having rejected Fodor's version of Modularity, Sperber is able to make this postulation without contradiction. His modules are encapsulated though the output of one is available as input of others. So there is no theoretical contradiction in claiming that there could be a module, so encapsulated, which can access the outputs of many different modules, as inputs for its internal processing.

[67] Though Robbins 2007, as we noted above argues against Borg's use of Fodorian modularity to characterize the semantic component of the language faculty. The particular arguments he mentions are answered by Borg, as we will see in section 5.

4.3.2 Fodor's Representational Theory of Mind

Borg takes Fodor's particular form of modularity of mind as a premise for her critique of contextualism. Modularity is a central part of his version of a representational theory of mind (RTM), which maintains that:

a. Psychological explanation invokes laws about relations between mental states.
b. Mental representations (and not ideas or images) are the primitive bearers of intentional content.
c. Thinking is (and not all causal processes are) computation
d. Meaning is information.
e. What distinguishes coextensive concepts is in the head.

There are a few points to note here: (a) importantly notes that relations between mental states are *governed by law*; in (c) Fodor isolates thinking, as opposed to other causal or cognitive processes, as being computational in character. Further (b) allows Fodor to maintain that mental representations are both abstract (given their role as *bearers* of content, and therefore as distinguished from the particular content) and particular (taking into account the particular content that they bear). They are abstract, since they are in a sense placeholders for (potentially different) intentional content; in other words, to the extent that they are distinguished from the particular content, they are abstract. They are particular since at any given moment, they would have particular content. This relates and leads to (e) when we consider (b) in conjunction with the less controversial idea that coextensive concepts can be distinguished intentionally. So, being the bearers of intentional content, it appears that what distinguishes two coextensive concepts are these mental representations themselves. Thus, Fodor concludes that: modes of presentation (MOPs – which are supposed to distinguish coextensive concepts) are mental representations. Fodor's MOPs, like Fregean MOPs, distinguish coextensive concepts; but unlike Fregean MOPs, they are not senses. They are also both abstract and mental, as noted above, and the "apparent tension is reconciled by taking MOPS-qua-things-in-the-head to be the tokens of which MOPs-qua-abstract-objects are the types." (1998:20) For Frege, senses were abstract and objective: abstract because "the sense of the sign" is that "wherein the mode of presentation is contained" (Frege 1993: 24); and objective because MOPs are objective ways of defining or designating the same referent.[68]

[68] "Somebody observes the Moon through a telescope. I compare the Moon itself to the reference; it is the object of the observation, mediated by the real image projected by the object glass in the interior of the telescope, and

By regarding MOPs as both abstract and mental, Fodor appears to be conflating the distinction between 'ways of conceiving' an object (an internal process) and 'ways in which the object is presented to us' (an external fact). There are certain important consequences of this conflation. Firstly, it conflates the mental and the physical, or claims about our minds and claims about physical objects (or the way that the object is presented to our mind). Secondly, if MOPs are mental particulars, it amounts to claiming that for a particular mode of presentation, there can be only one way of conceiving the object. There appears to be no particular reason why there must be only a single way of conceiving any particular mode of presentation.[69] .

Fodor comes to thinking about MOP's as mental particulars through the following reasoning: in order to individuate concepts, MOPs must correspond one-to-one with concepts. The only way for them to do so is if they are mental. One way to sympathize with this view would be to understand the MOP in question as something like the meaning of a linguistic expression as generated in the mind on occasion of an act of linguistic comprehension and therefore as a mental particular. Here it does not make much sense to distinguish between the way of conceiving an expression and the way that it is presented as both here would refer to the meaning generated, the fact being that meaning is necessarily generated only in the mind. This pushes us towards an internalist conception of meaning.

Fodor however, went further and claimed that MOPs are rather mental representations. With the notion of a mental representation comes the question what this representation is a representation of. The notion of a mental representation appears to inherently support a form of externalism: a representation is a representation of something external to itself. Let us try to grasp how he tries to maintain his attachment to externalism through exploring his claims about concepts.

In *Concepts* (1998), Fodor manages to combine three distinct positions: atomism, realism and nativism. We will consider Fodor's atomism and realism in the rest of this section. Nativism for Fodor is not of concepts but of mechanisms required for the acquisition of concepts, while realism is of object, and atomism applies to concepts in the way explained

---

by the retinal image of the observer. The former I compare to the sense, the latter is like the idea or experience. The optical image in the telescope is indeed one-sided and dependent upon the standpoint of observation; but it is still objective, inasmuch as it can be used by several observers. At any rate it could be arranged for several to use it simultaneously. But each one would have his own retinal image." (Frege 1993: 27)

[69] For instance, in Frege's example mentioned in note 1 (Frege 1993:27), the same real image in the interior of the telescope can give rise to different retinal images.

below. To attribute atomicity to something is to assert that it cannot be decomposed into parts. Thus to say that concepts are atomic implies that they cannot be decomposed into other concepts, the acquisition of which would result in the acquisition of the concept. So according to Fodor's form of informational atomism, there are no concepts the possession of which is metaphysically necessary to have the concept DOORKNOB. Rather, "what makes something the concept DOORKNOB is just: expressing the property that our kinds of minds lock to from experience with good examples of instantiated doorknobhood." (Fodor 1998: 137) He clarifies that these "good examples" are the typical doorknobs or instances of the doorknob stereotype. Further, "with *doorknob* as with *red*, all there is to being it is how things tend to strike us" Further "if a property is constituted by the way that things that have it strike us (under certain circumstances), then being locked to the property requires only that things that have it do reliably strike us that way (under those circumstances)." (Fodor 1998: 141)

Thus the physical object that is a doorknob has the property *doorknobhood.* This is understood as a single atomic property. Secondly, this property, doorknobhood, must be a real object, as it were, or exist in a real object, to which we can stand in a causal relationship. Thirdly, what constitutes *doorknobhood* is how objects that are doorknobs tend to strike us. So here, a real property is constituted by a mental representation. Lastly, acquiring the atomic concept DOORKNOB or being locked to the real but mind dependent property of doorknobhood requires only that the real object *doorknob* does reliably strike us as it is, i.e. as containing the (mind dependent) property. So what emerges here is that properties as well as concepts are for Fodor mind dependent. But he tries to merge this (internalist) mentalistic picture with realism about objects. His argument for this is of the following form:

a. Doorknobs, or Tuesdays, are mind dependent.
b. Therefore, if there are no minds, then there are no doorknobs.
c. But there are minds.

Now step (d) or the conclusion of this argument can only be: There may or may not be doorknobs or Tuesdays. To claim otherwise would be to commit the fallacy of negation of antecedent. However, Fodor claims that "doorknobs are real but mind dependent" (Fodor 1998: 150). This clearly goes beyond the premises just outlined. In fact, if it was a premise then it would follow that there are minds. But the other way round does not hold. Thus realism cannot follow from his argument and therefore must be part of the premises or already assumed in the argument. For Fodor, "Doorknobs are constituted by their effects on our minds, and *our minds* are in the world." (Fodor 1998: 149) And these premises lead him to the conclusion that

doorknobs are real. One way to construe this alleged reality of doorknobs would be to restrict it to the mental: i.e. to say that they *are* real and in the world, as the mind that constituted them is real and in this world, but they exist within the real mind. In other words they are as real as anything within a real physical object. To put it differently, the concept DOORKNOB would be in a sense as real as the object doorknob, as both are constituted by and depend for their existence on the mind. If this is the case, then there is no particular advantage in talking about mental as opposed to physical objects and any metaphysical reality whatsoever. We would here again be within internalist territory, and the externalist dimension of the representational picture under scrutiny proves to be no substantive addition to what the internalist can claim as well.

Fodor's conception of natural kind terms, on the other hand, is not conducive to an internalist explanation. "All the metaphysics of concepts possession requires, of our concept WATER or Homer's, is being locked to water. If you are locked to water our way, you have the concept WATER as a natural kind concept; if you are locked to concept WATER Homer's way, you have the concept WATER, but not as a natural kind concept... if you are locked to water either way, you have the concept WATER." Fodor 1998: 157) Although all that is required to have the concept WATER is being locked to water, however, the way that we are locked to the concept determines whether or not we have the concept WATER as a natural kind: only if we are locked to the concept WATER *our way*, or through the scientific concept H2O, do we have the concept WATER as a natural kind concept. This suggests that anyone who does not have the concept H2O does not have the concept WATER as a natural kind, and then having the WATER concept would be like having other concepts which are not of natural kinds (and therefore mind dependent as we claimed above). But Fodor argues that there are no concepts whose possession is metaphysically necessary for having WATER as a natural kind concept. This, he claims, is because science discovers essences and doing science thereby links us to natural kinds as such. So in the case of Homer, where these essences were not yet discovered, though one was linked to water, the water he was linked to was not water as a natural kind substance. We therefore need the essences that science discovers, in order to form any link to it. Even if we grant him that in today's world, our concept WATER is a natural kind concept, this means that WATER links us to a natural kind: H2O. But the fact that water really exists as a natural kind is just presupposed in his argument[70]. Following Frege, however, one

[70] Frege for instance argues that "when we say the Moon, we do not intend to speak of our idea of the Moon, nor are we satisfied with the sense alone, but we presuppose a reference... Now we can of course be mistaken in the

could argue that"in grasping a sense, one is not certainly assured of a reference." (Frege 1993: 24). But grasping WATER as a natural kind concept as Fodor does, already includes assurance of reference. This is because in grasping the concept as a natural kind concept, we are understanding our concept WATER not as being constituted by the effect of certain impressions on our mind, but rather by the relation to a natural kind: the physical substance water or H2O. So it is not the case of the argument presuming the reality of a physical substance, but the concept is in fact claimed to be *constituted* by the real physical substance. For Fodor, it is science that gives us the essences, the physical substances, and in acquiring concepts we form a relationship to these essences which science discovers.

What is important for acquiring the concept WATER as a natural kind, is that there are these essences to which natural kind concepts are constitutively linked, even though the scientific descriptions of these essences may change(based on new discoveries).[71] In other words, though it is necessary that water is H2O or 2 is a prime number, these are empirical facts, and not constitutive of the concept WATER or the number 2. We get a clearer sense of this idea in Fodor (1998: 74)[72] where he claims that although it is necessary that water is H2O, this is a fact about the *world*. What is, however, constitutive of WATER is the *substance water*, given that it is a natural kind term.

This is clearly not, for Fodor, merely the stipulation that WATER refers to *water* (which would be a deflationary statement) but rather, what it refers to is always identified by by an essence, or a scientific description in case of H2O, independently of whether these descriptions

---

presupposition, and such mistakes have indeed occurred. But... in order to justify mention of the reference of a sign it is enough, at first, to point out our intention in speaking or thinking. (We must then add the reservation: provided such reference exists.) (1993: 27-28)

[71] "The problem for informational semantics comes not from intuitions that the connection between *being Tuesday* and *coming before Wednesday* is *necessary*, but from intuitions that it's *constitutive* in the sense that one can't have one of the concepts unless one has the other. Compare *water is H20* and *two is prime*. Presumably though both are necessary, neither is constitutive. Accordingly it's possible to have the concept WATER but not the concept HYDROGEN, and its possible to have the concept TWO but not the concept PRIME. All of this is perfectly OK as far as informational semantics is concerned. It's perfectly consistent to claim that concepts are individuated by the properties they denote, and that *the properties* are individuated by their necessary relations to one another, but to deny that knowing about the necessary relations among properties is a condition for having the concept". (Fodor 1998: 74)

[72] "The problem for informational semantics comes not from intuitions that the connection between being Tuesday and coming before Wednesday is necessary, but from intuitions that it's constitutive in the sense that one can't have one of the concepts unless one has the other. Compare water is H20 and two is prime. Presumably though both are necessary, neither is constitutive. Accordingly it's possible to have the concept WATER but not the concept HYDROGEN, and its possible to have the concept TWO but not the concept PRIME. All of this is perfectly OK as far as informational semantics is concerned. It's perfectly consistent to claim that concepts are individuated by the properties they denote, and that the properties are individuated by their necessary relations to one another, but to deny that knowing about the necessary relations among properties is a condition for having the concept".

are actually formulated or not. If we are not locked to the concept WATER via the (scientifically described) essence then we are not in possession of WATER as a natural kind concept. This at least compromises informational atomism, as the concept WATER is no longer atomic but is rather identified through scientific descriptions. The point is that, even viewed as a natural kind, the kind intended cannot be atomic if it is identified by an essence We have seen in this section that much of Fodor's version of a representational theory of mind is capable of an internalist interpretation. We have found nothing in his arguments to preclude such an internalist interpretation, apart from the stipulation regarding natural kind terms being constituted by the natural substance, and we noted that his position on natural kind terms compromises his informational atomism. In the next section we shall see how Borg applies these Fodorian commitments to what she calls 'Organisational' lexical semantics.

4.4 Borg's Organisational Lexical Semantics

Fodor, as we noted above, combines nativism, atomism and realism. We also saw that his argument for natural kind terms may be rejected without affecting the rest of his theory. So his attachment to realism may be considered mild, as the realism he endorses is a realism of other minds as opposed to objects (barring natural kind terms). Words for him refer to concepts rather than objects and these concepts are atomic in nature. So word meanings are not a collection of properties/ uses, or even one typical property or use (as in Horwich 1998 for instance). Rather meanings are concepts which are atomic. So there is no particular property which makes the concept DOG, nor even a bundle of features (for instance +ANIMATE, +MAMMAL, +FOUR LEGGED etc.). DOG is conceptualized and understood as a single atomic concept. So there are no particular properties nor any bundle of properties, which are necessary.

In arguing against lexical semantics, Fodor and Lepore (2002) insist that accounts positing lexical complexity fail the compositionality constraint that they impose on any semantic theory. If 'dog' is analysed into features, one of which is, say, +TYPICAL PET, then that cannot contribute to expressions like 'wild dog'.]. In other words, if the meaning of 'dog' was understood as having +TYPICAL PET as a feature, it could not combine with 'wild' compositionally without contradiction. The point is that this feature does not enter into the compositional determination of complex meanings.

Borg, it seems, agrees with the criticisms of lexical semantics but points out that they only apply to a lexical complexity argument that views words as *constituted* by features or

properties. For Borg, meanings are atomic, but she also wants to allow for some form of lexical complexity. The way to allow for lexical complexity as well as atomicity without contradiction is, Borg insists, through "weakening the appeal to lexical complexity still further, taking the additional information required to carry the internal burden on semantics to be constitutive of neither concept content nor concept possession. Perhaps (following Fodor's 1998: 65 suggestion above) we could simply allow that the lexicon contains additional information above and beyond a mere list of word-denotation pairs but maintain that this additional information doesn't constitute the meaning of a word nor does it impose any conditions on the possession of that word's meaning, instead it imposes constraints on our competence with the word." (Borg 2012:193)

What Borg is suggesting is again a form of competence-performance distinction, though it is interesting that she calls the performance side of it *competence.* This usage is valid within her assumptions as simply possessing the atomic concept does not allow us even the *ability* to use the concept. The traditional competence/ performance distinction, as for instance for Chomsky, was however the distinction between 'knowledge of language' and 'use of language' (Chomsky 1965). What we seem to need minimally, in order to perform/use the concept, is the kind of syntactic information Borg puts within the extended lexical information not included in knowing/understanding/having the concept, though needed for full competence. So for her knowing the concept is distinguished from knowing the language, in that it appears that one may have concepts in isolation from knowing the language. She claims that word meanings or concepts should be explained referentially, but there are non-meaning-constitutive features which are lexically specified. Although these do not constitute the meaning, this information is contained within the "organization of the lexicon". Thus the name *organizational* lexical semantics.

> "So, for instance, a noun like 'dog' could be marked as, say, +AGENT, + ANIMATE, but these properties of the expression do not constitute part of the word's meaning, instead indicating the *kind* of expression the word is by revealing the categories into which it falls." (Borg 2012:194)

It needs to be made clearer which word varieties have atomic meanings: is it the root word or each form that has the atomic meaning? Given the direction of the theory, Borg could allow only the root word as having the atomic meaning, which is then shared by other forms of the word, though in each form it is combined with different non-constitutive features which

are lexically specified. So for instance: she may claim that *develop* and *development* share a core meaning, although lexically they fall under different categories, that of verb and noun respectively. This would avoid the problem of creating multiple meanings for related words.

So the idea is that "simple words have simple meanings whilst allowing that the framework into which they slot is itself a complex, cross-categorising system and thus that additional lexical complexity emerges at this point (external to the meaning of a word)…, [T]he additional information emerges as part of the organisational structure of the lexicon, telling us something about the meaning of words but where this information is in addition to, and extraneous to, the meaning of a particular linguistic item." (Borg 2010: 68)

So effectively she is making a distinction between

i. knowing the meaning: having the atomic referential concept,
ii. knowing the meaning as related to knowing the language: the logical form, or where the word should be syntactically placed,
iii. and knowing the language: having the ability to form novel sentences, produce and understand an infinite number of sentences, and do so systematically.

A way to argue against this position would be to show that the three-fold distinction doesn't exist: It is possible that any one or all of these steps are a false move, or that all three are inextricably linked, or (i) and (ii) merge, or (ii) and (iii) merge. It is for instance plausible to argue that (ii) and (iii) are both a part of syntax and (ii) does not provide semantic features which affect syntax as Borg argues. There is thus no semantic level information which is part of the process, as for instance in Hinzen (2007). On the other hand, one could also conflate (i) and (ii) and argue either that a concept is itself complex or that referential information is also lexically marked. Borg herself raises a similar objection as follows:

> "This objection emerges in two ways: first, why think the syntactic behaviour witnessed requires semantic level explanation at all (as Fodor and Lepore note, perhaps the recognition that 'devour' but not 'eat' requires an object is a purely syntactic realisation, with no semantic underpinning at all)? And, second, how do we know which elements of world knowledge deserve representation at the lexical level – why should the fact that kicking involves a manner of motion be lexically represented but not, presumably, the fact that being kicked is often painful?" (Borg 2012: 197)

Her answer to the first is essentially that we should not here resort to 'syntactic eccentricities' as this syntactic behaviour needs explanation that accounts for semantic similarities and differences. For the second she answers that there may not be a systematic way of determining why certain features are lexically marked other than the idea that "It is the realization of these features within the lexicon which explains how and why they come to have an impact on the syntactic behaviour of expressions." (Borg 2012:199) So the only possible way to determine which features should be lexically marked is through their effect on syntactic behaviour of the expression; in other words, the features that have syntactic effects are the ones that should be lexically marked. This explanation however, does not help us to identify those features which are lexically marked from those that are not, and the problem remains.

Even if we are able to identify these features via the syntactic difference they make, it remains unclear why we would understand them as constituting information which is non-syntactic. If these are features that make a syntactic difference, why is it that we cannot simply say that these constitute syntactic features are lexically marked? Thus, one could, as noted above, still argue that there is no semantic level of information involved.

We saw in this section that Borg attempts to argue for both atomism and a form of lexical complexity, and the way she purports to overcome the contradiction is through distinguishing between the meaning of a word, which is atomic, and the non-meaning-constitutive organizational information which is lexically specified. We also noted that it is possible to argue against this by pointing out that there is no semantic level of information here. If there is semantic information which is lexically marked, and is not such that it has syntactic effects, then there would be no way of identifying it, within Borg's framework.. Borg remarks that features which are syntactically relevant are the ones which are lexically marked. Even if we are able to identify these features (though no systematic way of such identification is available), it would not preclude the argument that there is no semantic level of information which is lexically marked. So reconciling Fodorian atomism with lexical complexity again leaves room for an internalist interpretation. Borg's externalism however is also the result of extending Fodor's realism from natural kind terms to all concepts, so that they are not just mental representations but also *really* attach to objects of reference. In the remainder of this chapter, we will consider Borg's externalist referential commitments, after first considering the underlying assumption of minimal semantics: propositionalism.

4.5 Extending Fodor's Commitment Beyond Natural Kind Terms: Defense of a Referential, Propositional Account Against Internalism

We have in the preceding sections dealt with point (A) from the introduction: arguments for and against contextualism, noting how minimalism can be defended from contextualist attacks. We pointed out that Borg's minimal account can survive even in the face of much neurological evidence for contextualism. We will in the rest of the chapter consider (B), (C) and (D) from the introduction: a defense of a propositional account of semantic content, which involves the claim that this content is referential, and consider internalist arguments against it. In this section we shall consider these points and how they are related to each other and the preceding discussion of (A).

Propositionalism forms a key assumption in minimalism, as it should be evident through the discussion in this chapter, and therefore it can be argued that minimalism depends on the truth of propositionalism. Propositionalism constitutes the "the idea that every well-formed declarative sentence expresses a complete proposition relative to a context of utterance." (Borg 2012: 73) This forms the underlying assumption in both the minimalist positions we have considered: Borg and C&L. C&L however do not provide any explicit defense of propositionalism, although the assumption of propositionalism is explicit in their works (2005a, b). They argue for instance that an utterance of 'Rudolf is a reindeer' is "true just in case Rudolf is a reindeer, and expresses the proposition *that Rudolf is a reindeer*" (2005a: 3). C&L, however, and importantly, point out in a footnote that "Semantic Minimalism, as understood in this book, need not take a stand on whether semantic content is a proposition, or truth conditions, or what have you. Throughout the book we try to remain neutral by couching the issues both in terms of truth conditions and in terms of propositions." (2005a:3, footnote 3) So C&L try to distance themselves from the metaphysical status of semantic content.

They do however recognize that their theory could raise metaphysical questions: for instance – what is the minimal proposition *that A is red*, or what state of affairs does it pick out, or what do its instances have in common. C&L insist however that the answer to these questions is just that A is red (2005: 158). They argue that these questions are themselves metaphysical and non-linguistic, and therefore do not have anything to do with semantics. What makes something an instance of red or of dancing is not a question that either arises or is answered by contextualism about red or dancing. (2005: 161) To put it differently, whether one

endorses propositionalism or contextualism, questions of metaphysics arise irrespectively, and need metaphysical and not semantic answers. C&L argue further that in case someone is not happy with the response that what instances of red things have in common is that they are red simpliciter, they must logically question "the very idea of two or more things sharing properties or engaging in the same activities." (2005: 163) C&L are, in their own words are "trying to push all these worries onto the metaphysicians. Whatever story in the end they settle on about what makes something red, that's what it takes to satisfy the semantic truth conditions of 'A is red.'" (2005: 164)

Consider here the discussion in Travis (2006) regarding the determination of truth conditions for 'Pia dances', mentioned in section 2.1 above. Consider the utterance 'Pia dances'. Travis (2006) argues that understanding the truth conditions of this sentence requires much more information than is provided by the proposition *that Pia dances*. We need rather to know what scenario would cause the speaker/hearer to apply the predicate 'dances' to Pia: does she need to be skilled, or simply eager etc., and this information is not contained in the proposition expressed by the uttered sentence. So the uttered sentence does not give us any determinate proposition, i.e. a proposition that has determinate truth-conditions. Borg (2004) argues against this kind of argument that this appeal to indeterminate truth conditions is just another case of speaker meaning as opposed to sentence meaning, and this sort of truth evaluation is a pragmatic matter (2004:256). Minimal propositions give us truth conditions, but these truth conditions need not be of the kind that enables us to pick out the world in which the sentence is true. For Borg: "'London buses are red' means that *London buses are red*. It is not expected to tell us why 'red' means *red* and not *blue*, nor is it required to guarantee that every subject who grasps this semantic information knows what is required to make something a bus, nor must they be able to tell, for every object they encounter, whether or not it is a bus" (Borg 2009: 39). This explanation brings us to point (C) above: an explanation of the referential nature of semantic content. It appears from this example that Borg-type truth conditions do not require us to be able to pick out either the world in which the utterance is true, or the object of reference, despite the fact that they are said to involve a referential relation. The way this relation is established is explained as follows: "There is a discrete language faculty, containing specialised bodies of knowledge and operations on that knowledge, dealing with phonetics, orthographics, syntax and semantics. This vast body of knowledge would make use of the common currency of thought (i.e. language of thought expressions) but in a constrained, or encapsulated way. So, for instance, the lexical information that 'cat' means CAT (where the capitalised item here

indicates an expression in the language of thought, which I will take to be a concept) would be a proper part of the language faculty, and, since it seems right to think that the content of this concept or language of thought expression is determined by its connection to certain objects, or a certain property, in the world, word-world relations will (somewhat derivatively) be a proper part of the language faculty". (Borg, 2004: 84)

Thus, firstly, the language faculty consists of various interacting modules. We already discussed the Fodorian modularity she takes for granted (section 4.3), and the potential problems with it. We have also pointed out in section 4.3.2 that much of Fodor's representation theory of mind is conducive to an internalist explanation. Derivatively, and as it appears in this quotation, much of Borg's picture is also open to an internalist explanation: in particular, no elements of the external world need to be referred to in order to characterize the language faculty as such. Secondly, the syntax interacts with the conceptual system, and this whole is a part of the language faculty. This again is a picture akin to internalism. Thirdly, the content of a concept is determined by its connection to certain objects or properties. Thus the language-world relation is also a part of the language faculty, as per Borg 2004 above. Moreover, as we already noted in 4.3.1, concepts are for Fodor mental particulars, and are internalistically determined: They are constituted not by the object or property they refer to (except the case of natural kind terms), but rather by the effect that certain experiences (of instantiations of properties) have on our mind.

Before we proceed, it may be useful to rehearse the conception of internalism and externalism as developed in chapter 1. Chomsky as we noted in chapter 1 distinguishes I-language from E-language, and maintains that linguistic theory is a study of "what we may call 'I-language', where 'I' is to suggest 'intentional' and 'internalized'. The I-language is what the grammar purports to describe: a system represented in the mind/brain… and is in this sense internalized; a system that is intensional in that it may be regarded as a specific function considered in intension – that is a specific characterization of a function – which assigns a status to a vast range of physical events" (Chomsky 1997: 10). E-language, on the other hand, is what philosophers have traditionally understood 'language' to refer to: "'E' is intended to suggest extensional' and 'externalized'. The definition is 'extensional' in that it takes language to be a set of objects of some kind, and it is 'externalized' in the sense that language, so defined, is external to the mind/brain" (Chomsky 1997: 7). Thus Chomskian internalism is distinguished from externalism about language through a restriction in the domain of inquiry: Chomsky

wants to restrict himself to studying brain/mind-internal structures, what he calls elsewhere (2000) the language organ or the faculty of language, as opposed to a conception of language which is external to the mind/brain. Borg's domain of inquiry is clearly what Chomsky (1997) calls E-language. As we noted in preceding discussion, for Borg, words are related (through concepts) to objects in the world. She argues that internalism claims either of the following:

(a) There are no words and objects.
(b) These do exist, but as mind-dependent entities.

However, internalism does not make claims about the existence or non-existence of words or external objects. It is a methodological decision and program for analysis in the empirical cognitive domain, not a metaphysical claim. It consists in the decision to investigate the empirical properties of mental states without reference to the external environment, or that what is mind internal is not individuated by what is external to the mind. Thus (a) would not be endorsed by internalists in the first place. Borg argues against (b) that it 'conflates questions of concept content with questions of how we identify objects as falling under those concepts'. However, as just noted, internalism does not seek to address the issue of existence of objects at all (whether they fall under a certain concept or not). Rather it is a methodology for studying what is mind-internal without any attempts of reducing all that is external to mind internal entities. In fact it does not take any metaphysical position on whether or not there are external things at all. Importantly, moreover, Borg's argument against (b) appears to be in conflict with the Fodorean position which Borg (2004) endorses, as Fodor himself claimed that most concepts other than natural kind concepts exist as mind-dependent, as noted in the above section.

To elaborate on her argument against (b) above: Borg argues that the reason why we must admit that word meanings are mind-independent is that there is an element of 'correctness' in the use of language (as mentioned above). "So when someone learns that 'contract' means *mutual agreement* rather than *written agreement*, they take themselves to be corrected about the meaning of the term, yet this behaviour only makes sense given the externalist perspective that what matters for word meaning can lie outside the individual." (Borg 2009: 46) Here Borg appears to endorse a non-atomistic decompositionalist position on lexical content in order to ground her externalism. However, if words are atoms, as she argues, then the word in question *automatically* cannot mean either of these things. Notwithstanding the internal inconsistency within this argument, it may still work against internalism. But

internalism does not deny that word meanings are learnt. The child must acquire the lexicon of a particular language through, and in agreement with other speakers of the language, in order to use it in that community of speakers. We will now consider Borg's defense of propositionalism and whether there can be an internalist argument against it.

Borg (2012) defends propositionalism against what she considers the prime candidate in recent work for motivating a rejection of propositionalism: the phenomenon of "the apparent existence of well-formed sentences which apparently fail to express complete propositions prior to rich contextual input." e.g. "Jill went to a local bar [local to whom?]" and "The café is on the right [relative to what?]" (2012: 75) So her defense of propositionalism is the same as her argument *against* contextualism, and *for* minimal content, which we have already detailed in the preceding sections. She does not deal directly with the metaphysical issues involved with the assumption of propositionalism. Borg does however divorce metaphysical issues from epistemological issues (2001) claiming that in the case of singular terms, there is a distinction between what it is to be a singular term and how we can tell whether an expression is a singular term. Following C&L we can perhaps simply transfer these issues to the metaphysician, and claim that it is not a semantic issue.

There are however other problems which arise, apart from the metaphysical, context shifting and incompleteness problems already mentioned[73]. An internalist would argue that there are many sentences we cannot grammatically utter in language. Consider the following:

(13) The clown entertained the children

(14) The clown laughed the children. (cf. Hinzen 2007)

Though we can grammatically utter (13), (14) is an ill formed expression. There is a sense in which (14) could be understood as having a meaning: *the clown made the children laugh*. However, language and not the world constrains that possibility. Similarly, the following sentences (16) and (18) are ill formed, though they may be understood as having the meanings (15) and (17) respectively.

(15) He put the books on the shelf.

(16) He booked on the shelf.

[73] See section 4.2.1

(17) A cow had a calf.

(18) It cowed a calf. (Fodor and Lepore 1999)

In a logically possible language, (16) might have the meaning (15), and (18) might have the meaning (17). Though it is not in principle impossible to interpret (16) and (18), and it may be possible to imagine a possible world where they could be true. However, "Grammar forbids assigning them the interpretations indicated." (Hinzen 2007:145-146) As Hale and Keyser (1993) argue "[i]t is well known that a subject… cannot incorporate into the verb that heads its predicate" (1993: 60). So subject lowering is illicit and therefore (16) and (18) cannot have the meaning (15) and (17). Fodor and Lepore (1999) insist that what is precluded is the *derivation,* for instance of (18), from 'A cow had a calf', and not the primitive meaning of 'cowed' being 'cow had': "By definition, impossible-word arguments purport to explain intuitions of the form "there couldn't be a word w that means 'e'" by showing that 'e' couldn't be a derivational source for w. But, on the face of it, that doesn't show that there couldn't be a word that means 'e'; the most it could show is that if there is a word that means 'e', then it must be primitive… Assuming that 'a cow had' is a well-formed meaning, HK's sort of account does not explain why there could not be a primitive lexical item cowed that means 'a cow had'." (Fodor and Lepore 1999: 449) So the claim is: to say that (18) cannot be derived from (17) is not to say that 'cowed' could not exist as a primitive word meaning 'a cow had'. But this does not show why impossible words are impossible; Borg (2012) agrees with this argument and notes that her kind of lexical complexity is not the one under attack. We discussed in section 4 how Borg integrates Fodorian atomism and lexical complexity and saw also that this integration leaves room for an internalist interpretation. Now we may point out that it is precisely this integration that she appeals to in order to counter internalist arguments. Borg is able to argue that, "impossible words are impossible since they would require the lexicalization of structures which are in some way illegal" (Borg 2012: 180) or rather that "there might be a prohibition on words whose meaning would require categorisation in a way not permitted by the language" (Borg 2012: 201).

Borg also uses this thicker lexical meaning which admits of some classification in the form of non-meaning constitutive features, to explain other internalist challenges to a referential semantics. These challenges arise for instance through the following examples from Pietroski which seem require more than the syntax and simple denotational content of lexical items:

(19) John is too clever to catch a fish

(20) John is too clever to catch

(21) John is eager to please

(22) John is easy to please.

Despite the surface similarities, these sentences require very different treatment. Although in (19) and (21) John is the agent, this is not the case in (20) and (21). Pietroski (2005) points out that non-ambiguity also needs to be explained: we need to explain the negative fact of why, for instance, (22) can't mean that 'It is easy for John to please us'. Further: 'eager' and 'easy' differ semantically in an important sense, which cannot be explained simply by saying that 'eager' means EAGER and easy means EASY. Borg (2012) argues that such facts can be explained by her introduction of semantic features to the lexicon. But as we already saw in section 4.4, this moves opens her theory to an internalist interpretation rather that providing a solution to internalist challenges.

In this section we have considered ((B), (C) and (D) from the introduction: a defense of a propositional account of semantic content, considered the referential nature of this content, and a defense from internalist arguments against it. We saw that while for C&L propositionalism is an undefended assumption, for Borg it is inextricably linked to her argument against contextualism. Both Borg and C&L however, distance themselves from metaphysical claims: C&L through distinguishing metaphysical from semantic issues, and Borg through distinguishing metaphysical from epistemological issues. The latter move also leads Borg to maintain that although her account is a referential one, we may neither be able to pick out the world in which an utterance is true, nor the object of reference[74]. So reference remains in a sense (and ironically so) a theoretical claim. Furthermore, Borg's arguments against internalist challenges rests on her integration of lexical complexity and atomism[75], which can in turn be interpreted internalistically. What remains however, is Borg's contention that semantics without reference to the world is simply not semantics.

Borg clarifies however that "what minimalism is committed to is the idea that semantic contents are the kinds of things which can be true or false and to get this result what is needed is just some kind of commitment to weak externalism, construed simply, as Rey 2005 suggests, as the view that ascriptions of some content depend in one way or another on relations a

[74] See section 2.2 for a preliminary account of how reference is accounted for in Borg (2004, 2012)
[75] See section 4 for details of this integration.

computational system bears to things outside itself [not in one way or another, I would say, but on a way that is exactly specified by grammar (for example, causal relations are not enough)]. That is to say, minimalism assumes a lexical semantics which contains some kind of symbol-world connection. What I'll be arguing for, then, is more a denial of internalism than presenting the case for a specific kind of externalism." (Borg 2012:144, footnote 2)

So here it appears that she is arguing for an indeterminate kind of externalism, while at the same time, importantly distancing herself from internalism. But in case she merely wants to stress the need for a referentialist account, it does not seem necessary to argue against internalism per se: for there are internalist strategies which can capture reference to some degree (Hinzen 2007 for instance)[76]. We may note here however, that in the weakened claim above, she sets herself the lesser task of arguing against internalism, and not at the same time defending externalism. Thus we may grant her an indeterminate symbol-world relation which common sense takes for granted. But in our discussions above we noted that it is precisely her argument against internalism which are not convincing. And much of her position including the Fodorian assumptions can be interpreted internalistically. Thus there appears to be no reason why a minimal semanticist cannot be an internalist. Borg's externalism does not inform any of her arguments or explanations, and remains a stipulation, although a central one.

Consider the sentence "Bill is ready". For Borg (2012: 202) "'ready' always means *ready*, the content of this expression is exhausted by the property in the world to which it refers. However, the term 'ready' gets characterised within our lexicon as belonging to a set of two-place adjectives, a set of terms which requires both a subject and an object." Thus, following her externalist referential stance, the meaning of 'ready' is simply denotational and refers to a property in the world. Being a property, it necessarily requires a subject, but not an object. The apparent incompleteness is addressed by the organisation of the lexicon: 'ready' is lexically marked as a two-place adjective which requires both a subject and an object. The reason why taking two arguments is not part of the constitutive (as opposed to organisational lexical) meaning is, Borg insist, because meanings are denotational.

Besides this stipulation, there appears to be no reason why taking two arguments should not be a part of the meaning of 'ready' as opposed to forming a part of the organisation of the lexicon. In fact, we intuitively recognise ready as needing both a subject and a particular object.

[76] Infact, Hinzen (2007) argues that if you want to capture what is humanly specific about reference (which is almost all of it), then the externalist elements simply do not enter in any explanatory role.

As Collins 2013 points out "*Bill is not ready* has the natural reading where whatever it is that Bill is not ready to do is construed as wide-scoping over negation, that is, one is saying that Bill isn't ready to do some particular thing, not that Bill isn't ready to do anything at all. Thus, it seems as if bare *ready* is not incomplete, but obligatorily existentially definite as regards its complement position" (2013: 154) Borg seems to recognise this and argues that "information associated with the term itself tells us that the logical form of sentences containing 'ready' *must always* have the form __ready__" (Borg 2012:202 emphasis added) But she explains this understanding of 'ready' through "information associated with the term" as opposed to the meaning of the term. This lexical information, Borg claims, contributes to the meaning of 'ready' in terms of how it behaves syntactically for example, but it does not *constitute* the meaning, which is denotational. However, if Borg's minimal propositions are sensitive to the kind of organisation that syntax provides, one argument against this could be that syntax provides too much structure (Collins 2007[77]), much of it unnecessary and unhelpful towards providing the truth conditions for the sentence (e.g. structures generated by *movement* and *copying*[78]). On the other hand, given that (perceived) incompleteness is a criterion for Borg to introduce lexical complexity, 'ready' and 'tall' for instance would not fall in this category, as these are not understood as incomplete, but rather as requiring two arguments: i.e. someone is ready for/to do something, and someone is tall with relation to a class. In other words, Borg does not provide enough/clear grounds for introduction of lexical complexity for certain terms. She argues that the introduction depends on the syntactic effects of purported incompleteness. However, as we just noted, these terms are not perceived as incomplete, and if syntactic behaviour is what we must go by, then there is too much information that syntax provides, and for all terms, which would need to by lexically marked. That 'ready' means *ready* does not seem to have any explanatory role in this argument. We note here that explaining and externalist stance may not be what Borg was after to begin with, as externalism and denotation forms one of her axioms. However, what interested us in this work was a possibility that an externalist referential account, involving minimal propositions, could answer to our original intuitions requiring both an internalist and externalist explanation. Borg, as we now see neither provides such *explanation* nor perhaps aims to provide it. We were seeking an explanation for what was to begin with, and remains, a stipulation, or rather an axiom within her work. We therefore cannot retain anything from this position.

[77] Here he argues against the understanding of propositions as that what syntax offers.
[78] See for instance Collins 2007 for more detail.

4.6 Conclusion

We set out in this chapter to seek a minimal referential semantic theory. We began with two main contenders: Borg (2004, 2012) and C&L (2005a,b) and noted that in order to defend their positions they would need to argue against contextualist and internalist[79] challenges, argue for propositional/truth conditional semantics, and explain referential relations. In section 4.2 we noted the minimalists argue that the truth-conditionally relevant contextual elements find their source in syntax (hidden variables for Stanley 2000 or the class of demonstratives for C&L 2005a,b) or can be interpreted in syntactic terms (Borg 2004). For instance, 'you' is understood as the contextually invariant form: 'the addressee of the sentence'. We saw also, that neither the contextualist arguments nor the empirical evidence discussed affect the minimal compositional core of meaning. In section 4.3 we considered one of Borg's main arguments against contextualism, namely that modularity is desirable as a theory of mind, and that contextualism cannot accommodate modularity; therefore, contextualism must be discarded. We saw however, that this argument is flawed in various ways. We saw furthermore that much of the Fodorian position Borg assumes is amenable to an internalist explanation. In section 4 we argued that Borg extends Fodor's commitments through extending his realism of natural kind concepts to all concepts, retaining his atomism while introducing a form of lexical complexity. This integration of atomism and lexical complexity, however, shifts Borg's position more into internalist territory. In section 5 we noted that Borg's defense against internalist challenges rests on her integration of atomism and lexical complexity, and this is itself conducive to internalism. Thus we can conclude that, in case we can have a minimal semantics, there is no reason why it must be externalist, or in other words, our discussion so far points to the possibility of an internalist minimal semantics.

There appears to be no argumental necessity for endorsing externalism which arises from our discussion of minimal semantics. So we may accept all of the following claims we discussed as belonging to a minimal semantic position: there is a minimal semantic level of interpretation prior to contextual interference. At this minimal level, we may not be able to identify the object of reference, though through the notion of character, we have a bridge between the linguistic and other cognitive domains which help attach a concept to the descriptive understanding of for instance 'addressee of the sentence'. Hence, through the interaction of cognitive domains we are able to pick out the object of reference. But through

[79] Task specific to Borg's version of Semantic minimalism.

all these steps we remain within mind-internal descriptions. The addition that allows the theory to retain its externalist character is the idea that the content of a concept is, as Borg maintains, the object itself. However, nothing within the structure of the theory necessitates this addition. Thus, if we take reference not merely as a primitive (as Borg does), but assume that we need to explain its human-specific features, all the explaining factors are internalist ones. As we noted, Borg's externalism does not inform any of her arguments or explanations, and remains a stipulation, although a central one.

4.7 Consequences Towards Our Original Intuitions

We saw in this chapter that Borg's theory did not turn out to account for both externalism and an argument against contextualism. This leaves open at least one of the two following possibilities: either externalism is not needed for arguing against contextualism, and we could thereby follow the line of internalists; or alternatively, a full blown externalist position could account for the intuition that an element of use creeps into what constitutes a proposition. However, both of these alternatives answer to only one of our two Wittgensteinian intuitions (we started out seeking a theory which accounts for both a content internal understanding of meaning, as well as the aboutness of language). Another option available to us here is to reinterpret the internalist/externalist divide. This is what we will attempt in the next chapter. We will attempt to formulate an account that could potentially answer to both of our initial intuitions, and be in line with our re-interpretation of Wittgenstein: that an element of use is already inherent in what constitutes a proposition. Thus, although we cannot retain anything from Borg's position, it does point to the kind of alternative we are seeking to answer our original intuitions.

Chapter 5: A Unified Theory and its Application

5.1 Introduction

The debate between internalism and externalism arises, as we noted in chapter 1, with respect to varied domains of inquiry. In the philosophy of mind and language, the debate arises with respect to the individuation of linguistic meaning, mental content, and propositional attitudes (Goldberg 2007) We have been concerned particularly with content internalism and externalism in this thesis.[80] As Lohndal and Narita (2009) aptly point out: "the core claim of internalists who, like Chomsky, seek a naturalistic theory of language is that the proper object of a serious linguistic science should be organism-internal aspects of human language (namely I-language). Internalists never deny that there are phenomena broadly related to language (in particular to language use) that are beyond the narrow confines of the architecture of the human mental grammar (I-language). Such phenomena would surely include prescriptive pressures from the linguistic community, speakers' intentionality for communicative success, and all sorts of other E-language phenomena […] What internalists doubt is rather the feasibility and/or legitimacy of providing a serious science of any mind-external phenomena such as these."[81] (Lohndal and Narita 2009: 326)

This doubt is what I seek to address in this chapter. Davidson, through his *unified theory*, gives us what appears to be a serious scientific way of approaching the mind external phenomena, though we would here be in the realm of the social sciences.[82] He combines a theory of truth, traditionally understood as an externalist enterprise, with decision theory, the domain of which is, at least partially, the mind internal beliefs and desires. I have, so far, attempted to reconcile two dominant ideas found in Wittgenstein's works: first, what I have called the internalist strand in Wittgenstein[83], and the second, the externalist element: that when we talk or use language, we refer to objects, events, places external to language itself.

[80] For details refer to section 1.3
[81] For a discussion on the difference between I-language and E-language, refer to section 1.4
[82] Though if one considers it more scientific to reduce talk of beliefs, desires and other mental phenomena, to something physical, Davidson's anomalous monism insists that there are no law-like correlations between phenomena classified as mental and phenomena classified as physical, though mental entities are identical, taken one at a time, with physical entities. (Davidson 1995: 120-21)
[83] exemplified, for instance, by his claim that "[The] content of the sentence is in the sentence"(BBB 167)

Davidson's unified theory thus particularly appeals to this enterprise. Through this chapter we will see whether we can adopt a Davidsonian decision theoretic approach, towards reducing or reinterpreting the internalist/externalist divide.

5.2 A Substantive Account of Truth as a Theory of Meaning

Davidson's seminal paper 'Truth and Meaning' (1967) has been widely discussed and reviewed over the past 45 years. What is particularly relevant for our purposes is that Davidson rejected the possibility that each word has some entity as its meaning. He suggested, among other problems, that this would lead to infinite regress while explaining how meanings of sentences are generated from these word meanings. Importantly, for Davidson, the sentence is the unit of meaning. And knowing the meaning of the sentence does not simply consist of knowing the meanings of words combined with the appropriate syntax, "for knowledge of the structural characteristics that make for meaningfulness in a sentence, plus knowledge of the meanings of the ultimate parts, does not add up to knowledge of what the sentence means."[84] (Davidson 1967: 21) But he also rejected the idea that the meaning of a sentence is what it refers to, as this would entail that all sentences with the same truth value would be synonymous.

What he starts with seeking is "a theory that has as consequences all sentences of the form 's means m' where 's' is replaced by a structural description of a sentence and 'm' is replaced by a singular term that refers to the meaning of that sentence" (1967:20) He then replaces 's means m' with 's means that p', where the latter is without the problematic singular term which is supposed to refer to meaning, and p stands for a sentence. What he wants from the theory, or what, according to Davidson, would make the theory adequate, would be that we have for every sentence (s) a sentence (p) which somehow gives the meaning of s.[85] It is clear that p can easily be replaced with the sentence s itself. But the predicate 'means that' is a non-extensional predicate. So he replaces that with an extensional predicate to get the following:

(T): *s* is *T* if and only if *p*[86] (1967:23)

[84] This problem arose, for Davidson, in the case of belief sentences, where word meanings do not help towards giving us the truth conditions of these sentences.
[85] "The theory will have done its work if it provides, for every sentence s in the language under study, a matching sentence (to replace 'p') that, in some way yet to be made clear, 'gives the meaning' of s. One obvious candidate for matching sentence is just *s* itself, if the object language is contained in the metalanguage; otherwise a translation of *s* in the metalanguage." (Davidson 1967: 23)
[86] Where '*s'* is replaced by the description of a sentence s, 'T' is an extensional predicate, if and only if is the logical connective and 'p' is replaced by the sentence s.

For (T) to be significant for a theory of meaning, the following conditions should be met: (i) it must not appeal to any more semantic notions, (ii) it should "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."[87] (1967:23) The reason why (T) would work for a theory of meaning is that any predicates which satisfy the above conditions, have the same extension. Now, Davidson contends that from the above we discover that "the sentences to which the predicate 'is *T*' applies will be just the true sentences of *L*, for the conditions we have placed on satisfactory theories of meaning is in essence Tarski's Convention *T* that tests the adequacy of a formal semantic definition of truth." (1967:23)

So the strategy is, to put it briefly, that we start with trying to find the meaning of a sentence, but realize that 'means that' is not extensional[88]; so we shift to an extensional predicate and impose certain conditions on it, for it to best generate *meanings*; and then we realize that these conditions lead us to Tarski's Convention *T*[89], and the predicate that satisfies these conditions is precisely Tarski's truth predicate. So we move from seeking a predicate that would generate meanings, to arriving at the truth predicate.

It is important for our purposes that Davidson's condition (T) above (1967:23) deals with sentences as a whole (and descriptions of sentences as a whole), and does not tell us, for instance, how the meanings of sentences depends on the meanings of its parts. Given the observation (1967:21) above, Davidson contends that word meanings and compositionality are insufficient for meaning as they are insufficient for identifying truth conditions. We get truth-conditions, which are clearly what Davidson wants from a theory of meaning, only from whole sentences. So a theory of meaning is also a theory of truth for Davidson[90]. Now, a theory of truth does, "in a minimal but important respect, do what we want [from a theory of semantics], that is, give the meanings of all independently meaningful expressions on the basis of an analysis of their structure." (1970:55) It is therefore clear that the structure of the sentence is also important for Davidson, but this is only because we need a finite way of capturing potentially infinite true sentences. He thereby places the following demands on the theory: (i) it should "account for the meaning (or conditions of truth) of every sentence by analyzing it as

[87] Where '*L*' refers to a language.
[88] The assumption is, for Davidson, that we would prefer that our theory of meaning be extensional.
[89] Tarski's Convention T: X is true if, and only if, p (with 'p' replaced by any sentence of the language to which the word "true" refers, and 'X' replaced by a name of this sentence)
[90] Though he insists that truth conditions and meanings cannot be equated: "at best we can say that by giving the truth conditions of a sentence we give its meaning" (1970: 56, footnote 3).

composed, in *truth relevant* ways, of elements drawn from a *finite stock*" (emphasis added); (ii) it should "provide a method for deciding, given an arbitrary sentence, what its meaning is" (Davidson 1970: 56)[91] Tarski's Convention T again satisfies these conditions. What is needed from the theory is to give a recursive account of truth[92]. Equipped with such as account, we would be in a position to say for any sentence what is required for it to be true, "and this amounts, in one good sense we can give to the phrase, to understanding the language." (Davidson 1967:24) In 2.1 below, we'll consider what this *semantic conception of truth* is for Davidson.

5.2.1 Davidson on Deflationism and the Nature of Truth

There are various forms of deflationism about truth (Horwich 1998, 2008; Field 1986,1994), but what is largely common to them is the claim that truth is not a substantive notion. Horwich, as we noted in chapter 3, deflated both compositionality and truth. Davidson (1990a,1996) has argued against Horwich and against deflationism, to establish his substantive conception of truth. Here, we will point out what motivated Davidson against deflationism and towards a substantive account of truth, in order to better understand the nature and kind of truth that Davidson's theory is able to offer. This understanding would enable us to see whether Davidson's truth theory can capture the aboutness of language.

Deflationism about truth, and specifically the kind that considers Tarski's definition[93] to be disquotational, can be understood in two ways: either as claiming that Tarski didn't capture all aspects of truth, or that the concept of truth is not as deep as we thought. According to Davidson, the former is the case, because nothing in Tarski's definition shows what these definitions have in common (Davidson 1990a). Deflationists believe that convention-T provides an adequate answer to the question 'what [do] Tarski's definitions have in common', but Davidson insists that this is not the case, and he is therefore not a deflationist. This is because, he argues, in cases where the object language is contained in the meta-language the requirement within convention-T is "merely syntactical" i.e. "it tells us something about the predicates, but not much about the concept. In other cases (where the object language is not

[91] There is a third condition he mentions here : "the statements of truth conditions for individual sentences entailed by the theory should, in some way yet to be made precise, draw upon the same concepts as the sentences whose truth conditions they state", but we omit discussing it for our present purposes.
[92] i.e. "to show that the syntax of the language is formalizable in at least the sense that every true expression may be analyzed as formed from elements (the 'vocabulary), a finite supply of which suffice for the language by the application of rules, a finite number of which suffice for the language." (1970:57)
[93] Refer to note 10 above.

contained in the meta-language) its application depends on our prior understanding of the notion of translation, a concept far more obscure than that of truth. The central point is this: aside from our grasp of the concept of translation, convention-T gives us no idea how to tell in general when one of Tarski's truth predicates applies to a particular language. He does not define the concept of translation." (1990a:296) An appeal to the undefined concept of translation already leaves out an important aspect of truth, and makes us unable to articulate what these definitions have in common.

Another criticism against Tarski's definitions is that they turn empirical or substantive truths into truths of logic.[94] The solution could be to add that "Tarski's predicate holds for all and only the true sentences of L" (Davidson 1990a: 292), where 'true' is the substantive undefined concept. But Davidson notes that the addition is not required: "we can just as well regard Tarski's truth predicate 's is trueL' as having the properties of our real-life predicate 's is true in L', as long as those properties do not create inconsistencies." (Davidson 1990a: 292) But, since the definition is no longer stipulative, we don't know what 'true' means and whether it would lead to inconsistencies. Davidson insists, however, that this is closer to what Tarski himself intended, as he claimed that his definition "does not aim to specify the meaning of a familiar word used to denote a novel notion; on the contrary, it aims to catch hold of the actual meaning of an old notion" (Tarski 1944: 341)

Davidson, then, maintains that Tarski's definition does not entail that the T-sentences are logical truths. They would be logical truths only if the truth-definitions are purely stipulative, or tell us everything there is to know about truth. And, whether the definitions are stipulative depends on the intention of the definer and can't be determined by studying the system. He puts forth the following analogy in order to illuminate this point: "Suppose we offer as a definition of the predicate 'x is a solar planet' the following: x is a solar planet if and only if x is just one of the following: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, Pluto. This entails the P-sentence 'Neptune is a solar planet'. Is this last a logical truth? One may as well say so if our definition is purely stipulative, otherwise not. The question whether it is purely stipulative is not one that can be answered by studying the formal system; it concerns the intentions of the person making the definition. If we were simply presented with the defining sentence, we could hardly fail to notice that, if we interpret the words in more or

[94] For instance, according to Etchemendy: although T-sentences seem to convey substantive truths, "they carry no information about the semantic properties of [the] language, not even about the truth conditions of its sentences" (Davidson 1990a:290) because T-sentences are truths of logic.

less the usual way, it expresses a substantive truth." (Davidson 1990a: 293-294) The point of contention is whether the P-sentence 'Neptune is a solar planet' can be logically derived from the definition, and is therefore a logical rather than an empirical truth. Davidson is arguing that whether it is a logical or empirical truth depends upon the intentions of the speaker: whether the speaker intended the definition to be purely stipulative or expressing a substantive truth. Davidson claims that all we need to do is work "within Tarski's system while acknowledging that the truth predicate has further essential properties, as long as we make no use of the unspecified properties." (1990a: 294) In this way, Davidson claims to give us a *theory* of truth, where "A theory of truth, in contrast to a stipulative definition of truth, is an empirical theory about the truth conditions of every sentence in some corpus of sentences." We can therefore use Tarski-like truth conditions, for instance, "'Snow is white' is true iff snow is white" and still claim that such truth conditions do not say everything there is to say about truth. The word 'true' used here is used to denote the intuitive substantive concept.

So for Davidson, "it is truth rather than reference [or satisfaction] that is the basic primitive." (1990a:299) This is because of "the key role of convention-T in determining that truth, as characterized by the theory, has the same extension as the intuitive concept of truth" (1990a:299) It is only in the implementation of convention-T for particular languages that we need the satisfaction predicate. And in fact he argues, in line with Quine's inscrutability of reference, that we can substitute endless other relations for any given satisfaction relation "without altering the truth conditions of any sentence or the logical relations among sentences." (Davidson 1997a:78) Thus Davidson considers the concept of truth as neither trivial nor shrouded in mysterious metaphysics. It is rather an intuitive concept which is readily available, unlike other semantic concepts like reference and satisfaction.[95]

5.2.2 Objectivity and Holism

If meanings are given by Tarski-like truth conditions, then the question arises, how is it that we can know that the conditions are satisfied at all, since even asking this question "would appear to require a confrontation between what we believe and reality; and the idea of such a confrontation is absurd" (Davidson 1983:137). Davidson's point is that "nothing can count as a reason for holding a belief except another belief" (1983: 141) We might want to say that a

[95] "[W]hat is open to observation is the use of sentences in context, and truth is the semantic concept we understand best. Reference and related semantic notions like satisfaction are, by comparison, theoretical concepts (as are the notions of singular term, predicate, sentential connective, and the rest)." (Davidson 1990a: 300)

sensation justifies the belief in that sensation; but this justification cannot apply to external objects. Davidson (1995a) follows Descartes[96] in the belief that what one knows for certain is that thought exists, but unlike Descartes, he does not cast doubt upon everything that we think we know[97]. So the world is, more or less, the way we think it is; or in other words most of our beliefs about the world are in fact correct.[98]

A shared, objective reality is thus a starting point, and not a reasoned consequence. The very existence of thought attests to (though it does not explain) the existence of an objective reality, since one cannot have a belief (the most basic form of thought) without understanding that a belief may be false. "Awareness of the possibility of surprise, the entertainment of expectations – these are essential concomitants of belief" (1995a:7) He argues that surprise "requires that I be aware of a contrast between what I did believe and what I come to believe. Such awareness, however, is a belief about a belief: if I am surprised then among other things I come to believe that my original belief was false."(1980: 104) Thus surprise comes with the realization that our earlier belief was false. So when we are in a position to be surprised, we are in a position to evaluate our own beliefs and change them based on new information regarding objective reality. But surprise cannot explain objective truth, as we already need the concept of objective truth in order to be surprised. The point is that recognizing the chance that we may be wrong is to recognize that beliefs can be tested. And "the notion of a proposition being true or false independent of one's beliefs or interests" is the notion of objectivity. Thus, "to have the concept of truth is to have the concept of objectivity." (1995a:10) None of this derives the existence of an objective reality.

It is not just the concept of truth whose possession entails another concept (objectivity), rather, for Davidson, it is not possible for a creature to entertain a thought on its own: "the identity of a given thought depends in part on its relations to other thoughts." (1995a:11) To have any single belief, is to have a multitude of other true beliefs on which it depends. For instance, "my belief, if it is to be about a snake, whether it is a true belief or a false one, depends on a background of true beliefs about the nature of snakes, of animals, of physical objects of

[96] Descartes rather began with doubting everything, and deduced from this that what I cannot doubt is the fact that I doubt, i.e. that I think.

[97] "I should begin, then, as I think we clearly must, in medias res, assuming that we have a roughly correct view of our surroundings and of the existence of other people with minds of their own… Of course we are apt to be wrong about many things, but the possibility of error depends on a generous supply of truths" (Davidson 1995a:5)

[98] "It is absurd to look for a justifying ground for the totality of beliefs, something outside this totality which we can use to compare with our beliefs. The answer to our problem must then be to find a *reason* for supposing most of our beliefs are true that is not a form of *evidence*." (Davidson 1983: 146)

the world." Thus to have any belief, we need a host of *true* beliefs; though, "with respect to most of our beliefs, any particular one may be false." (1995a:16) What saves truth from being 'radically non-epistemic'[99] is that "belief, through its ties with meaning, is intrinsically veridical."(Davidson 1983: 156) The idea is that although any particular belief may be false, they cannot all collectively be false. This is because belief and meaning are interconnected.[100] So Davidsonian Holism entails that "to have a concept, to entertain propositions, to be able to form judgements, to have command of the concept of truth. If a creature has one of these attributes, it has them all." (1995a: 9)

In Davidson (1983) he calls his theory of truth a coherence theory in order to stress that the concepts of meaning, truth and belief are interrelated: "each of these concepts requires the others, but none is subordinate to, much less definable in terms of, the others."(1983: 156) In 1987 Davidson added some clarifications as afterthoughts to this paper. Calling his theory a coherence theory, he clarified, was a way of making the negative point "that 'all that counts as evidence or justification for a belief must come from the same totality of beliefs to which it belongs'." (1987:155) But this does not lead Davidson, unlike others who hold this view, to the conclusion that reality and truth are constructs of thought.[101]

Davidson however is not arguing for realism of the correspondence theory variety. He agrees with C. I. Lewis that a correspondence theory of truth faces the objection that there is nothing for true sentences to correspond to, "that if sentences correspond to anything at all, it must be the universe as a whole; thus all true sentences correspond to the same thing."(1988: 183) Neither does he want to advocate the kind of coherence theory that ties truth directly to what is believed and is therefore subjective. But to reject it, he insists, "is not to say there is no connection whatever between belief and truth." (1988: 185) Rather "there are inescapable and obvious ties among the concepts of truth, belief and meaning. If a sentence s of mine means that p, and I believe that p, then I believe that s is true. What gives my belief its content, and my sentence its meaning, is my knowledge of what is required for the belief or the sentence to

[99] Here being radically non-epistemic "implies that all our best-researched and -established thoughts and theories may be false". (Davidson 1983:140)

[100] See sections 5.2.3 and 5.3 for more details on this connection.

[101] Another reason he identifies, to not stress coherence is that coherence is nothing but consistency, and although it is "in favor of a set of beliefs that they are consistent, but there is no chance that a person's beliefs will not tend to be consistent, since beliefs are individuated in part by their logical properties; what is not largely consistent with many other beliefs, cannot be identified as a belief."(Davidson 1983: 155)

be true." (1988: 188-9) Thus meaning comes from knowledge of the truth conditions of the sentence or belief.

For Davidson "All attempts to characterize truth that go beyond giving empirical content to a structure of the sort Tarski taught us how to describe are empty, false or confused. We should not say that truth is correspondence, coherence, warranted assertability… To the extent that realism and antirealism depend on one or another of these views of truth we should refuse to endorse either." (1988: 190) It is by giving empirical content to the Tarskian structure that he makes sure that his semantic conception of truth is substantive and not deflationary. We still need to explain how, given the observation (1987:155) above, we are able to go beyond the realm of beliefs and reach an objective reality. We take this up in 5.2.3 below.

5.2.3 Triangulation

"Our sense of objectivity" is, according to Davidson, a consequence of a sort of triangulation, which among other things requires two creatures. "Each interacts with an object, but what gives each the concept of the way things are objectively is the baseline formed between the creatures by language." (1980:105) Triangulation is "a prelinguistic, precognitive situation which seems to me to constitute a necessary condition for thought and language, a condition that can exist independent of thought, and can therefore precede it." (1997b: 128) Triangulation is a threefold interaction, which "involves two or more creatures simultaneously in interaction with each other and the world they share." (1997b: 128) One line of the triangle goes from the speaker to the object, another from the interpreter to the object, and the third from the speaker to the interpreter. The intersection of the first two lines is what locates the object in a shared space. Once these correlations (the triangle) are set up, each creature is in a position to expect the external phenomena (a table for instance) when it perceives the associated response of the other (the utterance 'table'). Error comes in when there is the occasional failure of expectation, i.e. the reactions do not correlate. The point isn't that consensus defines the concept of truth, but that it creates the space for its application. If this is right then thought as well as language is necessarily social.

Although necessary to thought, triangulation is not a sufficient condition: being pre-linguistic, it can also exist in animals. What's needed in addition, and what is essential to thought, is language: "Unless the baseline of the triangle, the line between the two agents is strengthened to the point where it can implement the communication of propositional contents, there is no way the agents can make use of the triangular situation to form judgements about

the world." (1997b: 130) It is through using a language which we understand, that we can interpret others: "the sentences of a language we understand have a known structure which we can use to understand other speakers." (1997b: 132)

This brings us to a fundamental principle of Davidson's work on triangulation: the principle of charity. In an interpretive situation, Davidson insists, following Quine, on the centrality of the Principle of Charity, which entails that we interpret others according to our own standards of rationality. The principle of charity "directs the interpreter to translate or interpret so as to read some of his own standards of truth into the pattern of sentences held true by the speaker." (1983: 148) For successful interpretation within a triangular situation, the speaker and the interpreter must recognize the similarity of their responses, i.e. what they naturally group together as similar must be very much alike. In other words, "A condition for being a speaker is that there must be others enough like oneself." (1992: 120) Thus triangulation demands that there must be at least two creatures interacting with one another, and further that these creatures must be greatly like each other. A single creature by itself is not enough: "there would be no saying what a speaker was thinking or talking about, no basis for claiming he could locate objects in an objective space and time, without interaction with a second person." (1992: 121)

Thus a lot of what we can know about truth, and objective reality, depends on our interaction with others. And much of this interaction depends on ascribing to others, our own standards of rationality. So, for Davidson, our understanding of truth is inextricably linked to our understanding of rational behavior. Davidson argued for a *unified theory of thought, meaning and action*, and a step towards achieving such unification was to put together a decision theory and a truth theory. As we noted in section 2.1, what is important for Davidson is the primitive (or intuitive) notion of truth applied to an interpretive context. But how are we then to confirm the truth of the T-sentences in such interpretive contexts? This is where decision theory becomes central to Davidson's enterprise: "to the extent that an agent is rational he will take what he believes is the best bet available (he "maximizes expected utility")." (1990a: 316) We will, in the following section, try to understand Davidson's application of decision theory to sentences.

5.3 Decision Theory and its Place in Davidson's Work

Everything we do involves a choice. Choices pervade actions, inactions, and even our very existence. If the ability to make choices is so central to our being, it is important to try to find a way to explain our choices. Decision theory attempts to do just that. It rests on the idea that if your actions are reasoned and not random, and you are a rational agent, you would attempt to choose that which you think is the best possible outcome (i.e. you maximize expected utility). At first glance, this idea seems rather natural and intuitive, nothing controversial. For our purposes, it is in the details of how this is (potentially) applied to a truth theory, that the problems arise.

Decision theory "shows how to extract both cardinal utilities and subjective probabilities from simple preferences", and "subjective probabilities when taken as applied to sentences, are enough to yield a theory of meaning." (Davidson 1995) Cardinal Utilities measure the value a person gives to something. Subjective probability represents "the observer's degree (or strength) of belief that the system will adopt a certain state." (French 1988/1993: 222). The claim then is that an individual's preference of one sentence over another, gives us both what a person prefers (or values more), as well as what they believe is (or will be) the case. Davidson (1995) argues that "conditional probabilities and entailments between sentences, by registering what the speaker takes to be evidence for his beliefs, provides the interpreter with what is needed to interpret more theoretical terms and sentences." (1995:127) But he continues that we never have sufficient evidence, and we make all sorts of assumptions about the speaker's beliefs and desires etc. The claim is that we use our own norms and standards of rationality in order to interpret others. This is the reason why a mere manual of translation (in line with Quine 1960) is not sufficient for interpretation, as we need a way to interpret such a manual, and in this interpretation, we use our own standards of rationality.[102]

Before we go any further, I think it is important to note that the entire enterprise, i.e. the attempt to put together Decision theory and Truth theory is, for Davidson, a conceptual exercise: "The approach…is not…meant to throw any direct light on how in real life we come

[102] Thus in *Turing's Test* (1990b) Davidson argues that for the interpreter to know whether the object of the experiment is thinking (as opposed to whether it is a person or a machine), the object needs to be brought out in the open so that its interactions with the world can be observed. Merely by knowing that they are human we can legitimately infer that (like in our own case) there are causal connections between the person's words and the world; or alternatively we can discover these connections through observing how the speaker interacts with the world.

to understand each other, nor on how we master our first concepts and our first language." He admits that it is a "conceptual exercise aimed at revealing the dependencies among our basic propositional attitudes." (1990a: 324-25) So applying decision theory to sentences is not meant to explain how we *actually* understand one another or learn to use language, but rather it potentially explains how truth, meaning, belief and desires are related.

Towards this aim, Davidson begins with the conceptual acceptance of the principle of maximization of expected utility (as suggested in Davidson 1990a:316). Again, conceptually, this principle works on the idea that in making decisions, rational agents try to choose what they believe is the best available option (i.e. they try to maximize expected utility). This choice in turn depends upon the comparative values the agent puts on the expected outcomes (or how much he desires the outcomes, i.e. the cardinal utility) and the degree of belief that it will occur (or the expected probability). The problem in applying decision theory to truth theory is that we do not know how to tell what degree of belief a person has in some proposition or what the relative strengths of his preferences among propositions are.

Ramsey's solution to this problem involves finding a proposition which is as likely to be as true as it is false (for instance 'The outcome of this coin toss will be heads'). This proposition is then used to construct a series of wagers, and choices among these show how the various options are valued. With these values (the desirability of the various options), it is then easy to calculate the degrees of belief. Davidson admits (1995b) that although this is theoretically possible, in reality we may never have sufficient evidence[103], i.e. we make all sorts of assumptions about people and their beliefs and desires. The whole theory, he contends, is built on the norms of rationality, and as per the principle of charity, you use your own standards of rationality to interpret others. The Unified theory, he admits further, leaves no room for irrationality, neither does it say anything about "the forms of rationality that deal with incorporation of new information into a system of thought." (133) The theory is holistic in the sense that "sentences are interpreted in terms of their relation to other sentences, beliefs in terms of their relation to other beliefs and so on." (1995b: 130)

Within this relational interdependent whole, we get our sense of objectivity through triangulation that requires two creatures and an object. "Each interacts with an object but what gives each the concept of the way things are objectively is the base line formed between the

[103] "[T]here is no escape from the fact that we cannot check up on the… objective correctness of our own norms by checking with others, since to do this would be to make basic use of our own norms once more." (1995b: 134)

creatures by language." (1980: 105) Triangulation involves: "two or more creatures simultaneously in interaction with each other and the world they share." (1997b: 128) But it is not only the presence of this triangular interaction which is important; it is also important that the interacting creatures are aware of this interaction: they recognize similarities of response to stimuli. For instance in the case of interpreting a child: "The child finds tables similar; we find tables similar; and we find the child's responses to tables similar. Given these three patterns of response we can assign a location to the stimuli that elicit the child's responses. The relevant stimuli are the objects or events we naturally find similar (tables) which are correlated with the responses of a child we find similar. It is a form of triangulation: one line goes from the child in the direction of the table, one line goes from us in the direction of the table, and the third line goes between us and the child. Where the lines from child to table and us to table converge, 'the' stimulus is located.' (1992: 119)

An extreme setting useful to comprehend Davidson's notion of triangulation is a case of *radical interpretation* (1973). Imagine a situation where we are confronted with a need to interpret someone who does not understand our language. Davidson asks us to understand his radical interpretation in line with Quine's radical translation (1960) where a linguist attempts to translate a completely foreign language by simply observing the community of speakers in their natural setting. Translation is, however, importantly different from interpretation, since the latter involves communication and a semantic connotation missing the former.

In the case of radical interpretation, we will proceed in much the same way as we would with the case of interpreting a child above (1992). The principle of charity is crucial here: we grant that the foreigner naturally finds much the same things similar as we do (for instance the tables in the example above). Given that the foreigner finds tables similar and we find tables similar, we are able to find the foreigner's responses to tables similar and thereby achieve interpretation.

5.3.1 Critique of Davidson's Decision-Theoretic Account

Zilhao (2003) contends that "Davidson's characterization of human agency is not compatible with his other requirement that truth should provide the essential link connecting environment, speech and action." He argues that Davidson's account of human behavior as rational is based on an analysis of the behavior of a gambler. The fact that gambler-behavior can be characterized as geared towards the maximization of expected utility, does not warrant that all human behaviour can be so characterised; and therefore the principal of maximization of

expected utility cannot be taken to be the definition of rational agency. As a possible response to this criticism, note the following passage in *The Structure and Content of Truth* (Davidson 1990a): "The approach… I have outlined is not… meant to throw any direct light on how in real life we come to understand each other, nor on how we master our first concepts and our first language. I have been engaged in a conceptual exercise aimed at revealing the dependencies among our basic propositional attitudes." (1990a: 324-325) Thus what Davidson attempts to do is show that our basic propositional attitudes are connected, and one way to do this is through the framework he describes. He himself admits that his account is not meant to explain how people actually behave or understand each other in real life, or even how we acquire concepts or language.

Zilhao (2003) argues further that beliefs and desires are not independently available, i.e. there is no way of identifying them without accepting the principal of maximization of expected utility.[104] As a reply to this objection, one could argue on behalf of Davidson that precisely what he is attempting to do is give a theory about one possible way of identifying beliefs and desires via what is observable, that is, simple preferences. Beliefs and desires are not merely independently not available, but are rather not *available* for observation at all, therefore it is a useful tool to have a theory that identifies them, through what *is* available for observation, namely choices or simple preference (of the truth of one sentence over another).

Another point raised by Zilhao (2003) is that one cannot admit of a decision theoretic definition of rational agency while maintaining that truth has a central role to play in the connection of language and the world. He argues that "one cannot have it both ways: to claim, in the hope of procuring a foundation for empirical semantics, that an essential connection of veridicality should hold between speech, environment and action, on the one hand; and to severe that connection in order to make the interpretational theory conceptually simple and mathematically tractable, on the other hand." (Zilhao 2003: 244) Again, on Davidson's behalf, one could argue that it is this *essential connection of veridicality*, or the centrality of truth, as the basic primitive intuitive concept which is available to us prior to any analysis, that lies at the heart of the *interpretational theory*.

[104] "The procedures set up within the theoretical framework of Bayesian decision theory in order to allow us to arrive at a positive identification of the agent's beliefs and desires are pro-cedures that make sense only if one assumes from the outset that agents act according to the principle of maximization of expected utility." (Zilhao 2003: 234)

I do however sympathize with this objection, as well as the first objection regarding the possibility of generalization of gambling behavior, as it is surprisingly unclear from Davidson's writings on the subject how the decision theoretic framework is to be applied in the case of sentences. He does not, for instance, explain the application with any actual study to serve as an example. The application is further shrouded in mystery since, at times, Davidson takes the decision theoretic framework to be a kind of *analogy* to his truth-theoretic claims. For instance, in *The Folly of Defining Truth* (1996) he suggests that: "We should think of a theory of truth for a speaker, in the same way we think of a theory of rational decision: both describe structures we can find with an allowable degree of fitting and fudging, in the behavior of more or less rational creatures gifted with speech. It is in the fitting and fudging that we give content to the undefined concepts of subjective probability and subjective values – belief and desire, as we briefly call them; and, by way of theories like Tarski's, to the undefined concept of truth." (Davidson 1996: 278)

5.4 Reinterpreting Use

It is also clear from our discussions in this chapter that, for Davidson, knowledge of the truth conditions of a sentence is related to understanding the sentence. Further, this notion of understanding and truth conditions leaves open the idea that truth conditions of a sentence (or the knowledge of truth conditions), is related to the way the sentence is *used*[105]. So there is the potential, within a Davidsonian framework, for a use-theoretic account of meaning as related to truth and understanding. Towards this end, we will, in the remainder of this chapter first try to understand what *use* can potentially contribute to meaning.

5.4.1 Understanding Use

What is use? How do we begin to analyze it? Clearly, the use of a word is its utterance at a certain moment in time. One could potentially argue, for instance that the *use* of a sentence is the instantiation of the sentence (in a particular context), and the meaning of each instantiation is the sentence itself (in the context of utterance). In other words, a sentence used is an uttered (or read or written) sentence within its context of utterance (or discourse). It is clear that in an

[105] As exemplified by passages such as the following: "What is clear, I think, is that someone who knows under what conditions a sentence would be true understands that sentence, and if the sentence has a truth value (true, false or perhaps neither), then someone who does not know under what conditions it would be true doesn't understand it. This simple claim doesn't rule out an account of meaning which holds that sentences mean what they do because of how they are used; it may be that they are used as they are because of their truth conditions, and they have the truth conditions they do because of how they are used." (Davidson 1999: 111)

instance of use, the analysis must take into account the context (we will shortly consider in greater detail what the context consists in). We may, as a preliminary step, now define use as follows:

A use of x is an instantiation of x

How do we distinguish meaning from use? It naturally follows from the above that meaning can be separated from use just in case, and just as much as, meaning can be distinguished from an instantiation. Meaning is commonly understood as something that can be abstracted from each use or instantiation, or that which is common to the various uses. Thus, what distinguishes meaning from use, to put it loosely, is that meaning is what remains intact even when the context of the actual utterance is removed, along with any additional information about the speaker, hearer, the object of reference etc.

Apart from the context of utterance, the pragmatic context, we must also consider the syntactic context: what comes before and after the word, as well as the context of the discourse. Let us consider the following example:

'It is red'

This is a fairly simple statement, and means that the object referred to by 'it' possesses the quality of 'being red', i.e. the colour red. Now let us consider the discourse within which it occurs:

'Look at the light! It is red!'

This helps us make our meaning more precise by adding that 'it' refers to light. However, ambiguity still remains by virtue of pragmatic considerations. The way to clarify the meaning would be to introduce the pragmatic factors: the speaker utters this while the hearer is driving a vehicle at top speed while approaching a red traffic signal, implying that the hearer should stop or slow down; or alternatively, the speaker utters this while walking on a beach during sunset, or entering a room lit with red lights etc. In these cases, the utterance could be followed by 'stop!', 'how beautiful!' and 'how ugly!' respectively.

We can create any number of such examples. The point is that with each step, meaning gets more refined. Is each of the above at a different level of cognition, or, does each correspond with a different stage of the analysis or comprehension? In other words, is the processing of the pragmatic considerations successive or simultaneous, that is, does it occur at a later time?

What we gather from ‘the peanut was in love’ analysis (Nieuwland & Van Berkum, 2006) considered in 4.2.3, is that the meaning of the whole is calculated simultaneously, i.e. taking into account the context of utterance, and the context is not factored in at a slightly later stage: “contextual appropriateness not only can ‘‘neutralize’’ animacy violations, but can even render the animacy-violating predicate more preferred than an animacy-obeying canonical predicate.” (Nieuwland & Van Berkum, 2006: 1107) In the context of the story of the dancing peanut, we are led to expect the predicate ‘in love’ and consider the animacy-obeying canonical predicate ‘salted’ as a violation. This suggests that what matters as far as understanding the meaning is concerned is the entire relevant discourse within its utterance context.

One way to understand this could be as follows. We have at our disposal certain features (phonological, syntactic, semantic) which account for the meaning of x. So we have *i* features (f_i) in our system and these are universal, in the sense that they can account for the meaning of each word. The word x occurs within a discourse context C, which captures what comes before the word and after it: Ya_bZ, where a and b are phrases which form the remainder of the sentence of which the word x is a part (so for the word ‘ran’ in Mary ran the marathon ‘a’ would be the noun phrase ‘Mary’ and ‘b’ would be the determiner phrase ‘the marathon’). Y and Z are sentences forming the discourse. Each of the features have a certain weighting (W) depending on the context. So the meaning of x in context C is the following:

1. $M(x/C) = W_{01}^{C}(f_1) + W_{02}^{C}(f_2) + \cdots$
2. Where C = Ya_bZ

Now, there are statements which are analyzable (provable) only when taken in context, for instance the ‘The peanut was in love’ (there is no problem with the syntax, even to begin with. Here the problem was only semantic, and the context of the story which provides the animate feature to the peanut solves this problem.) But there are also statements which cannot be analyzed given our present system, for instance a particular instantiation of a new use. So this instantiation of a new use of x will not be analysable within our system (which includes the context of utterance, since this is a completely novel use), though it still maybe true. This possiblytrue but not provable (or rather, analysable) statement (the instantiation of a novel use of a term) is what necessitates the concept of use and, if meaning is use, the non-reducibility of meaning to lexical features. This is because the syntax derives the meaning of x from its lexicon, in which each lexical item is a collection of features. In the case of the novel use U* this collection does not capture the meaning, and therefore it cannot be accounted for within

the syntax. However, we assume (?) that the relevant (?) statement is true. Understanding these statements means that we understand each instantiation of a word, each use of a word in its own right (to use Hinzen(2007)'s terminology, as an 'atom'). So the use of x is an instantiation of x; each use gives the meaning (in abstraction); and the meaning of x is an abstraction of its own instantiations.

Thus it is important to incorporate use as distinguished from meaning and the analysis into features as the latter does not account for novel uses. The 'use' and not the understood 'meaning' becomes important only in case of a novel use. This is because in the case of a new use, although we know the context, we are not certain regarding the value or weighting to be given to some of the features. Thus we would be unable to perform the calculations in (1) in order to arrive at the meaning. But this novel use will eventually feed into the meaning (in the sense that we would recognize the weightage of the various features) and we will no longer need to refer to the novel meaning of the word in that particular instance. Thus use remains a 'use' only until there is abstraction. It appears that though every utterance is a use (or instantiation), the 'use' aspect lasts only an instant. The use is any and every particular instantiation; but any analysis of it would lead to meaning.

When we speak thus of an appropriate use, we mean rather an instantiation which is in line with meaning. Horwich's use theory is interesting in this regard, as his use-properties, in defining meaning in terms of use, retain the element which we just argued is lost in any analysis of use. Horwich's use properties, which constitute meaning are in fact particular instantiations. These particular instantiations are special in that we intuitively accept these statements as true, and this intuitive acceptance elevates them to the status of meaning-constituting use-properties. So there need not be abstraction from the use or instantiation to arrive at the meaning, but the use itself constitutes the meaning, prior to analysis of the sentence. So it appears, as we pointed out in chapter 3 as well, there is a sense in which we would already know the meaning of the relevant term, which is what enables us to intuitively accept the use-property sentence. But our discussion in this section also suggests that the use-property sentence is accepted without being analysed for meaning. That is, it is accepted as a whole, prior to analysis. There is no way of specifying the kind of meaning that these use-properties lead to except actually *using* the sentence. Thus Horwich's use-properties are unique in that they actually result in understanding meaning in terms of *use* in the sense of use suggested in this section. Below, in section 5.4.3, we will develop further how we use and interpret this notion of use-properties towards answering our original intuitions.

5.4.2 Pietroski's 'most' Experiments: A Davidsonian Application of Decision Theory

A possible application of a Davidsonian-like framework, through the analysis of choice behavior, are the experiments related to the meaning of 'most', carried out by Pietroski, Lidz, Hunter and Halberda (2009). In these experiments, the subjects were asked to answer whether or not they thought the statement (A) 'most of the dots are blue' was true, given various combinations of two, and later more than two colours of dots. In the first set of experiments involving dots of two colors (yellow and blue) the analysis of the result indicated that subjects did not understand the meaning of 'most' in terms of one-to-one correspondence of the two colours of dots: by, for instance, judging which of the two sets (blue dots and yellow dots) has at least one more. Rather they appear to have approximated the cardinalities of the two sets using ANS (approximate number system)[106]. In a two colour experiment, it is not possible to determine what speakers understand as the meaning of most, whether and how it is different from more for instance. So in the next set of experiments (Lidz et al 2011), more number of colours were introduced, to see whether for instance the subjects understood the meaning of (A) to be A1 or A2 as follows

(A1) $>(|\text{DOT} \cap \text{BLUE}|, |\text{DOT} - \text{BLUE}|)$[107]

(A2) $>(|\text{DOT} \cap \text{BLUE}|, |\text{DOT}| - |\text{DOT} \cap \text{BLUE}|)$

A1 contends that A requires that the number of blue dots ($|\text{DOT} \cap \text{BLUE}|$) is greater than the number of non-blue dots ($|\text{DOT} - \text{BLUE}|$). A2 says that A requires that the number of blue dots ($|\text{DOT} \cap \text{BLUE}|$) is greater than the result of subtracting blue dots ($|\text{DOT} \cap \text{BLUE}|$) from the total number of dots ($|\text{DOT}|$). Thus A1 requires one to determine the cardinality of both the blue dots and the not blue dots, while A2 selects only the blue dots. Given the observation from Halberda et al (2006) "that at rapid presentation rates, the visual system can accurately track a maximum of 3 sets" (Lidz et al 2009: 24) it was predicted that in A1 accuracy should decline as number of colours increase, and in A2 the number of colours should have no impact. The results were consistent with A2 and reflect the use of ANS: The performance

[106] The authors note, interestingly, that ANS is available to other animals as well, and not limited to humans.

[107] Where '||' signify cardinality and '∩' is for intersection of two sets. So, '> (|DOT ∩ BLUE|, |DOT − BLUE|)' means that the cardinality of the intersection of the set of dots and the set of blue things, is greater than the cardinality of the subtraction of the set of blue things from the set of dots.

declines as a function of ratio, with better performance on easier ratios (weber ratio). The performance is unaffected by the number of colours.

What these experiments do is provide a way to distinguish between and choose from truth conditionally equivalent semantic representation (like A1 and A2) through constraints imposed by interfacing cognitive systems like the visual and number systems. For our purpose, we note that these experiments provide a way to gage the meaning of a sentence (albeit one importantly containing a quantification term) from the preference behavior of participants, given two possible use-contexts (the visual representations). So here the observable phenomenon - the choice - is between contexts and not sentences as in the case of Davidson. But what Pietroski is aiming at is also the meaning of words (in this case 'most') and not the meaning of sentences. The different use contexts offer different truth conditions for the relevant term, given that the sentence is kept constant.[108] Thus the preference of one use-context over another, gives us the truth condition of the relevant term.

The reason why the above mentioned experiments become interesting for us, is that we can potentially imagine similar conditions for testing Horwich-like use-properties. Use-properties, we may recall, were based on our underived acceptance of a certain sentence containing the word in question. For example: "'red's meaning stems from our underived propensity to accept 'That is red' in response to the sort of visual experience normally provoked by observing a clearly red surface." (Horwich 2004:351). So given a clearly red surface we see, we intuitively accept the sentence 'that is red' in response. Thus we could potentially create an experiment in order to test these intuitions. For instance, if we consider that the meaning of red is a particular shade of colour, then we can create an experiment around that on the lines of the meaning of most experiments. We could show subjects a spectrum of colour cards with the hues of orange yellow red and brown, and ask them to pick which colour card they think best fits the sentence 'That is red'. Such experiments inform our intuitions regarding use-properties, by showing that these use-properties are testable empirical assertions.

[108] If the choice was over sentences, keeping the context of use constant, then the truth conditions would pertain to the meaning of one sentence over the other.

5.4.3 Re-interpreting Use-properties

Use, as we saw denotes an instantiation, an utterance of a sentence (or word) in a context. For Horwich words are assigned use-properties[109] as their meaning, and sentence meanings are trivially obtained through compositionality; i.e. the individual word meanings (use properties) are inserted into a compositional schema to obtain the meaning of the sentence. I argued in chapter 3 that a substantive account of compositionality is needed in order to derive meanings of words from use properties. Either we need such an account, or use properties lose their special status, as all we need to say is that in using terms, we already need to have an implicit knowledge of their meaning. In fact, even in case we admit a substantive account of compositionality, such an account could be given in each particular use of the term, again rendering use-properties superfluous.

For our purposes here, we want to retain Horwich-like use properties, while discarding the rest of Horwich's account, along with his deflated accounts of compositionality and truth. The idea we preserve is that use properties connect what we intuitively accept as true with the meaning of the word. This general idea is connected to or even exemplified in the experiments conducted by Pietroski et al (2009) and Lidz et al (2011). This connection has interesting consequences for a Davidsonian unified theory as we shall see now.

The experiments discussed in the previous section suggest a possible way in which the meaning of 'most' arises from the truth conditions of the sentence, and it provides a way in which this can be tested. This, we suggested briefly in this chapter, is a possible application of Davidson's decision theoretic procedure. In the experiments, the truth conditions and thereby the meaning of 'most' is determined by analyzing people's preferences. Since the aim is to arrive at the meaning of a word, the context of discourse[110] is kept constant while the use-context (or the external situation)[111] is varied. Davidson , on the other hand, aimed at reaching the meaning of a sentence, by figuring out which sentence fit the external truth condition better, based on the simple preference of one sentence over another, given a context of use.

---

[109] For example: "'bachelor''s meaning is engendered by our underived acceptance of the sentence, 'The bachelors are the unmarried men'" and "'red's meaning stems from our underived propensity to accept 'That is red' in response to the sort of visual experience normally provoked by observing a clearly red surface." (Horwich 2004:351). See chapter 3, section 3.2 for details.

[110] That is, in this case, the sentence ('Most of the dots are blue' for instance).

[111] The cases of blue and other coloured dots.

There are a number of important points regarding the kind of use-properties we get from the above, which we want to carry over into a reinterpretation of a use theory. Firstly, as it appears in the above formulation, meaning is not atomic, i.e. our knowledge of meanings of words is not isolated from our understanding of sentences. That we understand and accept a certain sentence gives us the meaning (of a word). Similarly for Davidson, meaning arises at the level of sentences. Words have meanings in the context of the sentence in which they occur. Secondly, meaning is essentially connected with propositional attitudes, which, if we are to agree with Davidson, are mainly beliefs and desires. Our intuitive acceptance (or preference) of one sentence over another gives us both the degree of belief as well as desire.[112]

In chapter 3 we noted that one of the problems with Horwich's conception of use-properties was that there appeared to be no systematic way of picking which sentence constitutes the use-property. With the kind of experiments mentioned in Pietroski et al (2009) and Lidz et al (2011) we now potentially have a systematic way to make this decision. We could for instance choose two sentences which we ourselves may be inclined to intuitively accept as the meaning of the word (in line with Davidson's use of the principle of charity) and then test to see whether others would prefer one over the other, keeping the context of use (or the external situation) constant. Such experiments would not only give us the meaning of the word, as in the case of Pietroski's experiments, but also the meaning of the sentence, or the conditions under which the sentence would be true, along with the subjective probability and expected utility of the sentence (applying a Davidsonian decision theory). Thus we now have a way of connecting the Davidsonian framework with a use-theoretic understanding of meaning. As we also saw in 5.4.1 above, using Horwich-like use-properties would also ensure that we understand meaning simply in terms of use itself, and the instantiation or use, is not lost to abstraction. Davidson himself was open to the idea that truth conditions of a sentence is related to the way the sentence is used, as is evident by passages such as the following: "What is clear, I think, is that someone who knows under what conditions a sentence would be true understands that sentence, and if the sentence has a truth value (true, false or perhaps neither), then someone who does not know under what conditions it would be true doesn't understand it. This simple claim doesn't rule out an account of meaning which holds that sentences mean what they do because of how they are used; it may be that they are used as they are because of their truth conditions, and they have the truth conditions they do because of how they are used." (Davidson 1999: 111) So for Davidson, knowledge of truth conditions of a sentence constitutes

[112] Refer to chap section 5.3 above.

understanding the sentence, and what I have tried to do in 5.4.1, 5.4.2 and 5.4.3 is show how this knowledge of truth conditions could be related to use, through using Horwich's use-properties and testing them through Pietroski-like experiments under the framework of Davidsonian decision theory.

How does any of this help us achieve what we set out to do: to reconcile internal content and external significance? What we achieve from this experimental approach is a relation between meaning and our propositional attitudes. If truth is in this way related to what we believe and what we desire, it appears that locating ourselves as situated in this world which we share with others is central to any external significance that we attach to them. To put it differently, much background precedes the establishment of any external significance (to any symbol, linguistic or otherwise). We need to assume to begin with that while providing this external significance, we are ourselves located in a world that we share with other beings much like ourselves, i.e. our rationality is for all purposes the same as that of other normal human beings. Thus, given our adaptation of Davidsonian triangulation and decision theory, what is content internal to our mind/brain, i.e. our propositional attitudes, is what determines and is inextricably linked to external significance or truth. What is crucial here, as in Davidson, is that we understand ourselves as located in a shared world, which we share with others like ourselves.

5.5 Wittgenstein and Davidson on the Impossibility of Private Language

The necessity of a world shared with others was felt even in Wittgenstein's PI, where he argued there can be no private language or one that only the speaker can know and understand. The reason for this is basically that without a way to verify meanings of words, without objective criteria for identifying them, even the speaker would not be able to identify them. Wittgenstein's private language argument, as set forth in §244–271 of the *Investigations*, is well known and extensively discussed since it appeared. Given that I have nothing particular to add to the argument, I have not rehashed it in this work. But it is important for us, that language (and not just language use), for Wittgenstein was not something that can occur in isolation, and requires a community of speakers and hearers, and rules and conventions of use. I would just like to briefly sketch the argument so as to bring together the conclusion of this work.

A private language, for Wittgenstein, is one where "The individual words… refer to what can only be known to the person speaking; to his immediate private sensations. So another person cannot understand the language.' (PI §243) For example: We can imagine a person A growing up alone on an island. There would be here no human being to communicate with. However if he got hurt, he may groan in pain, and he may laugh while watching a dog go round and round while attempting to catch his tail. He may even tame a dog or a horse, and may utter certain particular sounds when he wants them to go away or sit down etc. To carry our example further, if someone from the mainland comes to this island and befriends A, who, for the sake of our example, tries to communicate with him through his own usages of sounds, demonstrating what he means by telling the dog to sit down first, or uttering the sound and then sitting down, and doing this repeatedly till the person sits down. However, when the person sits, he has, for himself, translated the sound to mean 'sit down'. There is successful communication, and the language is thus no longer private. Even when A was previously telling the dog to sit down etc, there was communication, what he had to say was not *incommunicable.* Wittgenstein argues, through the example of private sensations such as 'pain' that it would be a mistake to think that only the person feeling the pain can understand what the word means when he utters it. To put it briefly, he argues that to have a private language for our immediate sensations, i.e. a language which is unintelligible to anyone but us, is impossible. This is because such a language would be unintelligible to us as well, without any objective criterion for identifying which of its words refers to which sensation. If on the other hand, such objective criteria exist, then the language will no longer be private as others would also be able to understand it.

As another example: A person writes a diary and comes up with a new word to describe how she felt on the first day of the job, as she thought no known words could express it. Now she may use the same word again, every time she felt that way. One day her friend reads her diary and asks her what the word means. Now she explains the situation to her friend when she felt that feeling, and said 'its something between fear and nervousness, excitement and elation, importance and insignificance...' here we cannot be certain that B's friend completely got her import, but we can be sure that she understood what the ordinary language words mean, and through imagination she could succeed in combining them. Further, in order to make sure she remembers the exact meaning of the word, B writes down the situation in which she felt it, and also the explanation she gave her friend. Now there is an objective criterion to check whether

her usage in later situations is correct. However, once it is written, it is no longer private; the rules of its usage can be read and understood by others.

In fact, I agree with Kripke (1982) although not with his skeptical conclusions, that the crux of the private language argument appears earlier in the Investigations than most believe, i.e. in PI §202: " 'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately': otherwise thinking one is obeying a rule would be the same thing as obeying it." I do not think that this leads to the skeptical conclusion that there is no such thing as 'obeying a rule'.[113] A practice involves repetition of performance. So a rule needs to be obeyed more than once for an action to qualify as obeying a rule. In the case of a new rule, it must be accepted as a rule by more than one person, i.e. it needs to be initiated into a 'way of life' of a community. The idea is that if one obeys a rule known only to oneself, then there would be no way to judge whether we are correctly following it. For objectivity, we need to consult other people.

Davidson argued further that our sense of objectivity itself is derived from this triangular relation formed by our simultaneous interaction with other people and the world. In other words, with Davidson's Triangulation (see section 5.2.2 and 5.2.3), we see that objectivity, or at least *our sense* of it, involves at least two people (a speaker and hearer) simultaneously in interaction with each other and the world. We have a similar necessity within Wittgenstein, via the rejection of a private language. Both stress on communication within a community of speakers, in order for there to be objectivity. However, I do not want to harp on this similarity here, and point it out only to buttress the implication for my argument that sharing our world with others is deeply connected to understanding language.

Before we conclude, I would like to point out an alternative way in which Internalism and Externalism has been reconciled in recent times. This is the position of Un-Cartesian Linguistics. I only just briefly mention it here so as to point to an alternative to the Cartesian view I have been assuming.

5.6 Cartesian and Un-Cartesian Linguistics

Commonly, a variety of domains are thought to be forms of expression- art, music, dance, theatre; and language are among them. In what way do these qualify as forms of expression?

[113] I have in no measure exhausted all there is to say about *Rule Following*. But my purpose here was simply to give a general gist of the argument in order to indicate how it relates to the direction of this thesis.

To speak of forms of expression, entails that there is something, importantly distinct from these various forms, and we use these forms to express it; i.e. the modes of expression must be distinct from the object of expression. One often hears of the various emotions expressed by these forms, or even questions and commentaries on social or political structures. Thus it seems that we can express through these varied media, our emotions, beliefs, fears, hopes etc., i.e. our thoughts. Being mind-internal, these various structures need a medium of expression in order to be shared, given that we are social beings and live in this shared world. Thus intuitively, it appears that language is, like art and music, a mode of expression of our thoughts (which are thereby distinct from language).

What does it entail to call language a form of expression? It would mean, at least, that we use language (or other modes of expression above) to express something other than itself. One may get the feeling, time and again, of being unable to express oneself. I often feel that there are thoughts in my mind, and I'm unable to find the right words for their expression. This intuitive distance between language and thought has dominated much dialogue in philosophy of language. Chomsky (1966), for instance, contended that human language is "an instrument for the free expression of thought". (Chomsky 1966: 65)

Though Chomsky argued that Cartesian mind-body dualism was no longer a problem, as the distinction between mind and body can no longer be formulated, he was however Cartesian in at least two important senses. Firsty for Descartes, as for Chomsky (1966) grammar is specifically human. No animals have propositional thought such as in humans, though animals may be said to have thought to some degree. In our immediate environment we may have noticed the pattern in the behavior of our pets. People who have pets can sometimes even distinguish between the barking noises that indicate (or even communicate) to them what exactly they want. A large range of research indicates that non-human animals communicate with each other using signals of a variety of sorts (Hauser et al. 2002). But as noted in 1.2 animal communication is not like human language as it lacks generative capacity and is not stimulus independent.[114] Thus language (grammar) is unique to humans.

Consider the research on primates, e.g. Savage-Rumbaugh's work on bonobos. The acquired vocabulary of these animals can apparently consist of about 300 to 500 distinct symbols, and they can even use these in appropriate contexts. (Hauser et al 2007:6-7) This

[114] Chomsky's (1959) criticism of Skinner stresses that human language, unlike animal communication, cannot be explained in terms of stimulus-response.

would be greatly significant if these animals are able to combine this limited vocabulary to form an unbounded array of expressions. This however does not seem to be the case, "their lexicon never takes off as it does in very young children... rarely do animals combine two or more symbols, and when they do, the system appears highly constrained."(Hauser et al 2007: 7)

Further, Boysen and Matsuzawa's work on Chimpanzees indicates that though their ability to 'map the number of objects onto a single Arabic numeral' is impressive, however, 'they require thousands of training trials and often years to acquire the integer list up to nine'. (Hauser et al 2002:1577) Human children, having acquired the numbers 1,2 and 3 go on to acquire all the numbers, as they have grasped the idea of the successive function. This however, is not the case with chimpanzees. Their acquiring of the integers is similar to the acquiring of any other symbol with its referential properties: "The system apparently never takes on the open-ended generative property of human language. (Hauser et al 2002:1577)

The point here is that there is evidence to suggest that Descartes and Chomsky were right in that human language is unique to the species. We have also intuitively followed the further Cartesian argument that language and thought are distinct, and language expresses thought. This is, as noted above, also in line with Chomsky. Chomskian Biolinguistics also understands language on analogy with a biological organ, or an independent faculty which interfaces with other systems like the conceptual-intentional system and the sensory-motor system, which lie outside the core faculty of language of FLN (see figure 1).

An alternative to this Cartisian position is the Un-Cartisian approach (Hinzen 2006, 2007, 2012, Hinzen and Sheehan 2013) which contends rather that language constitutes thought, and is therefore not distinguishable from it. Hinzen emphasizes that human mode of thought is inherently grammatical. According to Hinzen 2012 given that a human mode of thought is intrinsically grammatical, there is the possibility that "the theory of grammar may upgrade to an actual theory of thought" (2012: 643) Thought, in Wittgenstein, as well as traditional semantics, was meant to have external significance: it is propositional, where a proposition is one which can be judged to either agree or disagree with reality. By making thought inherently grammatical, Hinzen effectively abolishes the divide between syntax and semantics: he promotes the view that semantics, which traditionally analyses propositional thought in order to derive linguistic meaning, needs to realize the significance of syntax, and the domains need to merge. Hinzen and Sheehan 2013 argue that syntax gives us our mode of

reference to the world, i.e. the way we understand and classify the world is grammatical. Thus grammar gives us an ontology and is not merely the vehical for an expression of thoughts. For example, consider the following:

1. Mary smiles.
2. Mary's smile is beautiful.

These sentences may be uttered given the same external situation. They also contain the same root word smile. But the (1) picks out a state of affairs while (2) picks out an object or property. Our mode of reference (to the same reality) is linguistically distinct, and this, the Un-Cartesians claim makes an ontological difference. The idea is that grammar enables us to refer to the world in very specific ways; our mode of reference is grammatical, and grammar makes available specific ontological categories: "[G]rammar *narrowly constrains* the ways in which words can be used to refer, making available a small number of discrete options in which this can happen, ranging from purely predicative nominals to quantificational, to referential, to deictic, and finally, to personal ones" (Hinzen and Sheehan, 2013: 211)

I have here only briefly mentioned this vast and fruitful enterprise so as to indicate that there are alternatives to what I have generally accepted as an axiom: that language and thought are distinct. That *grammar narrowly constrains the ways in which words can be used to refer* is I believe quite acceptable. Grammar constrains the way language is used, and thereby also the way we use words to refer and insofar as thought is linguistic, how we think about and grasp the world around us.

The task I took up in this thesis was to find a way to reconcile internal content and external significance. The Un-Cartesian position could be one way to go towards this aim. By understanding language as constitutive of thought, the Un-Cartesians abolish the syntax/semantics divide (as noted above). The claim is that syntax, traditionally understood as constituting internal significance, and semantics, traditionally understood as the domain of external significance, merge in the Un-Cartesian position, and they are therefore able to reconcile internal content and external significance.

My problem with this is that, I do not believe that abolishing this distinction leads to much explanatory benefit. To say that grammar constrains truth and reference, does not amount to an explanation of how we achieve truth and reference, but is rather a description of the actual

phenomena.[115] The Un-Cartesian may reply that grammar gives you truth and reference and it constitutes external significance when it joins forces with the lexicon. I sympathise with the idea that our mode of reference is grammatical, but I do not think it has the philosophical significance that Hinzen (2012) assumes.

5.7 Conclusion

We saw that for Davidson, it is a sentence that is the unit of meaning (unlike for instance, Horwich 1998) as truth conditions pertain to whole sentences. The reason why parts of a sentence become relevant is essentially that one wants a semantic theory to generate potentially infinite meanings from finite resources. The predicate that is able to do this job, turns out, for Davidson, to be Tarski's truth predicate. The conception of truth, for Davidson is our intuitive concept of truth, which is substantive and not deflationary. Further, it is only in the presence of another creature much like ourselves, that we are able to locate objective reality. But this does not mean that reality is a mental construct. It is rather that we begin with a shared world, filled with other people who possess the same standards of rationality as ourselves. All this is implied in the principle of charity, which is essential to achieve interpretation. Truth, meaning, desire and beliefs are all interconnected for Davidson, strung together by the principle of charity. We can use Decision theory, or an analysis of simple preferences, to extract subjective probabilities as well as cardinal utilities and the former, when applied to sentences gives us meanings. We mentioned the experiments regarding the meaning of 'most' by way of a possible application of Davidsonian decision theoretic framework.

It is also clear from our discussions in this chapter that, for Davidson, knowledge of the truth conditions of a sentence is related to understanding the sentence. Further, this notion of understanding and truth conditions leaves open the idea that truth conditions of a sentence (or the knowledge of truth conditions), is related to the way the sentence is used. So there is the potential, within a Davidsonian framework, for a use-theoretic account of meaning as related to truth and understanding. The reason why Pietroski's 'most'-experiments are particularly interesting in this regard is that they provide a bridge between Horwich-like use-properties and a Davidsonian unified understanding. I believe this method is more explanatory than the Un-Cartesian position discussed.

[115] Un-Cartesians believe that such a description of the phenomena of reference and truth, is, as much of an explanation that is possible.

Chapter 6: Concluding Remarks

Traditionally, much of philosophical as well as scientific enquiry arises from concerns regarding the nature of things or the relations between things. It is then perhaps natural to wonder how philosophy can contribute to a naturalized discussion of things and relations. Let us for instance consider the field of language acquisition. How we acquire language has been a topic of interest to Philosophers since the Vedic period in ancient India. We can however, quite easily get data from linguists working in language acquisition, which tells us, for instance, that the way children learn language is uniform across the species. Babies a few days old can discriminate between their mother tongue and a foreign language from a different class; around two weeks, they distinguish between voices and unvoiced consonants, and at about 4 or 5 months they can distinguish their own language from a foreign language in the same rhythmic class. From 18-24 months, children have telegraphic speech (two word utterances) and fully grammatical utterance appear uniformly by 4 years of age.[116] That there is an innate mechanism that develops within poverty of stimulus conditions is evident, for example, from the case of three deaf children who constructed a sign language secretly, against their parent's wishes, and this language was found to be similar to spoken language of children of the same age. (Goldin-Meadow and Fieldman 1977, Glietman and Newport 1995)

The point is that there is a lot a linguist can tell us about language acquisition and the nature of the language organ or faculty. We can also get information regarding the relation of vision or perception with the acquisition of, as well as the use of language, through for instance brain imaging techniques. What then could be the significance of a philosophy of language with all this knowledge at our disposal? It would, I think, be useful to ask whether any questions remain unanswered through a natural scientific enquiry into language acquisition or even linguistics. The question remains, what does it mean to know that a particular part of the brain lights up when we see a particular image, for instance. Or what can we know about reference if we know what kind of a linguistic structure is required to achieve reference?

I believe a philosophical addition to the naturalistic discussion would be in terms of noticing connections. Traditionally, philosophy has questioned the fundamentals of any field,

[116] See for example Stromswold (2000) for a survey of such research into language acquisition.

what do the axioms of any field mean in terms of a deeper understanding of the elements and relations involved. Philosophy is I believe particularly instructive in helping to understand the depth or importance of previously known connections. And this is, I hope, what I have done through this work. Not many would doubt (except for extreme Cartesians) that we live in a world that we share with others. But that *this* particular fact is deeply connected to our use and understanding of language is a separate and less obvious claim, and one I hope I have at least pointed towards.

To briefly trace the progression of the thesis: we started out seeking to understand the intentionality of language; how is it that we are able to reach out to the world through language. That we have an internal Faculty of Language which facilitates this reaching out, was not (as we also pointed out above) the answer we were looking for. Biolinguistics specifies the internal structure that language uses, but the question remains, how does language reach out to the world. The answer we thought was both obvious and important: language reaches out to the world through its use. We tried to develop this idea through an (re)interpretation of Wittgenstein, where we understood the determinate relation of the Tractatus as grammar of the later Wittgesnstein. Here grammar in the broad sense already involved an element of use, and thus use was part of what constitutes a proposition (i.e. the determinate relation). Further through an analysis of Horwich, we attempted to see whether it is possible to systematize a use theory of meaning. We found that Horwich's account could use strengthening that would either take away from the use-theoretic aspect of his account, or his account would be rendered inconsistent. We then moved on to externalist referential semantics, specifically to minimal semantics, and focused on Borg, as she appears at first glance to have a theory which incorporates both internalist as well as externalist elements. We found however, that there is nothing within the theory that necessitates Borg's externalism. She does not attempt to explain reference but rather just stipulates that words have external significance; that the word 'dog' gets its content from dogs. For a theory that was able to scientifically explain content ascription, we turned to Davidson. Davidson's decision theoretic account of what he called triangulation presented itself as a candidate to better explain how we ascribe a particular content to a particular word or sentence. This framework, we argued, was exemplified in Pietroski's "Meaning of 'Most'" experiments. We contended that Horwich's use-properties could form the test statements for Pietroski-like experiments, which would yield a possible way to relate meaning to what we consider true. And this result is only achieved with the important assumption that we live in a world which we share with others much like ourselves.

In the introduction we noted two related problems: first was the problem of intentionality: how is it that we are able to reach out to the world through language. The obvious answer, we thought, was through use. The second problem was scientific enquiry: whether or not it is possible to scientifically study the object of externalist inquiry, the use of language for instance. The kind of theory I try to point towards, both incorporates how we are able to reach out to the world through language; i.e. explains use, and also aims to do so scientifically: through experiments which indicate how use determines meaning. Beyond this argument, I hope that this work was also able to at least point towards the fact that: what enables us to use language to reach the world is that we ourselves are situated in this world which we share with others much like ourselves. To put it differently, it is language use *within a community of users*, which serves as a background for understanding language.

References

Antony, L.M., Hornstein, N. (2003) *Chomsky and his Critics*. Blackwell Publishing.

Baker, M. C., (2001) The Atoms of Language, Basic Books

Baron-Cohen S., S. Wheelwright, V. Stone and M. Rutherford (1999) "A Mathematician, a Physician, and a Computor Scientist with Asperger Syndrome: Performance on Folk Psychology and Folk Physics Tests" *Neurocase* 5:475-83

Baron-Cohen S., S Wheelwright, V. Scahill, A. Spong, and J. Lawson (2001) "Are Intuitive Physics and Intuitive Psychology Independent? A Test with Children with Asperger Syndrome" *Journal of Developmental and Learning Disorders,* 5:47-78

Bellugi U., L. Lichtenberg, W. Jones, Z. Lai, and M. StGeorge (2000) "The Neurocognitive profile of Williams Syndrome: A Complex Pattern of Strengths and Weaknesses" *Journal of Cognitive Neuoroscience* 12 (suppl): 7-29

Boghossian, P.A. (1993) 'Cognitive Science and the Analytic/Synthetic Distinction: Comments on Horwich'. *Philosophical Issues*, Vol. 3, Science and Knowledge, pp. 135-142

Borg (2001) "The Metaphysics and Epistemology of Singular Terms" *Philosophical Papers* 30 (1): 1-30

Borg, E. (2004). *Minimal Semantics*. Oxford: Oxford University Press.

Borg, E. (2009). 'Must a semantic minimalist be a semantic internalist?' In Proceedings of the Aristotelian Society Supplementary Volume LXXXIII: 31-51.

Borg E (2007) "Minimalism versus Contextualism in semantics" In G. Preyer and G. Peter (eds) *Context Sensitivity and Semantic Minimalism : Essays on Semantics and Pragmatics.* 546-571.

Borg, E. (2010). 'Minimalism and the content of the lexicon'. In *Meaning and Context*, ed. L. Baptista and E. Rach. Bern: Peter Lang. 51-78.

Borg, E. (2012*) Pursuing Meaning.* OUP

Bowler D. M. (1992) "The Theory of Mind in Asperger Syndrome" *Journal of Child Psychology and Psychiatry.* 33: 877-93.

Brandom, R.B. (1994) *Making It Explicit. Harvard University Press.*

Brandom, R.B. (2000) *Articulating Reasons: An Introduction to Inferentialism***.** Harvard University Press,

Brandom, R.B. (2007) “Inferentialism and Some of Its Challenges” in *Philosophy and Phenomenological Research*, Vol. LXXIV No. 3, May 2007.

Cappelen, H., and E. Lepore (2005a). *Insensitive Semantics*. Blackwell.

Cappelen, H., and E. Lepore (2005b). “A Tall Tale in Defense of Semantic Minimalism and Speech Act Pluralism,” in G. Preyer and G. Peter (eds.), *Contextualism in Philosophy*. Oxford University Press, pp. 197-219.

Cappelen, H (2005). ‘Pluralistic Scepticism: Advertisement for Speech Act Pluralism’ in *Philosophical Perspectives*, 19, Epistemology.

Carruthers P. (2006) “The case for Massively Modular Models of Mind” in R. J. Stainton (Ed) *Contemporary Debates in Cognitive Science.* 3-21

Carstairs-McCarthy (1999) *The Origin of Complex Languages: An Inquiry into the Evolutionary Beginings of Sentences, Syllables, and Truth.* OUP.

Chomsky N. (1959) "A Review of B. F. Skinner's Verbal Behavior" in Language, 35, No. 1, pp26-58

Chomsky N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N., 1980, *Rules and Representations*, Columbia University Press, New York.

Chomsky, N., 1986, *Knowledge of Language: Its Nature, Origin, and Use*, Praeger

Publishers, New York.

Chomsky, N. (1997/2004) “Language and Mind: Current thoughts and ancient problems.” Published in. L. Jenkins (ed) *Variation and Universals in Biolinguistics*, (2004), 379–405.

Chomsky N. (1997). “Language and Problems of Knowledge”, in *Teorema* Vol. XVI/2, pp. 5-33.

Chomsky, N., (2000) Chapter: “Linguistics and Brain Science”. In A. Marantz, Y. Miyashita and W. O’Neil (eds.) *Image, Language and Brain*.

Chomsky, N., (2005) “Three Factors in Language Design” inLinguistic Inquiry, Volume 36, Number 1:1–22.

Collins, J. (2001)‘Truth Conditions without Interpretation.’ *Sorites*, 13: 52-72.

Collins, J. (2003)‘Horwich’s Schemata Meet *Syntactic Structures*.’ *Mind*, 112 (447): 399-432.

Collins, J. (2007)‘Syntax, More or Less’ Mind, 116 (464): 805-850.

Collins, J. (2013) Review of Pursuing Meaning by Emma Borg. In *The Review of Metaphysics*, Vol.67(1), p.153-154

Cosmides, L., Tooby, J., & Barkow, J. H. (1992). “Introduction: Evolutionary psychology andconceptual integration” In J. Barkow, L. Cosmides, & J. Tooby (Eds.), in *The adapted*

mind:Evolutionary psychology and the generation of culture (pp. 3–15). New York: Oxford University Press.

Crain S. and Pietroski, P. (2001) "Nature, Nurture and Universal Grammar" in *Linguistics and Philosophy* 24: 139-186.

Davidson, D. (1967) 'Truth and Meaning.' *Synthese,* 17: 304-323

Davidson, D. (1970) 'Semantics for Natural Languages' in D. Davidson *Inquiries into Truth and Interpretation*. OUP 1984. 55-64.

Davidson, D. (1980) 'Rational Animals' in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001. 95-105.

Davidson D. (1983) 'A Coherence Theory of Truth and Knowledge' in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001. 137-153.

Davidson D. (1987) 'Aterthoughts' to 'A Coherence Theory of Truth and Knowledge'(1983) in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001. 154-158.

Davidson, D. (1988) 'Epistemology and Truth' in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001.

Davidson, D. (1990a) 'The Structure and Content of Truth' *The Journal of Philosophy.* Vol. 87, No. 6, pp 279-328

Davidson, D. (1990b) 'The Turing Test' in D. Davidson's *Problems of Rationality.* OUP 2004. 77-86.

Davidson, D. (1992) 'The Second Person' in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001.

Davidson, D. (1995a) 'The Problem of Objectivity' in D. Davidson's *Problems of Rationality.* OUP 2004.

Davidson, D. (1995b) 'Could there be a Science of Rationality' in Donald Davidson's *Problems of Rationality* OUP 2004. 117-134

Davidson, D. (1996) 'The Folly of Trying to define Truth' *The Journal of Philosophy,* Vol 93, Issue 6, pp 263-278

Davidson, D. (1997a) 'Indeterminism and Antirealism' in in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001. 69-84.

Davidson, D. (1997b) 'The Emergence of Thought' in D. Davidson's *Subjective, Intersubjective, Objective* OUP 2001. 123-134

Davidson, D. (1999) 'The Centrality of Truth' in J. Peregrin (ed) *Truth and Its Nature (If* Any). Kluwer Academic Publishers.

Devitt, M. (2011). Deference and the Use Theory. ProtoSociology, 27:196-211

Dummett, M. (1959) “Truth” in Proceedings of the Aristotelian Society 59 (1):141-62

Dummett, M. (1991) *The Logical Basis of Metaphysics*. Duckworth

Dummett, M. (2002) “Meaning in terms of Justification” in *Topoi* 21 (1-2):11-19

Dummett, M. (2004) *Truth and the Past*. Columbia University Press.

Dummett, M. (2006) *Thought and Reality*. OUP

Edwards, S. (1994) Externalism in the Philosophy of Mind. Avebury, Ashgate Publishing Company.

Egan F (2003) “Naturalistic Inquiry: Where Does Mental Representation Fit in?” In L. M. Antony and N. Hornstein (Eds) *Chomsky and His Critics*. Blackwell Pubishing.

Field, H. (1986). ‘The Deflationary Conception of Truth’ in G. MacDonald and C. Wright (eds.) *Fact, Science and Morality*. Blackwell

Field, H. (1994). ‘Deflationist Views of Meaning and Content’ *Mind*. 103 (411): 249–84.

Fitch, W., Hauser, M., & Chomsky, N.(2005) ‘The Evolution of the Language Faculty:

Clarifications and Implications’, *Cognition*, Vol. 97, pp. 179-210.

Frankenhuis W.E., and A. Ploeger “Evolutionary Psychology Versus Fodor: Arguments For and Against the Massive Modularity Hypothesis” Philosophical Psychology Vol. 20, No. 6, December 2007, pp. 687–710

Frege G. (1993). ‘On Sense and Reference’. As reprinted in *Meaning and Reference, Oxford Readings in Philosophy*. Moore, A.W. (ed). Oxford University Press.

Fodor J. A.(1983) The Modularity of Mind. MIT Press.

Fodor, J.A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford University Press.

Fodor and Lepore (1999) “Impossible Word?” *Linguistic Inquiry* 30: 445-53

Fodor, J.; Lepore, E. (2001) “Why compositionality won’t go away: Reflections on Horwich’s Deflationary Theory.” *Ratio* XIV 350-368

French, S. (1988/1993) *Decision Theory: An Introduction to the Mathematics of Rationality*. Ellis Harwood Limited.

Gallistel, C.R.(1998) “Insect Navigation: Brains as symbol processing organs” http://www.nyu.edu/gsas/dept/philo/courses/representation/papers/Gallistel.pdf
Also in S. Sternberg & D Scarborough (Eds.) Conceptual and methodological foundations. Vol. 4 of An invitation to cognitive science. (D. Osherson, Series Editor) MIT Press

Goldberg S.C. (2007) *Internalism and Externalism in Semantics and Epistemology*. Oxford University Press.

Goldin-Meadow, S. and C. Mylander(1998) “Spontaneous sign systems created by deaf children in two cultures” Nature. Vol 319

Goldin-Meadow, S.; and H. Fieldman (1977) "The development of language-like communication without a language model" in *Science* 197:401-403.

Gupta, A. (1993) 'A Critique of Deflationism', *Philosophical Topics,* Vol 21, No. 2, pp. 57-81.

Gupta, A. (2003) 'Deflationism, the Problem of Representation, and Horwich's Use Theory of Meaning', *Philosophy and Phenomenological Research,* Vol. 67, No. 3, pp. 654-666.

Hagoort, P; Hald, L; Bastiaansen M; Petersson K.M. 2004. 'Integration of Word Meaning and World Knowledge in Language Comprehension'. In *Science* Vol 304, pp 438-441.

Hagoort, P. & Van Berkum, J.J.A. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society, B, 362,* 801-811

Hale K. and Keyser S. J.(1993) "On Argument Structure and the Lexical Expression of Syntactic Relations". In K Hale and SJ Keyser (Eds) *The view from Buildin 20.* MIT Press. 53-109

Halberda, J., Sires, S.F. & Feigenson, L. (2006). 'Multiple spatially overlapping sets can be enumerated in parallel'. *Psychological Science, 17,* 572-576.

Harnad, S. (1990) "The Symbol Grounding Problem" *Physica D 42: 335-346.* http://cogprints.org/615/1/The_Symbol_Grounding_Problem.html

Hattiangadi, A. , (2006) "Is meaning Normative?", Mind and Language, Vol. 21 No. 2, pp 220-240

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, Vol. 298

Hauser, M., Barner, D., & O'Donnell, T. (2007). "Evolutionary linguistics: A new look at an old landscape". *Language Learning and Development*. 3(2): 1-32.

Hinzen, W. (2006) *Mind Design and Minimal Syntax*. OUP

Hinzen, W. (2007) *An Essay on Names and Truth*. OUP

Hinzen, W. (2012) 'The philosophical significance of universal grammar', *Language Sciences,* 34: 635-649.

Hinzen, W. & M. Sheehan (2013) *The Philosophy of Universal Grammar*. Oxford: Oxford University Press.

Horn (eds) *Handbook of Pragmatics*. Oxford : Blackwell.

Horwich, P. (1997) 'The Composition of Meanings'. *The Philosophical Review,* Vol. 106, No. 4, pp. 503-532

Horwich, P. (1998) *Meaning.* Oxford: Clarendon Press.

Horwich, P. (2001) ‘Deflating Compositionality’. *Ratio* XIV, 4 December 2001, 0034-0006. Oxford: Blackwell Publishers.

Horwich, P. (2004) ‘ A Use Theory of Meaning’. *Philosophy and Phenomenological Research.* Vol. LXVIII, No. 2, March 2004, 351-372.

Horwich, P. (2008) ‘What’s Truth Got to do with it?’ *Linguistics and Philosophy* 31 (3), 309-322.

Horwich, P. (2010) *Truth, Meaning, Reality.* Oxford: Clarendon Press.

Horwich, P. (2010) ‘Regularities, rules, meanings, truth conditions and epistemic norms’ in *Truth, Meaning, Reality.* Oxford: Clarendon Press. 113-142.

Kamp, H. 1995. “Discourse Representation Theory”, in: J. Verschueren, J.-O. Östman & J. Blommaert (eds.), *Handbook of Pragmatics*, pp. 253-257.

Kaplan D. (1978) “On the Logic of Demonstratives” In *The Journal of Philosophical Logic* 8: 81-98

Karmiloff-Smith, A., E. Klima, U. Bellugi, J. Grant, and S. Baron-Cohen (1995) “Is there a social Module? Language, Face-Processing, and Theory of Mind in Individuals with Williams Syndrome” *Journal of Cognitive Neuroscience*, 7:196-208.

Langer, S. 1942. Philosophyin a New Key. Harvard University Press.

Lohndal, T. & Narita, H. (2009) “Internalism as Methodology”, *Biolinguistics*, Vol. 3, No. 4, pp. 321-331.

Majors, B. & Sawyer, S. (2005) “The Epistemological Argument for Content Externalism”, *Philosophical Perspectives*, Vol. 19, No. 1, pp. 257-280.

Mill, J. S. (1973) “A System of Logic, Ratiocinative and Inductive”, in J. Robson, ed., *The Collected Works of J. S. Mill* (Volumes 7–8). Toronto: University of Toronto Press.

Millikan, R. G., (1998) “Language Conventions Made Simple”, Journal of Philosophy, Vol. 95, No. 4, pp 161-180

Mukherji, N. (2010) *The Primacy of Grammar* (Cambridge, MA: MIT Press).

Newman, A. (2005) ‘Two Grades of Internalism (Pass and Fail)’ *Philosophical Studies* Vol.122 No. 2, pp.153-169

Nieuwland, M.S.; Van Berkum, J.J.A. (2006). ‘When Peanuts Fall in Love: N400 Evidence for the Power of Discourse’ In the Journal of Cognitive Neuroscience 18:7, pp. 1098–1111

Pagin P. and Pelletier J (2007) “Content, Context and Composition” In G. Preyer and G. Peter (Eds) Context-sensitivity and semantic minimalism: new essays on semantics and pragmatics. Oxford University Press.

Perry, John (1986). 'Thought without Representation', *Supplementary Proceedings of the Aristotelean Society*, 60: 137–52.

Pessin, A. & Goldberg, S. (Eds.)(1996) *The Twin earth Chronicles: Twenty Years of Reflection on Hilary Putnam's 'The Meaning of 'Meaning''*, M. E. Sharpe, NewYork.

Pietroski P.(2000). The Undeflated Domain of Semantics. *Sats: The Nordic Journal of Philosophy* 1:161-76. Reprinted, in part, in *Reading Philosophy of Language* (Hornsby and Longworth, eds.), Blackwell, 2005.

Pietroski P. (2005) 'Meaning Before Truth', in Preyer and Peter: *Contextualism in Philosophy*. Oxford University Press. 253-299.

Pietroski, P.; Lidz, J.; Hunter, T.; Halberda J. (2009) 'The Meaning of 'Most': semantics, numerosity, and psychology' in *Mind and Language*, 24:554-85.

Pinker S. (1995) *The Language Instinct*. Penguin.

Prinz J. (2006) "Is the Mind Really Modular?" in R. J. Stainton (Ed) *Contemporary Debates in Cognitive Science.* 22-36

Putnam, H. (1973). 'Meaning and Reference' In *The Journal of Philosophy*, Vol. 70, No. 19, pp. 699-711.

Putnam, H. (2013). "The development of Externalist Semantics" in *Theoria***,** Volume 79, Issue 3, pages 192–203, September 2013

Putnam, H., 1975 'The Meaning of 'Meaning'', *Language, Mind, and Knowledge, (Minnesota Studies in the Philosophy of Science)*, Vol. 7, pp. 131-193, University of Minnesota Press. [Reprinted in Pessin & Goldberg (1996)]

Quine W.V. (1951) "Two Dogmas of Empiricism" reprinted in Quine (1961) From a Logical *Point of View, 2nd ed.* Harvard University Press.

Quine, W. V. O. (1960) Word and Object, Cambridge. MA- MIT Press.

Recanati, F. (2004). *Literal Meaning*. Cambridge: Cambridge University Press.

Recanati F (2001) "What is Said" In *Synthese*, Vol. 128, No. 1/2, Pragmatics (Jul. - Aug., 2001), pp. 75-91

Recanati F (2005) "Literalism and Contextualism: Some Varieties" In G. Preyer and G. Peter (eds) *Contextualism in Philosophy: Knowledge, Meaning and Truth*. Oxford University Press. 171-196.

Robbins P (2007) "Minimalism and modularity". In G. Preyer and G. Peter (Eds.), *Context-Sensitivity and Semantic Minimalism*. 303–319. Oxford. University Press.

Schiffer, S. (2000) 'Review: Horwich on Meaning', *The Philosophical Quarterly*, Vol. 50, No. 201, pp. 527-536.

Skinner, B.F. (1957) Verbal Behavior. Copley Publishing Group.

Sperber D (1994) "The modularity of thought and the epidemiology of representations". In L.A. Hirschfeld and S.A. Gelman (Eds) *Mapping the Mind: Domain Specificity in Cognition and Culture.* Cambridge University Press.

Sperber D (2001) "In Defence of Massive Modularity". In E. Dupoux *Language, Brain and Cognitive Development: Essays in Honour of Jacques Mehler.* MIT Press. 47-57

Sperber, D. and Wilson, D. (1986). Relevance: communication and cognition. Oxford: Blackwell.

Stanley J. (2000). Context and Logical Form. In *Linguistics and Philosophy* 23: 391–434.

Stanley J. (2005). Semantics in Context. In G. Preyer and G. Peter (Eds.) Contextualism in Philosophy: Knowledge, Meaning and Truth. 221-254. Oxford University Press.

Stromswold, K. (2000). The cognitive neuroscience of language acquisition. M. Gazzaniga (ed.), The new cognitive neurosciences, second edition. pp 909-932. Cambridge, MA: MIT Press

Tarski, A. (1944). The Semantic Conception of Truth: And the Foundations of Semantics. Philosophy and Phenomenological Research, Vol. 4, No. 3 (Mar., 1944), pp. 341-376

Tomasello, M (2000). First steps toward a usage-based theory of language acquisition. Cognitive Linguistics 11-1/2 (2000), 61-82

Travis C. (2008) *Occasion Sensitivity: Selected Essays*. Oxford University Press

Travis C. (2006) "Insensitive Semantics" Mind & Language, Vol. 21 No. 1 February 2006, pp. 39–49

Travis C. (1975). *Saying and Understanding* (Oxford: Blackwell).

Wilson, D., and Sperber, D. (2004). Relevance theory. In Horn, L., Ward, G. (Eds.). The Handbook of Pragmatics (pp.607-632). Oxford Blackwell.

Wittgenstein, Ludwig, (1974) Tractatus Logico Philosophicus [TLP], Tr. D.F. Pears and B.F. McGuinness. International Library of Philosophy and Scientific Method

Wittgenstein, Ludwig, (1958a) Blue and Brown Books, Ed. Rush Rhees, Oxford: Basil Blackwell

Wittgenstein, Ludwig, (1958b) Philosophical Investigations [PI], Ed. G.E.M. Anscombe and R. Rhees, tr. G.E.M. Anscombe. Second edition.

Wittgenstein, Ludwig, (1967) Zettel, G.E.M. Anscombe and G.H. von Wright, (Eds.) Oxford: Basil Blackwell

Wittgenstein, Ludwig, (1966) *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*. Ed. Cyril Barrett Oxford: Basil Blackwell.

Zilhao, A. (2003) “From Radical Translation to Radical Interpretation and back” Principia 7 (1–2) 2003, pp. 229–249. Published by NEL – Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.