Language and *The Language of Thought* -

A linguistic perspective on the philosophy of Fodor
# Table of Contents

Introduction ........................................................................................................................................... 2  
Overview ............................................................................................................................................... 3  
1. Basic tenets of The Language of Thought ..................................................................................... 4  
   1.1. Language acquisition .................................................................................................................... 5  
   1.2. Translation: from mentalese to natural language ....................................................................... 6  
2. Shortcomings in establishment ........................................................................................................ 9  
   2.1. Learning a language is about forming and confirming hypotheses ........................................... 9  
   2.2. A predicate is learned when we learn its truth-definition ......................................................... 10  
   2.3. To acquire a natural language presupposes competence in a mental language ..................... 11  
   2.4. The basic vocabulary .................................................................................................................. 12  
3. The relation between mental and natural language ....................................................................... 14  
   3.1. Some methodological concerns .................................................................................................. 15  
   3.2. A redundant assumption ............................................................................................................. 17  
   3.3. Why mentalese is not a language ............................................................................................... 18  
4. Conclusion ......................................................................................................................................... 20  
   4.1. Can the notion of an internal language be saved? ..................................................................... 21  
References ......................................................................................................................................... 24
Introduction

The status of folk psychology and its concepts has come under scrutiny in contemporary philosophy of mind and psychology (cf. Ryle 1949; Nisbett and Wilson 1977; Churchland 1986; Dennett 1987, 1991). In an attempt to unify a naturalized view of the mind with folk psychology, Fodor (1975) presented *The Language of Thought Hypothesis* (LoT) – a hypothesis stating that thought and thinking are carried out in a physically realized innate mental language: mentalese, which is said to serve as the basis of rational agency as well as natural language acquisition and comprehension.

The assumed mental language consists of a compositional symbolic or representational system with semantic content (cf. Ibid: 200, Fodor 1987; Fodor 1990),\(^1\) governed in their composition by syntactic specifications, thus providing our cognition with a linguistic structure. According to Fodor, all research within relevant fields point towards the assumption of an innate mental language (Fodor 1975). The findings concerned suggest that rational agency has much in common with a language, these are amongst others the systematicity and productivity found both in thinking and in language. In conjunction with Fodor’s aim to vindicate folk psychology, LoT is said to be the best explanation; “the only game in town” (Fodor 1975: 55). Thus, the theory states that by having representations bearing (narrow) semantic content, capable of facilitating intentionality, folk psychology is saved. The syntax of mentalese on the other hand forms the laws of psychology, which constitute the explanatory role of cognition in behaviour as well as the base upon which rational thinking and agency is founded (Fodor 1987).

Following Fodor’s (1975), LoT is something of a research paradigm in philosophy of mind and language as well some fields of both theoretical psychology and linguistics (cf. Loar 1982, Stich 1983, Block 1986, Rey 1991, Aydede 2005). This thesis intends to show that LoT is an unsound model, unable to explain perennial questions of language and mind. Therefore, the acceptance of the theory will be detrimental both to empirical research in behavioural sciences as well as further theorizing. To unveil this, we will address LoT from an often overseen angle: how the assumed mental language is related to natural language, i.e. how are syntactically defined thought processes, containing semantic content, transformed to natural language utterances (and vice versa)?

There are two central issues for LoT from this perspective: (I) how do we acquire a natural language based on an innate mental language, i.e. how is a natural language acquired on the basis of a mental language? (II) when do we have a natural language, how is the relation realized? In other words, how do transformations occur from natural language structure to mentalese equivalent? Fodor answers (I) by arguing that natural language acquisition and comprehension would be impossible without prior fluency in a mental language (Fodor 1975: 86) – a defence that raises further question: what aspects of a mental language is innate and why fluency in a mental language is necessary to explain language acquisition? The response to (II) is given in terms of a machine analogy, i.e. transformations from natural language utterances to mentalese equivalents are analogous to the process of compiling programming code to machine code (ibid: 67). This computational analogy is highly troublesome insofar as the split in machine and programming language entails a double methodological problem: the machine language is superfluous to the theory, and the theory in itself makes it impossible to empirically validate the machine analogy.

This thesis will argue that the notion of mental processing as the kind of language proposed by Fodor cannot be maintained. We will reveal this by looking primarily at the large differences between the characteristics of natural language and mentalese. Our inquiry will show that LoT’s solution to (I) is not solid enough to assume that language acquisition requires fluency in a mental language. At the same time, LoT faces theoretical and methodological issues in its attempt to explain how thought and thinking are carried out in a mental language, concluding that LoT lacks a reasonable explanation to (II). All this taken together suggest that LoT proposes a redundant intermediary layer in form of a mental language.
Overview

In the following, Section 1 begins with a layout of LoT, as presented by Fodor’s (1975). This rather brief survey is followed by a presentation of the two areas of interest: language acquisition (1.1) and the relation between natural language and mentalese in terms of a machine analogy (1.2) – with a formulation of the questions in general and how they have been tackled from the perspective of LoT. The former of these areas of interest, language acquisition, will be discussed in detail through Section 2. We will look at Fodor’s description of how a mental language is necessary in order for language acquisition to occur. It will be revealed that this description provides an unsatisfactory explanation as to how a language is acquired. The general problem for LoT is that Fodor's view of natural language is rather crass, leaving out several important dimensions of becoming a competent speaker, more specifically aspects of language as communicative and dependent upon social factors in its everyday usage.

Section 3 will critically examine how LoT explains translation from mental to natural language. These passages will highlight the aforementioned double methodological problem. This discussion focus on what constitutes an actual language and what would characterize a private, or at least internal, language. We will find out that mentalese cannot be investigated in any reasonable way, as well as being a theoretically redundant layer. Thus, the reasoning has two parts: if the description of language acquisition cannot withstand critique, while the relation postulated by the theory is hard to maintain, then there is no need to assume the kind of mental language proposed. Section 4 ends the thesis with a summary and some general remarks concerning the notion of internal languages.
1. Basic tenets of The Language of Thought

This survey of LoT takes Fodor’s (1975) formulation as its version of choice. The reasons are simple: it is the first explicit version of LoT, while being both the most accurate and precise version. At the same time, it is “canonical” insofar as subsequent literature on LoT follows this version.

One can find at least two motives in Fodor’s writings to formulate the theory at hand: (I) “update” the common-sense folk psychology of the mind (e.g., we believe things and want things, which is sufficient to explain behavior) to fit the view of (some disciplines within) cognitive and behavioural sciences (Fodor 1987: chap. 1; App). (II) Provide a naturalistic explanation as to how thinking is actually possible and how we can argue for the rationality of thought from such a perspective (ibid; Fodor 1975). Taken into conjunction, these motivations lead Fodor to propose LoT.

In brief, LoT is a claim that cognition is realized as processes of computation over compositional mental representations. A simple definition of mentalese is: “a metalanguage in which representations of object language predicates are formulated” (Fodor 1975: 65). What makes mentalese a language is that semantic components – mental representations with intentional and causal powers – are combined according to syntactic rules. Whereas thoughts are regarded, LoT is primarily concerned with the nature of propositional attitudes (according to the theory, propositional attitudes are at the heart of the cognitive sciences quest to explain rational human agency). In order to give a materialistic account while still maintaining folk psychology, LoT suggests that propositional attitudes are realized due to the relation an organism holds to its internal representations. That is, to be able to hold propositional attitudes one must have representations capable of facilitating intentions (a version of the so called The Representational Theory of Mind).

Propositional attitudes are, according to the theory, defined by certain functional/computational relations (e.g., a belief-relation). To believe that John F Kennedy was assassinated in Dallas is the tokening of the compositional representations of JFK and Dallas with the predicate of assassination in a computationally defined belief-relation. That is, mental computational processes in terms of beliefs occur with the proposition “JFK was assassinated in Dallas” as its outcome, thereby forming the belief.

And what the [belief]-box does is, it churns and gurgles and computes and causes and the outcome is that you behave in a way that (ceteris paribus) makes it true that P. (Fodor 1987: 136).

In order to get tokened, our representations must be causally sensitive to different conditions, amongst other environmental. To believe that I perceive a cat can be tokened by visual or auditory stimuli of a cat. These tokenings are also causally sensitive with respect to each other; the tokening of one representation might token some other, in accordance with the defined syntax as well as semantic and epistemic liaisons (thereby supporting the systematicity and productivity of thinking). Here, the status of mentalese as a language become obvious: the syntax of a natural language provides restrictions on the well-formedness of sentences. Just consider the following:

(1) *Romeo loves.

(2) Romeo loves Juliet.

As a transitive verb "to love" requires a direct object, analogously cognitive computations require well-formedness over internal representations.

In formulating LoT, Fodor takes advances in cognitive science into consideration. According to him, empirical findings within this field comport to a unified explanation in terms of a mental language. The advances primarily considered regard theories stating that cognitive processes are construed as computational (sometimes called the Computational Theory of Mind, first proposed by Putnam 1980[1961]).
By assuming cognitive processes as computational, the step to a representational system is not far: every computation presupposes a medium – a representational language (Fodor 1975: 27). If these considerations are taken into account, Fodor's conclusion is that the workings of the mind has much in common with a language: there can principally be an infinite amount of representations (since they are compositional), while still being systematic and productive in the composition of representations (since they are governed by syntactic rules and how they are tokened). Though, mentalese is not a language as natural ones are. The differences between the two linguistic structures are captured in terms of a machine analogy: That is, the mental language is the language upon which cognition operates. To continue the machine analogy, one might say that mentalese as the basis of cognitive processing is hard-wired, whereas input/output communication is handled by natural language.

To end our overview, propositional attitudes are instantiated as a certain relation between mental representations (see note 2, regarding Fodor's usage of the term propositional attitude) as well as the domain which mental processes operate on, most prominently according to causal and compositional factors.

Claim 1 (the nature of propositional attitudes)

For any organism O, and any attitude A toward the proposition P, there is a ('computational'/'functional') relation R and a mental representation MP such that

\[ MP \text{ means that } P, \text{ and } O \text{ has } A \iff O \text{ bears } R \text{ to } MP. \] (Fodor 1987: 17).

... 

Claim 2 (the nature of mental processes)

Mental processes are causal sequences of tokenings of mental representations. (ibid: 17).

The two areas of interest will be reviewed in the following paragraphs: how LoT explains language acquisition and how the relation between natural language and mentalese is illustrated in terms of a machine analogy.

1.1. Language acquisition

One major question for general linguistics as well as developmental psychology (and to some extent philosophy) is how human beings can learn (or acquire) a language. As a reaction to behaviourism, the field has for quite some time been influenced by nativist theories, such as The Critical Period Hypothesis (Lenneberg 1964), which later on became incorporated into the framework of Universal Grammar (Chomsky 1981). Being a nativist theory, LoT presents similar arguments and modes of reasoning as mentioned claims. In The Language of Thought (1975), one of the most important arguments for an innate mental language is that the assumption of a mental language helps explaining the course of (natural) language acquisition. As we shall see later on, LoT develops a model for cognitive processing compatible with folk psychology by asserting an innate basis for natural language.

In short, LoT states that learning a first language is a process of acquiring the semantic properties of the target language's predicates; a process that follows the general theoretical suggestions of concept learning (cf. Solomonoff 1964), viz. by the formation of hypotheses and their confirmation (Fodor 1975: 80). Even though there are several different theoretical suggestions within the field of concept learning, Fodor merely gives one relevant reason, viz. concept learning as the only kind of learning that fits into the process of language acquisition (ibid: 34). This is captured by an entailment of concept learning: a confirmational relation between the learning situation and what is being learned. In the case of language acquisition, the hypotheses are concerned with the semantic properties of a predicate, most importantly when a predicate is true. Fodor considers truth to be defined by extension. Thus, what we learn when we
acquire a language is when its predicates are true according to their extension. When one learns the semantic properties of $P$, one learns the truth rule for $P$.

Learning what the predicates of a language mean involves learning a determination of the extension of these predicates. Learning a determination of the extension of the predicates involves learning that they fall under certain rules (i.e. truth rules). (ibid: 64).

Following this, Fodor gives an account of why mentalese can harbour the necessities to acquire a language. Mentalese as a metalanguage is assumed to be powerful enough to express the coextensive representations of natural language predicates.

What, then is being denied? [...] that one can learn a language whose predicates express extensions not expressible by those of a previously available representational system. Still, less roughly, that one can learn a language whose predicates express extensions not expressible by predicates of the representational system whose employment mediates thinking. (ibid: 86. Emphasis in original).

When we learn that “dog” denotes the extension of “domestic canine”, we learn when ”dog” is true (more precise, we learn when to apply the correct truth rule for the predicate). This procedure begins with the formation of a hypothesis (in order to form a hypothesis we need some means to formulate it) about the conditions under which “dog” might be true. If correct, the hypothesis is confirmed – the predicate “to be a dog” has been acquired.

There is more to this story: no one can learn that “Fx is G” unless one already has a way to represent G. One accessible explanation is that there must be another language, a metalanguage, whose representations are coextensive with the natural language concept of G (ibid: 81).

[...] one cannot learn that $P$ falls under $R$ unless one has a language in which $P$ and $R$ can be represented. So one cannot learn a language unless one has a language. In particular, one cannot learn a first language unless one already has a system capable of representing the predicates in that language and their extension. (ibid: 64).

In other words, to acquire natural language predicates is much like matching them onto pre-known templates. An important remark is that since all representational means are facilitated within mentalese, our expressive power can never increase by the acquisition of a natural language. It is however plausible that some cognitive processes are more economically facilitated by natural language processing than mentalese processing (Viger 2005: 314).

1.2. Translation: from mentalese to natural language

The step from language acquisition to the relation between mentalese and natural language is not all that far: We might ask ourselves why no meta-metalanguage is required to represent the truth rules for mentalese representations (Fodor 1975: 65). Apart from merely stating that mentalese is innate, there is a slightly different way to reason: by appeal to a machine analogy.

The machine analogy is based on the distinction between programming and machine code – internally machines operate on machine code (i.e. the lowest-level of abstraction for representing a computer program). In the case of machine code, a system of instructions and data is directly executed by the Central Processing Unit). Input/output processes are said to be handled by the programming code (i.e. an artificial language with a defined syntax and semantics, used to control or instruct the behaviour of a machine). The transformations from program code to machine code are made by compilation, i.e. translation from source code written in programming language to machine code (decompilation, or perhaps maybe more accurately execution, if the direction is opposite). Hence, if we transfer the analogy to human beings; the internal code (mentalese) operates “in the head” and does not have to work
between a speaker and a hearer, while operations with the environment is handled by program code, i.e. natural language. Thus, a meta-metalanguage is unnecessary since our cognition is structured to operate in mentalese.

What happens when a person understands a sentence must be a translational process basically analogous to what happens when a machine "understands" (viz. compiles) a sentence in its programming language. (ibid: 67).

As with machine and programming language, we must have some kind of comprehension of mentalese (M) and natural language (L) before we can begin to formulate how the relation is realized. It could be so that M and L are a lot alike, but then the transformations from one structure to the other would be more of an object-by-object translation, rather than compilation. Thus, there must be some qualitative differences between M and L – apart from M being the linguistic system in which truth-definitions and coextensive representations of natural language predicates are couched. Otherwise, it would be possible to assume that M was necessary only for the acquisition of L, while later operations was handled without M, i.e. directly in natural language parsing. By introducing the machine analogy, M is comprehended as necessary for cognitive processing and rational agency.

Fodor's way of addressing the issue of how L and M are related is something of a reply to The Private Language Argument (Wittgenstein 1953). As known, one part of Wittgenstein's argument stems from its originator's view of natural languages. Just as games are defined by public conventions and criteria for correctness, so are languages. In a sense, LoT is an attempt to defend the notion of a private (or at least internal) language, whose expressions are available to a subject in virtue of the relation she holds to her concepts (in this case: mental representations). As far as Ls are concerned, Fodor agrees with Wittgenstein, i.e. any given L has a communicative nature, hence dependent upon conventions in its application. However, a language such as M cannot depend on conventions. If cognitive processes are computational over representations, they must be defined by something else than conventions. In a sense, conventions are arbitrary, which in order to be comprehensible, computational processes cannot be. We will return to this in 2.3.

In order to show that there is relation between L and M, Fodor wants to reveal how it can be a language of thought, not public conventions that mediates between linguistic forms, i.e. expressions in L, and propositional attitudes (as would be the case with the later Wittgenstein). The reasoning is based on the mind's depiction in selected parts of the cognitive sciences, where cognition is comprehended as computational. If cognition is computational (which, according to Fodor, it must be), then the cognitive sciences can attempt to explain the nature of propositional attitudes by providing for each kind of propositional attitude nomological conditions in terms of computational relations between the organism and the formulae of internal representations (recall that a certain kind of propositional attitude, e.g. believing, is to be in a certain computational relation). Thus, any organism capable to hold propositional attitudes are in two relations: one to the proposition expressing the attitude, and another to the formulae. These relations are so arranged that the latter is causally responsible for the former. That is, to be in a certain relation to formulae causes one to be in a certain relation to the proposition (Fodor 1975: 77).

The befits of this reasoning is that we can account for the right types of correspondences between the attitudes one bear to propositions and the relation to formulae of the mental language. This is how M and L differs: in L the correspondence is due to conventions governed by L. In M, on the other hand, the correspondence is founded on the structure of cognitive processes. In turn, this leads Fodor to propose something of a condition of L and M:

(S uses [a is F] to represent a's being F) just in case ((S believes that a is F just in case S assents to [a is F]) is conventional)

Now consider the case where [a is F] is a formula of the internal code. Then there will be a condition which holds for the formula and which differs only in that (a) 'assents to' is replaced by a sequence of one or more of the basic relations from which computational relations to internal formulations are constructed and (b) 'is conventional' is replaced by 'is nomologically necessary' (ibid: 78).
At first glance, this quotation might be a bit formal and not easily grasped. Basically, in M computations over representations occur as a fact of psychological law, i.e. there are no conventions; cognitive operations are in terms of a language which is necessary for (human and the like) cognition to occur at all.

The next step would be to show how transformation from a conventional system to a nomologically necessary language comes about. LoT is not provided with a clear-cut picture of the transformations, instead merely some general outlines are presented. Mostly, inspiration is drawn from Transformational Grammar (Chomsky 1965). One of the essential principles of Chomskyan linguistics is the distinction between surface structure and deep structure, where all languages share, more or less, the same deep structural level. Differences between languages are shallow insofar as they depend primarily on the values of parameters in surface structure. An example would the head directionality parameter: Either the head follows (head final, e.g. Japanese, Korean, Persian and Basque) or precedes (head initial, e.g. Germanic and Romance languages) the phrase in forming longer phrases. If the correct transformational rules are applied, we can reveal that all languages share the same deep structure.

Since LoT adheres to nativism and computationalism, Chomsky's model is a good candidate for explaining the relation between L and M. Even if Fodor clearly offers a reasoning separate from Chomsky's, we can find agreement between them in the way transformations occur between L and M, which seems to be much on par with Transformational Grammar. Due to the limited space, this discussion will not be continued at length. An addition from Fodor's side is that the vocabulary of our internal representations is used as definitions of our natural language expressions. That is, in expressing “Bill is a bachelor” we make use of the representation as a definition of what a bachelor is. Natural language facilitates definable terms, whereas mentalese representations provide the definitions of these terms, much like a dictionary where truth-definitions are couched.
2. Shortcomings in establishment

This section offers a critique of LoT's description for language acquisition. Fodor's description of language acquisition will be presented as well as this description's shortcomings. The arguments of this section will stress that (I) LoT lacks an explanation as to how we can learn to actually use a language and (II) That it is unreasonable to assume mentalese as the basis for natural language since there are such differences between the mental language and natural language. The goal of this section is to reveal that LoT's description for language acquisition is invalid on both a theoretical and empirical level.

In short, LoT's description of language acquisition is three conjoined statements, each with some accompanying arguments. All of these statements will be given its own part of this section where we will both discuss and criticize them. The section ends by revealing a major problem in Fodor's reasoning, stemming from his view of language as such.

It is a rather crass analysis of learning a language that is presented by LoT. Due to scientific advances in the fields of language acquisition and developmental psychology since Fodor's (1975) one must surely doubt whether the acquisition of language depends on facilitating coextensive representations in a mental language, i.e. to acquire that 'Fx is G' is to match F with some prior linguistic means to represent G. One might find the critique presented in the following as misdirected insofar as scientific progress has been made. However, Fodor and other LoT-theorists has had every chance to incorporate these advances in to the programme of LoT, but has chosen not to do so. Instead, it is imperative to be aware of Fodor's explicit defence of LoT in later writings (cf. Fodor 1987), while analysis from the perspective of a rather unmodified version of LoT is not unknown to contemporary philosophy (cf. Rey 1995). Throughout this section the notion of an internal language will not be doubted – rather that we can explain language acquisition by assuming competence in a mental language. We will look at Fodor's model for language acquisition, which as mentioned is made up by three conjoined statements:

- Learning a language is about forming hypotheses and confirming them.
- Hypotheses are confirmed when we learn to apply the correct truth rule of a predicate (according to the extension of the predicate).
- We can neither form hypotheses nor acquire a without prior fluency in a mental language.

These three statements will be scrutinized in the following. So, let's take it from the top.

2.1. Learning a language is about forming and confirming hypotheses

That language acquisition follows hypotheses formation and confirmation is based on the notion of concept learning (see 1.1). Since Fodor proposes the relation between natural language and mental language to be in terms of a machine analogy it is understandable that LoT draws inspiration from the literature on concept learning. As discussed in 1.1., LoT means that language acquisition is concerned with learning truth rules for predicates in order to match them against co-extensive representations. One reason for Fodor to propose the notion at hand is due to concept learning as a way to maintain a confirmational relation between the learning and what is learned (ibid 34). That is, we do not happen to learn a language by being hit in the head with a hammer. In order for language acquisition to follow the guidelines of concept learning, Fodor assumes that formation of hypotheses and representation of data occur in a representational system: a language of thought in which both data is assembled and justified inductive inferences are made in order to make the correct natural language predicate “match” with its co-extensive mental representation.

As will be shown by some empirical examples, LoT disregards several essential aspects of language acquisition, in particular the social aspect. As fruitful as Fodor's suggestion might be in theorizing over the abilities of machines (and to some extent humans), forming and confirming hypotheses can hardly be considered as mastering a language. By equating language acquisition with
concept learning LoT neglects essential constituents in the process of language acquisition. To some extent it could be reasonable that we induce when to correctly apply a concept by forming some hypotheses, but language is mostly employed for other purposes than labelling objects in the world or uttering truth-functional statements. We say things like “hello” or “pass the sugar, please”, which cannot be true. They are merely used correctly or not – there is no way we can form a hypothesis about the truth-rule of “hello” (or why we sometimes say “hi” or “how do you do”). In a mundane meaning languages are about communicating, not conveying truth-functional statements. It is in science we need defined concepts, i.e. we cannot have homonyms like “bank”. Though, in our everyday life such ambiguity rarely causes problems. Thus, as will be clear, equating language acquisition with the inductive process of concept learning is troublesome.

Throughout history there has been a number of so called “feral children”. Such children have for some reason lived in isolation with none or little human care. The lack of social contact have several consequences, among others a rarely remedied lack of linguistic competence (cf. Fromkin, Krashen, Curtis, Rigler and Rigler 1974; Lane & Pillard 1978). If acquiring a language is just about confirming one's hypotheses it should be sufficient to put an infant in an attic with the necessities to avoid malnutrition and a dictionary. Apart from the biological prerequisites, it seems as we must be in a social environment to develop a language. It is of course difficult to separate just the lack of linguistic competence from the absence of other social skills found in feral children. Though, since LoT is silent about critical period and other relevant notions of language acquisition it is hard to know Fodor's position in the question. Together with the following argument we have a strong case for assuming that social contact is necessary to develop a language, thus supporting the notion that developing a language is much like other social behaviour which cannot be explained solely in terms of inductive and scientific reasoning from the subject acquiring a language.

Baldwin (1991, 1993) constructed an experiment in joint attention which demonstrates that children acquire new words only in a social context, otherwise not. The procedure of the experiment is simple: a child is presented with a, for her, novel toy, A. Meanwhile, the experiment leader sits with another novel toy, B, which she calls 'modi'. In this situation the child will understand the joint reference to toy B as being 'modi', and not toy A. On the other hand, if the child is on her own playing with a toy, A or B, when a disembodied voice calls out 'modi' repetitively the child will not react to the voice. At least no inference will be made to 'modi' as the name of the toy. That is, the child doesn't even begin to form some hypotheses.

Learning to grasp a new concept is not just a matter of hypotheses formation and their confirmation (with respect to the extension under which a predicate is true). If one wants to maintain the criticized notion one must give an account about the conditions that ignite hypothesis formation. As long as LoT does not address the issue, the best we can say is that empirical data is not incompatible with the theory.

2.2. A predicate is learned when we learn its truth-definition

That language acquisition involves learning the truth-definition for predicates is not extensively covered by LoT. Fodor merely states that “...[it] will be granted by anyone who is willing to suppose that there is anything to the notion of semantic properties as real” (Fodor 1975: 59) – i.e. it should be entailed from LoT as a version of intentional realism. In language acquisition this means that we have learned a sentence only when we have understood the conditions under which the sentence is true. Since LoT is compositional, it will apply to less complex linguistic units that sentences (e.g. predicates; though not unambiguously linguistic entities as such, Fodor treats them so). For his reasoning, Fodor assumes that truth is defined by the extension of a predicate, i.e. to acquire a predicate is to learn which objects falls under the set of the predicate. If so, truth is most reasonably defined in a Tarskian manner, i.e. the truth-value of 'P v Q' is entirely dependent on the truth-value of P and Q (Tarski 1956).
There is a problem in establishing coextensive expressions as the bridge between mentalese and natural language (with the entailment that we learn truth-definitions when we acquire a language). Just consider the case of “animals-with-heart” and “animals-with-kidneys” (cf. Quine 1956). As familiar, these two expressions are coextensive. Imagine a natural language which groups animals based on kidneys, thereby forming two words “with-kidneys” and “without-kidneys”. Mentalese will couch a coextensive expression to “with-kidneys” (we will see later on that we cannot even begin to form the notion of a basic vocabulary, so we might as well assume that any reasonable aspect of natural language forms part of mentalese) which might be “with-kidneys” as well as “with-heart”. Thus, the natural language predicate and mentalese equivalent are coextensive, but can apparently differ in intension. Even more so, it might be so that the representations couched in mentalese only share extensions with their natural language counterparts. A mental representation R and natural language predicate P might be true of the same thing at the same time, but just out of coincidence. That is, they could mean two different things but both be objects of sets true at the same time. If one still wants to maintain the coextensive criteria, there are rather few occurrences of “coextensive by necessity”, such as animals-with-heart-and-kidneys. Therefore, it is rather hard to comprehend a mental language practically capable of couching representations coextensive with natural language expressions. Irrespective of forming coextensive expressions, how are we to acquire them? That is, how do we know that “with-kidneys” correspond to a mentalese representation, e.g. “with-heart”? This is primarily an empirical matter to discover – not a question for linguistic analysis.

Though, any given theory about semantic properties can replace the extensional criteria without complications for LoT (Fodor 1975: 59). If so, the previous argument might be overcome by changing the criterion for truth. The most striking would be to shift the extensional criterion to an intensional criterion, i.e. a language of the mind able to harbour representations of all possible instances of a certain predicate. With such a criterion, the relation between natural language and mentalese would be established on the basis of all possible things a predicate can describe. The effect of such a shift in truth-definition would imply a distinctively different mental language. Therefore, even if considered irrelevant for LoT, it is as a matter of fact relevant how mental representations and natural language predicates are related to one another.

2.3. To acquire a natural language presupposes competence in a mental language

Fodor’s last statement is by far the most important to the notion of an innate mental language. As such, it is rather straightforward: learning that “Dumbo is an elephant” involves to learn when to correctly apply the truth rule for the predicate “being an elephant” (ibid: 63). As familiar by now, truth definitions are said to be based on the extension of the predicate under acquisition, which been found a highly dubtable claim (2.2.). Let’s skip this for a moment and direct our attention on what follows.

No matter how languages are acquired, LoT states that we must be able to represent what is being acquired in an internal language. If one is to learn that “Dumbo is an elephant”, one must have means to represent elephants, otherwise one can never learn that Dumbo in fact is an elephant. Now, imagine a child that’s never seen nor heard of an elephant before she sees one while watching Animal Planet. Thus, the child have no linguistic label for the big, grey animal on the screen. To explain how “elephant” can be acquired, the only solution possible to Fodor is to assume that the English lexical entity “elephant” is mapped onto a (compositional) coextensive representation in the child’s innate mental language. To learn a natural language (and to be able to form hypotheses about when a predicate is true) presupposes preceding competence in a mental language.

At first glance this might seem reasonable. However, it is problematic due to the claim that computation occurs over the representations of our mental language. At the same time, mental representations have defined truth rules. This surely adds up to the view that our representations most reasonably are distinct and defined, otherwise it would be unable to provide fruitful and comprehensible output. That is, in order for computation to yield comprehensible output we cannot have vagueness or
ambiguity in our representations. If we have ambiguous representations, then there’s an inherent possibility that separate computations yield output with different truth values – as would be the case with mental representations as “bank”. That is, which of the different meanings of ‘bank’ is to be considered? Even though it is rather easy from the perspective of natural language context to clear the ambiguity, how is a machine to resolve such ambiguity? On the other hand, if our representations are vague, then we wouldn’t know where to draw the line. Just consider “large”: what is the criteria for being large? We could of course construct many different concepts, large1, large2, large3 and so on, all with different truth conditions. But, as a whole, the cluster of the concept “large” would still be vague. Thus, it seems reasonable to assume that our mental representations are distinct and defined.

But, if our internal language is defined, refined and distinct, then how come that our natural language classifications often are fuzzy and arbitrary? Consider a rather extreme example: the Dyirbal word “balan” denotes, among other things “women, fire and dangerous things”. In this word, it is close to impossible from a linguistic analysis in terms of extensions to provide an explanation as to how this word came about (though, we can surely explain how it came about by appeal to social and historical factors). So, we have got defined mental representations and alleged coextensive expressions in natural language. Our mental language pre-dates the acquisition of natural language in ontogeny (and phylogeny?). But if so, how can it be that we start with defined expressions and end up with fuzzy and arbitrary counterparts? It would be reasonable if natural language mirrored the mental language, since we operate in such a language. If languages express how we perceive the world (cf. Itkonen 1991), it seems to follow that languages can say something about our mental categorization (cf. Lakoff 1987).

LoT might be defended in two ways: (i) The differences between natural language and mentalese are superficial (remember the distinction in surface and deep structure mentioned in 1.2). (ii) it is irrelevant whether there are differences, since the mental language is alike machine language, i.e. it will not care what enters – it just operates. Against (i) we can raise the same objection as before: natural language does not inhibit distinct predicates. Therefore, it seems strange that our mental language would not affect natural languages to some extent. It could however, by some weird coincidence be so that our mental representations are alike natural language predicates – i.e. unrefined, indistinct and so forth. Then we could have nonsensical attitudes like “I believe that x might, among other plausible candidates, be a woman, a fire or a dangerous thing”. Another way of reason would be to assume that ambiguity and vagueness are due to composition. That is, every atomic representation is distinct and defined, but when we combine then they might be given aspects lacking in each and every participating representation. Then, the whole has something missing in its parts. If so, what mechanisms add these extra dimensions to a complex representation? A fortiori, how are we to make the reductive analysis of a complex representation into atomic representations?

Lets move on to (ii), which is based on the machine analogy (see Section 3 for a thorough discussion). Basically, mentalese does not have to function between speaker and hearer, i.e. it is sufficient if the language only functions for internal operations. This might of course be true, but if two machines were to communicate, it seems reasonable and more economical to assume that they do so in a language based on their internal code. That is, the expressions of input/output language will not be different in such respects as the differences between natural languages and the mental language. Instead, Fodor can assume that mentalese is inaccessible, which would be a strange way of reasoning. If the content of our mental language is principally inaccessible to us, then how can LoT be maintained without abandoning folk psychology? The implication would be an all too deterministic solution to fit Fodor’s purposes.14 Even more so, our mental representations must be accessible to us – otherwise we would be unaware (or at least less aware) of what our attitudes are about – a position that surely is logically possible (cf. Dennett 1981).15 Since it is only through mental representations that propositional attitudes are realized, Fodor’s model doesn’t allow attitudes without explicit representation.

It could be so that the mental language is completely deterministic and without principal access to the organism, but what kind of language would that be? This might be true of the machine, i.e. it does not care or reflect over what the input, it just makes the computations it is designed to do. But, we are aware
of our natural language usage, while the machine is unaware of the meaning embedded in the programming code.

### 2.4. The basic vocabulary

We end this section by discussing some general problems in Fodor's description of language acquisition. To accomplish this we will follow a method endorsed by Fodor: by looking at how natural languages are constructed we will gain insight in the layout of the putative mental language (Fodor 1975: chap. 3). To acquire a language is said to involve the acquisition of when to apply the truth rule of a predicate. Since mentalese is innate, it is reasonable to assume that there are means to represent the predicate acquired, even if Fodor is rather vague on exactly what mental representations are like. Most reasonably would be if there either are actual pre-linguistic semantically valuable representations, or some general means to represent extensions of natural language predicates. The latter of these suggestions would have it that mentalese inhibits categories or the like to represent linguistic expressions (and other cognitive capacities) – maybe analogous to Kantian categories or categorization in terms of cognitive prototypes (cf. Lakoff 1987).

Although not directly addressing this question, LoT states that the elementary vocabulary of any given natural language L is included in the vocabulary of the mental language M (which must be rich enough to express the extensions of the predicates of L (Fodor 1975: 81-82)). Later on in language acquisition, fragments of L can become incorporated into the computational system “by something like a process of abbreviatory definition” (ibid: 85). However, the basic vocabulary of L is part of M. Hence, an elementary vocabulary is innate and the rest is acquired on this base, together with rules for composition. Now we are faced with a dilemma. What is a basic vocabulary? Having words for kinship, numbers and other close-to universal semantic fields in the languages of the world (though, see Everett 2005 for a discussion of Pirahã – a language assumed to lack almost every aspects of a “basic vocabulary”)? If the basic vocabulary is the sum of shared meaning in close-to all languages, then we will soon enough notice that there are too few words and semantic fields shared by all languages. Just consider the project of Natural Semantic Metalanguage (cf. Wierzbicka 1972). This project attempts to find semantic atomic, primitive meaning present in all of the world's languages. As of today, the list consists of 61 primitives. Even though this is a, in comparison, rather vast collection of semantic primitives, we cannot even begin to construct an actual language with the help of so few primitives.

Thus, languages differ to such a large extent that we find an elementary vocabulary on which to base an actual language is close to impossible to accomplish... Meanwhile, a basic vocabulary for one language is dependent on social, historical and cultural factors (cf. Croft 2002) – differences that cannot be accounted for by Transformational Grammar or the like. Therefore, a less narrow concept of basic vocabulary will differ from language to language. Further, the encoding of the same semantic fields within related languages will from time to time still differ in how the field at hand is mapped onto morphemes. Just consider the following:

<table>
<thead>
<tr>
<th>Swedish</th>
<th>English</th>
<th>Danish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Träd</td>
<td>Tree</td>
<td>Træ</td>
</tr>
<tr>
<td>Trä</td>
<td>Wood</td>
<td></td>
</tr>
<tr>
<td>Skog</td>
<td>Forest</td>
<td>Skov</td>
</tr>
</tbody>
</table>

Table 1. Different semantic mappings of the semantic field of tree-wood-forest in three languages. (From Allwood 1983).
These three languages are closely related (genetically and typologically). At the same time, tree-wood-forest should be a rather simple part of a vocabulary, as long as trees are present in the environment. Still, these languages differ in their mapping of this field. If there is a relation between elementary vocabulary in L and the vocabulary of M, then how can these differences occur? Somehow the semantic field of tree-wood-forest must be cut up in M as well; we can’t just have one representation meaning ‘tree’. After all, there are large differences between a pile of wood and a forest. If mentalese is the foundation for natural language, then how can it be that even typologically and genetically languages cut up semantic fields differently; and even more so, how are we to match these words onto mental representations when they differ in intension and extension?

In summary, the notion of a basic vocabulary in M is highly troublesome in that:

- The actual basic vocabulary shared by all languages concerns a limited number of semantic fields. If such a shared vocabulary is to reveal the foundation of the mental language, then we would not be able to understand or express all too much.

- A wider concept of a basic vocabulary will reveal differences between languages, thereby yielding a relativistic mental language, or one with enormous potential.

- When comparing closely related languages we will find differences in their semantic mappings, which will make it hard to represent the field at hand in M in the same manner.

The morale is that languages do in fact differ with ramifications for Fodor’s theory. If we assume that the acquisition of a natural language is dependent on a mental language we come across issues concerning linguistic diversity. It is close to impossible from the perspective of natural languages to find enough words being basic so of which we can form a defined extension (or other theory for truth-definition). To have concepts only denoting defined sets is not a distinctive quality of natural languages, once again consider the Dyirbal word “balan” denoting, among other things, “women, fire and dangerous things”. These three nouns cannot reasonable share some quality that make them all fall under one and the same (defined) extension. We conclude by stating that language acquisition cannot be explained by appeal to a mental language capable of expressing coextensive representations of the extensions of natural language predicates. Therefore, the notion of an innate mental language a’la LoT becomes hard to maintain.

The conclusion is not a refutation of the logical possibility of a mental language. Instead, the kind of mental language outlined by Fodor cannot account for natural language acquisition. Since language acquisition is one of the cardinal arguments for mentalese, we must doubt the entire framework of LoT. The only way LoT can be saved is by a reasonable explanation as to how natural language operates on a mental language. We will look closer at this conundrum in the following section.
3. The relation between mental and natural language

As you might remember, there were two concerns for LoT concerning the relation between natural and mental language. Apart from language acquisition, we are still faced with language comprehension: how is the mental language actually related to natural language? Fodor's way of addressing the issue is by a machine analogy. We will see that there are both some methodological and theoretical inadequacies of the analogy. The section ends with discussing whether mentalese can be said to a language at all.

The machine analogy is introduced in order to describe how mentalese is related to natural language on an intra-personal level, i.e. how do transformations (analogous to compilation) occur from natural language structure to a mentalese equivalent? The machine analogy arises from the need to explain why an additional meta-metalanguage is unnecessary – the machine only need machine code to operate. In other words, input/output language is compiled to machine code on which the machine is designed to operate. The analogy states that the same is true for human beings, natural language is “compiled” into a mental language upon which our cognition operates. This mental language is hard-wired in the same way as machine code is in the machine, i.e. there are certain modes of operations that the mind can accomplish. Such rules of inference as modus ponens and hypothetical syllogism are plausibly hard-wired to administer the core of LoT: reasoning and rationality.19

Fodor must give an account of both mentalese and natural language before the issue of transformation from one kind of linguistic structure to another (or from data structure to program structure) can be addressed. As a reminder, Fodor considers natural languages as conventional where mentalese is nomologically necessary (see 1.2.). In other words, natural languages are to some extent arbitrary while mentalese complies with psychological laws. The next step would be to describe how the transformations actually occur and why they are necessary to explain rational agency. As mentioned in 1.2., this is done by appeal to Transformational Grammar. When it comes to LoT's goal to outline a similar analysis for semantics, we find that the field, Generative Semantics, is long abandoned (and never endorsed by “canonical” Chomskyan linguistics).20 Due to the lack of unanimity, together with the contemporary unwillingness to continue such research, we will leave it at that (however, see Chomsky 1957, 1965, 1981, 1995 for different versions and suggestions).

Instead of dwelling at Transformational Grammar, let's move on to the discussion of the relation between mental and natural language in principle.21 After all, it's been 30-some years since the publication of The Language of Thought and it's an open question whether we someday can find good enough methods to describe the alleged relation in terms of transformations. In the following, the heart of the machine analogy will be addressed, viz. the theoretical assumption that natural language expressions are transformed to mental language by a process analogous to compilation. We begin by discussing some methodological problems of the analogy in 3.1, before revealing in 3.2. that mentalese is a superfluous layer, thus a theoretically redundant assumption. Together with the glaring lack of satisfying description of language acquisition (see Section 2), the conclusion is that LoT is a theory that's served its days.

3.1. Some methodological concerns

If there is a language of the mind as depicted by Fodor, how are we to empirically investigate the claim? It is an explicit statement of LoT that mentalese can express coextensive expressions of natural language predicates:

One can learn L only if one already knows some language rich enough to express the extension of any predicate of L. (Fodor 1975: 80).

For each elementary predicate of L, there will be at least one coextensive predicate of M. (ibid: 81).
Fodor states that all L-expressions have coextensive expressions in a mental language M. However, we can turn it around and state that, if not all, at least many predicates of M have coextensive L-predicates. This might not be endorsed by Fodor, but since L’s elementary vocabulary is based on M it seems reasonable to assume. In other words, some predicates of M are expressible in L as well – otherwise we would not be able to communicate our attitudes. Then, how are we to find out whether M actually exists? Every time a sentence is uttered it will be in a natural language (with the exception of Esperanto and other constructed languages, which are based on natural languages).

As briefly discussed in 2.3, mentalese might just be the “language of the machine” insofar as cognitive processes operate on input in such a language. In other words, just like the machine is unaware of its computations, so are we of our cognitive mechanisms; we know at most what enters and what exits. But, if folk psychology is to withhold, then some degree of conscious access to one’s thinking is necessary. Otherwise, we wouldn’t be able to know anything about our beliefs and desires. If thinking is merely the language of the machine, then our attitudes might be real but they are only the effect of causation (as mentioned, Fodor doesn’t accept attitudes without explicit representations). To some extent, this is the view Fodor acknowledges (cf. Fodor 1990), but only insofar as computation goes. That is, we might be unaware of the actual cognitive mechanisms behind the formation of attitudes. Still, LoT cannot be completely deterministic.

Neither you nor I have any access to the mechanisms underlying the formation of our beliefs – but we do know, at least in a common sense meaning, what we believe and why we hold a particular belief. Since this is the case, we are unaware of the syntax of mentalese, but so are we mostly in our native tongue as well.26 On the other hand, just as we know the meaning of the words we use, it is reasonable to assume that we are aware of the semantic content of our representations. If we do not know what our beliefs are about, folk psychology would not be saved. Further, from a methodological perspective, how can we know that our mental representations bear (narrow) semantic content if mentalese is inaccessible? If mentalese is to be understood only as a language on which cognitive processes operates it is a rather odd language. Even more so, it would be a meaningless assumption, since it would not only be methodologically inaccessible, but principally. Thus, it would only be available in second-hand, through our natural language expressions. One way to, if not to save LoT, then at least give the theory some peace of mind would be by a reasonable account as to how transformations occur from one linguistic structure to another. This is said to handled by some cognitive mechanism analogous to a compiler.

The compiler

A question that needs to be answered in some way is by which means the transformations from one linguistic structure to another can come occur. LoT answers the question by assuming that some mechanism, analogous to a compiler is realized somewhere within the apparatus of language comprehension and speech production. There are several unresolved issues concerning this assumption, which will be briefly addressed in the following.

It is a fact of computational theory and practice that there aren't any semantically valuable objects in the operations of the machine unless it was written into the source code. If we transfer this back to LoT the implication is obvious: the machine analogy can only be valid only if the representations of the mental language are dependent on its natural language counter-expressions. But, mentalese pre-dates natural language in ontogeny and in speech production, thus there cannot be any semantically valuable objects. It follows that our mental representations bear semantic content only if their natural language counterparts do so, which is incompatible with LoT. Further, if we were to form another version of mentalese, which were dependent on natural language, then we would have fuzzy and unrefined expressions in computation (see 2.3). Such a proposal is hard to imagine from a computational perspective, since a compiler of unsurpassed qualities would be required.

The compiler must further have the ability to translate from any language of the world (and uninstantiated, but possible languages) to mentalese. This require the compiler to be an extremely
powerful tool – compare it to one computer compiler capable of compiling all possible Turing-complete languages. Fodor never states where or how the compiler is realized: if it's part of mentalese the analogy becomes odd – no machine has an inbuilt compiler; the only reason to have a compiler is for simplicity. The machine cannot do more if programming code is written in C or Assembler – the machine still operates on machine code. That is, the process of compilation is due to simplicity, not a necessity to communicate between man and machine. Following the point recently made: the compiler of our mind can in principle compile all natural languages. If so, then why are we unable to understand a foreign language? Even if we cannot make out the meaning of words, we should be able to compile sentences in a language unknown to us and access at least basic syntactic properties, most importantly where a word starts and ends (even though it is hard to know where one word starts and ends in spoken language). But, often we cannot even start to understand such aspects of a foreign language. Since such basic operations cannot be made, how are we then to assume that the compiler can accomplish anything? A fortiori, why are we to believe in its existence and in the validity of the machine analogy?

3.2. A redundant assumption

To the author's knowledge, all kinds of assumed internal languages are said to mediate between human agency, the internal and the external world. That is, I see a grey, big, animate object with a trunk and large ears – for myself I label this object “goggabonk”. In my internal language, this specific label mediates between perception and thinking. However, due to the nature of natural language as communicative, each time I am to tell you about “goggabonk”, I must adapt my internal language utterance to our commonly shared language – otherwise you wouldn't understand me (unless I offered you my stipulation). What would exactly be the reason to have “goggabonk” as my internal way of representing in favour over the conventional and commonly shared “elephant” (apart from its silly, yet appealing sound)? The answer is, out of simple reasons, none. We have two types of linguistic systems: S1, of which we have evidence for its existence in the fact that it is used constantly by literally billions of people everyday, and S2, which is said to be the foundation of S1. We have, as discussed in 3.1, no direct evidence for S2. Why are we then to assume the existence of the system? From a theoretical perspective, it seems rather superfluous, since S2 serves the same purpose as S1, with the difference that it only operates internally, whereas the evidence for S1 comes from its external usage. Since we can use S1 in our interaction with other humans it seems reasonable that we can use it to communicate with ourselves as well. It must be rather uneconomical to harbour two mental lexicons instead of one.

Let's apply this to LoT. Natural language can express everything that mentalese can. Then, why are we to assume its existence? Following the just preceding argumentation, with a hint of the machine analogy, we are fine with the programming language (since it can express everything of the machine) and the electrical states (corresponding to our neurological and cognitive states).

The point is that logical operations can be defined for input language propositions directly, despite Fodor's apparent denial of this for natural language. (Fortescue 1979: 72. Emphasis in original)

Thus, the flaw of the machine analogy is that mentalese in some respects resembles natural languages. Then, the machine analogy cannot explain the need to establish a language of the mind. There might be neural processes corresponding to the electrical states of a machine which are needed for execution. But, if rules for computation are embedded within the input language then we hardly have any need to assume mentalese in the first place.

LoT might be defended by stating the necessity of a mental language to explain propositional attitudes. However, there is nothing controversial in assuming that propositional attitudes take linguistic forms (Stoecker 2001). What is odd is to assume that these linguistic forms are in a distinct non-natural language. There is no inherent need for propositional attitudes to be formulated in a mental language – they can equally well take the form of natural language. The question is why it is not so, i.e. why does
propositional attitudes take another linguistic form than the kind provided by natural language?

3.3. *Why mentalese is not a language*

Mentalese must resemble natural language in several crucial respects, otherwise it could hardly be “the language of thought”. We have seen that neither arguments from language acquisition nor the notion of a machine analogy can aid LoT in its project of providing a naturalized basis for folk psychology. To wrap it up, let’s find out whether mentalese can in fact be called a language.

For quite some time, linguistics has been the discipline in charge of natural language studies. If natural languages are “based” upon mentalese we should be able to study mentalese by employing the same means as when investigating natural language. Even if several dimensions of contemporary linguistic research might be inapplicable at mentalese, we should be able to apply some of the most basic linguistic concepts on mentalese as well.

One of the most basic and plausibly most essential notion in all linguistic (and semiotic) research is the distinction between *signifier* and *signified* (see Saussure 1977[1916]). The distinction is a general claim about both written and spoken linguistic signs. If we are to look at orthography or phonology we will soon enough find that there are no specific reason as to why a certain string of phonemes or graphemes are attached to a certain concept (i.e. ‘elephant’ might as well be called ‘goggabonk’). The relation between concept (signified) and its corresponding sign (signifier) is to some extent arbitrary. For a system to be a natural language it is close to a necessity that this criterion is met, otherwise the relation between representation (sign) and what’s represented (concept) collapses. In mentalese, they are one and the same – the concepts are instantiated in their representations. Thus, one of the most basic distinctions in natural language cannot be found in mentalese. This takes us back to a discussion in Section 2.3: if mentalese is the foundation of natural language, how can it be that our mental language doesn’t affect, or mirror, our natural language?

The distinction of Saussure leads us to another remark: when we do have an arbitrary designator for something, then we can begin to talk about meaning. By equating the designated with its designator, the concept of meaning as comprehended in all kinds of linguistic and philosophical analysis loses its purpose, which doesn’t seems to bother Fodor at all.

I’m not particularly anxious that the theory that naturalizes the semantic properties of mental representations should work for smoke, tree rings or English words. On the contrary, I’m prepared that it should turn out that smoke and tree rings represent only relative to our interest in predicting fires and ascertaining the ages of trees[...], and that English words only represent only relative to our intention to use them to communicate our thoughts. I’m prepared, that is, that only mental states should turn out to have semantic properties *in the first instance...* (Fodor 1987: 99. Emphasis in original).

For someone schooled in linguistic and philosophical analysis of language this doesn’t make sense. We can’t equate the concept of natural meaning and non-natural meaning in the way Fodor does. Smoke doesn’t “mean” fire in the way ’dog’ means ’domestic canine’. This objection set aside, the suggestion that a semantics (albeit naturalized) is only applicable to mental representations is outrageous. If this was the case, then we would have to abandon our search for (natural and non-natural) meaning, since it’s useless “in the first instance” and only relevant relative our interests. If we talk about semantics in the way Fodor does, then all basic and essential notions employed in linguistic and philosophical analysis is pointless, since “real” semantics doesn’t apply to such analysis. If we want to study language we cannot without convincing arguments abandon concepts fundamental to our research. Distinctions such as signifier/signified and meaning/reference collapse when we comprehend semantics in terms of LoT.

Though, it could be so that mentalese and natural language doesn’t share all that many traits, but they are analyzable in the same terms. That is, by dissecting natural language and mentalese we might be able to reveal that they are similar (and related). To attain this, one suggestion could have it that both are
comprehended in terms of some (even more) basic level, such as neurological. However, LoT explicitly states that natural languages are compiled to mental code, thereby disapproving the counter-objection. At the same time, what would the proposal exactly yield? If some means of analysis can show that mental language and natural language are comprehended on equal terms, then we would still be facing the redundancy problem raised in 3.2.
4. Conclusion

The attempts of this thesis was to look at the relation between an assumed mental language and natural language. One must say that LoT is incompatible both with empirical and theoretical issues. The two central aspects for LoT, at least when it comes to natural language, were covered in this thesis. These two aspects were language acquisition and how the relation between natural and mental language is analogous to the relation between machine language and programming language. We found that the description of language acquisition is at best not incompatible with empirical findings concerning the relevance of a social context in acquiring a language. At the same time, the description faces the problem of discrepancies between mental representations and natural language concepts, insofar as the latter do rarely share the distinct and defined qualities required of mental representations. Since mentalese is the basis of natural language, it is rather odd that mentalese characteristics is lacking in natural language. As previously argued, the only way to save LoT would be by giving a reasonable account of how transformations from the linguistic structure of natural language to a mentalese equivalent occur. However, we found out in that mentalese is a redundant layer which cannot be empirically nor given a theoretically meaningful interpretation. Further, we have to question whether mentalese really can be called a language. Since several basic notions in linguistic research seems to be inapplicable at mentalese, it is hard to maintain the notion of mentalese as “the language of thought”. All in all, the description of language acquisition fails to aid LoT, while several theoretical inadequacies harm the theory in such ways that we can only conclude that LoT has served its days.

So, how did we accomplish all this? The first part of the critique, Section 2, dealt with language acquisition, i.e. does LoT's description for language acquisition fit with how we comprehend these issues? Though, this aspect can be broken down: do we have an innate capacity to represent our surrounding world and is this representational means in the form of a language? Emphasis was on the latter, while the former fell beyond the scope of this thesis. That we do have some means to represent the world cannot be overlooked, however that it would take the form of the mental language as outlined by Fodor is highly doubtful.

Instead of continuing down the line of LoT, we can conduct research in psycholinguistics and neurolinguistics to hopefully enhance our knowledge of language and mind. If we are to draw the contours of such a map, the fundamental notion is to acknowledge the fact that language mirrors how the world is perceived by human beings. By deepening our knowledge of language and its role in cognition, we will also increase our knowledge of the mind.

LoT's theoretical base in terms of the machine analogy introduces a methodological problem insofar as it's principally inaccessible to study in overt behaviour. Introspection, apart from being a dubious method, cannot help us either. We can of course learn more in cognitive science and neuroscience, but another objection arises: why are we to assume an intermediary layer (where propositional attitudes are realized) in terms of a language between brain states and behaviour? There might be reasons, e.g. to maintain rational agency from a physicalistic perspective. However, the machine analogy proposes a superfluous layer. A physicalistic, machine-inspired world-view is fine with just the neurological states, corresponding to electrical states of a machine, and the overt behaviour (i.e. the output of the machine). As you're all aware, a lot questions still needs to be answered within this field. Though, further theorizing from the paradigm of LoT can never aid, only harm the scientific process.

The purpose of this thesis might not be at the heart of LoT's matter – the nature of propositional attitudes. Instead, we have looked at another aspect, viz. the assumption of an internal language and its relation to natural languages. As far as the concerns of this thesis goes, it's still an open question whether rational agency and intentionality can be explained in terms compatible with folk psychology. LoT might or might not be “the only game in town” as far as unifying folk psychology with scientific progress goes. However, when looking closer upon the relation between natural and mental language, LoT becomes close to impossible to maintain.
All this taken together, I do believe that it is about time we start to realize that language is not separated on a cognitive level from other kinds of behaviour. Following Zeitgeist, I find that Fodor is wrong in treating language as a special mode of our mental life (i.e. as a module, see Fodor 1983). Instead, scientific progress is made by acknowledging the non-unique status of language. If language is “just” distinctively human behaviour and plausibly of other apes, then we can start to elaborate on which aspects are related in the founding and forming of rational agency and the relevance of language to such a machinery. By this shift of perspective related topics, apart from a conceptual analysis of language, will be incorporated in our research – such as anthropological linguistics, developmental psychology, psycholinguistics, sociology, biology and philosophy to mention a few.

4.1. Can the notion of an internal language be saved?

Whether or not we can maintain the concept of an internal language as a reasonable notion has to be answered with “it depends”. In particular, how one comprehends “internal” and “language”, but most importantly, what purpose one’s alleged language is supposed to serve. In the following, let us have LoT as our point of origin. LoT has a semantic side (mental representations) and a syntactic side (computational mental processes). We can then deny or at least downgrade either of these sides. If the semantic aspect is denied we end up with a model where cognitive processes take the form of syntactic parsing, irrespective of content. The reason for denying semantics could be that its unnecessary from a scientific view of the mind (cf. Stich 1983). However, if we would just remove the semantics from mentalese we would end up with correct computation and composition, but without any concept of what’s being parsed – if only the rules of inferences are available we wouldn’t be able to explain rational agency (that is, the phrase “colourless green ideas sleep furiously” is as reasonable as “Berlin is the capital of Germany”).

The opposite version is certainly possible – i.e. representations without rules for computation. However, this would at most be some version of intentional realism. We would have representations, bearing the power to explain rational agency. We could hardly call it a language, since the representations are not governed in their composition by rules – combinations would only be due to semantic content and epistemic liaisons.

Another problematic aspect of LoT and related models are the innate status of the language, as noted by Fortescue (1979: 68):

Does he [i.e. Fodor] seriously propose, for example, that a baby ‘knows’ the concepts ‘flying’ and ‘machine’ (if not ‘airplane’) prior to any encounter with an exemplar of such an object – as opposed to merely having the perceptual apparatus for recognising the reoccurrence of entity types. (Emphasis in original.)

Throughout this thesis, the objection of Fortescue has been mostly overlooked due to more pressing issues. Nevertheless, it is a relevant remark and must be met by every strong nativist model such as LoT. Fodor defends himself by stating that LoT is “the only game in town” and any theory is better than none. I, for one do not know how to explain the close to miraculous process of first language acquisition – but this thesis has revealed that it cannot be done by appeal to an innate language such as mentalese. It might be enough to assume that we have the innate capacities needed to acquire a language – just as we have the capacities to walk on two legs or to learn trigonometry.

From the viewpoint of LoT and the purposes of the theory its hard to construct a workable notion of an internal language that’s supposed to be (i) innate, (ii) distinct from natural languages and (iii) the basis of both natural language and rational agency. To sum up, it seems reasonable that we have some way to represent our surrounding world, however that it should take the form of an innate language is a more dubious notion.
An known, the implications of the thought-experiments of Putnam (1973) and Burge (1979) questions the status of internal semantic content. Trying to adhere to some version of internalism, Fodor claims that at least “narrow” semantic content is internal (as opposed to wide, external content).

Please note Fodor's idiosyncratic usage of the term propositional attitude: attitudes do not need to be directed towards an abstract proposition (cf. Russell 1918). Instead, a propositional attitude is said to be the effect of mental processing over mental representations, i.e. a propositional attitude is “the object” of a certain relation between mental representations.

This is Fodor's version of The Representational Theory of Mind – LoT's only concern is intentionality and propositional attitudes, thus being silent about whether other kind of mental states such as qualia can be represented in the same way (although, note Rey (1991)). That is, according to LoT, mental representations are intentional and bear causal power.

Please note a tricky but imperative side to LoT: believing is functional, though to believe that P is not functional, rather defined by the properties of P together with the outcome of mental processing (Fodor 1987: 83). Since Fodor proposes an anti-functionalistic stance concerning mental state, to believe 'that P' cannot be reduced to functional relations, even if believing as such is the effect of functional/computational processes.

Principally, anything can cause the tokening of a certain representation. To avoid meaning holism, Fodor proposes that any tokening of a representation is asymmetrically dependent upon its “proper” cause. Perceiving a cat might cause the tokening of the representation 'dog' iff dogs token the 'dog'-representation (see e.g. Fodor 1987, 1990; Devitt 1991).

If the picture isn't complicated, LoT's view on cognition and its processes is much like the analysis of language provided by Chomsky's Generative Grammar (Chomsky 1965).

LoT is however silent about important notions in the empirical research on language acquisition, such as LAD (Language Acquisition Device) and The Critical Period Hypothesis.

There is a difference between mentalese and Wittgenstein's depiction of a private language. In the latter, the meaning of a concept is stipulated by the subject, while the former states that the same basic semantic content is realized in all subjects in virtue of same innate representational and computational means.

In later writings (Chomsky 1995), the distinction in deep structure and surface structure has been replaced by Logical Form (LF) and Phonetic Form (PF).

The head is the morpheme that determines the category of a compound, or the word that determines the syntactic type of the phrase of which it is a member.

Chomsky's theory is too extensive, as well as under constant revision, to even begin to elaborate within the limited space. One major difference between Chomsky and Fodor is the latter's emphasis on the semantic aspects of a mental language. The original version of Transformational Grammar only deals with the syntactic aspects of language. Fodor, on the other hand, wants to provide a similar analysis even for the semantic part of language (See Section 3).

Interestingly enough, the findings of Baldwin was used as an argument against connectionism – in many respects the opposite of LoT.

Note that children with high-functioning autism acquire 'modi' as the name of the toy even with the disembodied voice.

Intentional realism states that mental phenomena is about something in an irreducible sense. Thus, our mental states have real content which cannot be reduced in functional or neurological terms.

LoT has received considerable critique for being an all too deterministic model of the mind – especially from the connectionist encampment (see Aydede 1997 for a discussion on this debate)

The question of aboutness is central in Fodor's writing. The topic is though all too vast to cover within the frames of this thesis.

The appeal to the notion of deep structure isn't much help either; the last bastion of Transformational Grammar: recursivity is missing in the aforementioned language of Pirahã.

Just consider the empirical impossibility in constructing a persistent Proto-World language.

See Stanford Encyclopaedia entry on LoT.

However, some of the ideas emanating from generative semantics are proposed by contemporary Cognitive Linguistics (see Newmeyer 1986: 138).

A problem with following a transformational method alike Chomsky's is what level of analysis we should endorse. Fodor seems to treat predicates as more or less the level most fit for analysis, while Chomsky prefers phrases (i.e. a single unit in the syntax of a sentence).

It is close to an axiom in much linguistic research that our metalinguistic competence is low.

As in Wittgenstein's private language in which private sensations are given linguistic labels.
References

Allwood, Jens

Aydede, Muhrat

Baldwin, D. A
1993 Infants’ ability to consult the speaker for clues to word reference. Journal of Child Language 20: 395-418.

Block, Ned

Burge, Tyler

Chomsky, Noam

Churchland, Patricia

Croft, William

Dennett, Daniel C

Everett, Daniel L.

Fodor, Jerry A

Fortescue, Michael

Fromkin, V.A., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M.

Itkonen, Esa

Lakoff, George
1987 Women, Fire and Dangerous Things: What Categories reveal about the Mind. Chicago, IL: Chicago

Loar, Brian 1982 Must beliefs be sentences? PSA, vol 2: 627-43.


Nisbett, Richard E. and DeCamp Wilson, Timothy 1977 Telling more than we can know: verbal reports on mental processes. Psychological Review vol. 84 nr. 3: 231-259.


Stoecker, Ralf 2001 How individual are intentional states really? Language and Communication 21: 167-175.


Electronic references

Stanford Encyclopaedia entry on Language of Thought:

http://plato.stanford.edu/entries/language-thought (retrieved 092507)