The Social Science Review
A Multidisciplinary Journal
ISSN: 2584-0789
(Open-Access, Peer-Reviewed, Refereed, Bi-Monthly, International e-Journal)
Homepage: www.tssreview.in

THE LANGUAGE OF THOUGHT HYPOTHESIS

Dr. Sudipta Goswami
Assistant Professor, Department-Philosophy, Balarampur College, Purulia, West Bengal, India

Corresponding Author: Dr. Sudipta Goswami

Abstract

The famous American Philosopher of mind, Jerry Alan Fodor defended a functionalist view of the mind. Fodor claimed that thinking is performing computational operations on mental representations. These inner representations form a system with many of the basic properties of a language, so the system can be called 'the language of thought'. This inner language of thought is not public, like English, but innate. According to Jerry Fodor's famous hypothesis of language of thought, we can say about the nature of a certain class of mental states, those variously known as intentional states, states with content, or propositional attitudes, namely, that they are sentence like. The language of thought hypothesis comes in two parts. The first concerns how a mental state gets to express a certain proposition; the second tells what attitude is being taken to that proposition — whether it is believed, hoped or feared or whatever. According to Fodor, there are three advantages of admitting the language of thought hypothesis: (a) to explain systematicity and productivity of thought (b) thought must be linguistic to be able to explain its complex feature and (c) language is a vehicle of thought; since natural language cannot be such a vehicle, mentalese must be taken as such a vehicle. The language of thought is the systematic expression of different propositions. The language of thought is also productive because, we can think about new thoughts and there is no upper limit to the new thoughts that we can think, except those imposed on us by our status as finite beings. Beside these thoughts are complex; different thoughts have different effects. Finally, thought evolves causally over time and one thought leads to another by causing it.

Keywords: *Philosophy of Mind, Functionalism, Thought, Hypothesis, Language, Sentence Like Structure, Productive, Evolve Causally*

The Language of Thought Hypothesis

Jerry Alan Fodor was one of the most influential figures in the "Philosophy of Psychology" and "Cognitive Science" during the latter part of the twentieth century. His primary concern has been to argue (vigorously) for a certain view of the nature of thought. According to him, thinking is an information processing within 'the language of thought'. The mind can be understood as a computer which directs action with the aid of internal representations of the world. Like several other philosophers, Fodor in the 1960s defended a functionalist view of the mind. Functionalism makes possible a physicalist world-view but does not seek simply to reduce sciences such as psychology to physics. The computer model of the

mind has been central to functionalism. Fodor, however, developed a more literal application of this model than other functionalists.

For Fodor, the value of the computer-model lies in the light it sheds on reasoning. Belief, planning and other intelligent thought processes (He does not see functionalism as a way of explaining the first-person 'feel' of experience). Fodor claims that thinking is performing computational operations on mental representations. These inner representations form a system with many of the basic properties of a language, so the system can be called 'the language of thought' (Fodor 1975). This internal language is not identical to any public language, such as English. Rather, it is used in learning public languages. Thus, the language of thought is both innate and universal.

Fodor holds that to believe that limes contain vitamin C, for example, is to have in one's head a certain sentence-like formula in the language of thought. A formula is a belief that limes contain vitamin C in virtue of both its internal causal role, which makes it a belief rather than a hope or desire, and also in virtue of connections to the external world, which determine its content. These formulae are made up of mental 'terms' with their own properties of meaning and reference. The terms combine to generate the truth condition of the whole. There are also 'implicit' beliefs which follow trivially from the explicitly represented ones, but only explicitly represented beliefs contribute to thought processes.

Though the inner formulae have semantic properties, their causal contributions to thought depend only on their formal or 'syntactic' properties; these are the properties relevant in computational processed. Indeed, these semantic properties depend partly on the nature of the thinker's environment, and what is outside the head cannot directly affect the production of behaviour. Nonetheless, Fodor insists that there are many psychological generalizations that are naturally and perhaps, necessarily expressed in terms of the semantic content of thoughts. For sometimes Fodor posited inner code with representational properties, but did not have a theory of how such representations were possible. He tried to solve this problem with his 'asymmetric dependence' theory of meaning (Fodor 1990). This is a variety of informational or indicator semantics, based on causal and law-like connections between thoughts and their objects. We cannot simply say that whatever can cause a representational state is represented by that state, since then error would be impossible. Fodor's proposal is that an inner symbol 'horse' represents horses if occurrences of this symbol do so because of the connection between the 'horse' and 'horses'. These are the central elements of Fodor's view, or the Computational Representational Theory of Mind.

In the most famous version of the computational model, intentional states are thought of as requiring the occurrence of sentence-like structures in the head; a language of thought or mentalese, in Jerry Fodor's terms. Just as to say that it is raining involves the occurrences of some sentences, for example, "it is raining", and to believe that it is raining involves the occurrence of some mental items analogous to such a sentence. Such states have a structure like syntax. They include recurring structural types ordered in systematic ways, just as in a language proper. There are other models, however apart from inner representations having a sentence-like shape. Some philosophers, for example, have argued that the inner structures might function more like a map than a sentence. Here we will concern ourselves primarily with the language of thought model, partly because of its prominence, and partly because it illustrates the task facing those who wish to reduce rationality in a particularly clear way.

To begin with, Fodor's language of thought is a hypothesis. Fodor is a scientific psychologist who thinks that empirical research will explain the workings of the mind. In his view the language of thought hypothesis best accommodates the data. For Fodor, the data requires us to adopt a degree of intentional realism, the view that there really are intentional states with genuine intentional properties bearing rationalizing links to each-other, which are causally productive in relation to behaviour, and that a scientific psychology will contain laws formulated in terms of such states. Such a view, however, also has to be put together with a physicalism in which at the most basic level, the world can be described without reference to intentional properties. The solution is to reduce intentional properties to something else.

For Fodor, something like inner sentences is the best candidate for such a reduction. They retain at the level of non-intentional realization the surface characteristics in terms of which beliefs and desires, the

propositional attitudes, are in relations to sentence-like entities. This explains how different intentional states can be focused on the same content and the features of productivity. We are able to think and understand a potentially infinite number of thoughts of the ingredients are systematic combinations of basic ones. It also explains the parallels between intentional contents and spoken and written sentences. It also can accommodate the opacity mentioned above, since different symbols in the language of thought can have the same referent (on which, more below), but distinct intentional states will be individuated if they involve relations to different symbols and, of course, such a model respects the rationality constrains, so that the causal interaction between states hold in virtue of formal structures seen as the analogue of logical structures.

According to Jerry Fodor's famous language of thought hypothesis, we can now say something rather general about the nature of a certain class of mental states, those variously known as intentional states, states with content, or propositional attitudes -namely, that they are sentence-like. Propositional attitudes are referred to as "intentional states". The term comes from the fact that a propositional attitude is about what the relevant proposition about. If you desire that it rains soon, your desire is about what the proposition expressed by the sentence 'it will rain soon' is about, namely, rain in the near future. According to the language of thought hypothesis, not only do certain sentences serve to give the propositional objects of beliefs, desires and thoughts in general, but also in addition, the thoughts are themselves sentence-like. What Fodor proposes is that these mental states have constituent structure such as sentences do. A sentence may be viewed as made up of syntactically significant parts-the parts we recognise when we parse it by identifying subject's verb and so on -put together according to certain rules. In the same general way, according to Fodor, thoughts have parts put together in certain ways. The idea is not, of course, that thoughts look like sentences, that, if we dive into someone's head, we will see states that look like writing in grey matter, as opposed to writing in ink on paper or that look like the patterns of soundwaves that are spoken sentences. It is that, thoughts have a sentence-like structure in that, for instance, the thought that, 'snow is white' and the thought that 'snow is cold' may share a part and this does not mean merely that the sentences used to report them, share a part, but that, the thoughts themselves do and it is by virtue of sharing this part that both get to be thoughts about the very same stuff: namely snow.

The language of thought hypothesis comes in two parts. The first concerns how a mental state gets to express a certain proposition; the second tells you what attitude is being taken to that proposition – whether it is believed, hoped, feared, desired, entertained or whatever. The story about how a mental state gets to express a certain proposition is that it is made up in a combinatorial way from atomic components which themselves have representational or semantic properties- that is they stand for things, properties and relations, much as the part of a natural language sentence do, and thereby get to have meanings. Words are the atoms of natural language. In English 'biscuits' represents biscuits and 'crisp' represents some properties of biscuits, which makes them crack satisfyingly when chewed. But these atoms do not express propositions by themselves. We must combine them in various ways to make claims capable of being true or false about the world. Thus, 'biscuits are crisp' makes a claim about the word which is true (or, equivalently express a proposition that is true) just if the things 'biscuit' stands for have the property 'crisp' stands for.

So, it is supposed to be with thoughts. Completed cognitive science will identify physical properties of brain states, which will allow us to re-identify syntactic tokens of the same type. Let's suppose, then, that brain state 'a' represents a biscuit, and brain state 'b' represents the property of being crisp. The idea is that there is some purely syntactic operation in the brain, which combines these two representations in a way which is analogous to predication to form a subject-predicate 'sentence' in mentalese, as it is commonly called. We can look into someone's head and see that an instance of state 'a' is connected to an instance of state 'b' in a way which creates an instance of 'crispness' being predicted of the 'biscuit': b(a). We now have a representation which is not atomic, but rather structurally molecular, and which expresses a proposition because it has a truth condition. An important question, obviously, is how the

states in the brain that are atomic components- the 'words' of mentalese – get to stand for what they stand for; and the same question arises for the part of a natural language, of course. The account we have just re-hearse is about how, on the language of thought hypothesis, to derive the representational content of a molecule from the representational content of its atomic parts.

The second part of the story tells us what makes it true that one or another 'attitude' is being taken to the proposition. On the language of thought hypothesis, what attitude you take to a proposition depends on the causal functional role of the sentence token in the head that encodes it-b(a)in our example. If the encoding sentence token that expresses the proposition that biscuits are crisp is a connected up in your brain so as to make you behave as though biscuits are crisp, then you believe that they are crisp. If on the other hand, the token is concerned up so as to make you try to bring it about that they are crisp (perhaps by putting them in the oven), you desire biscuits to be crisp. Often this is encapsulated in the metaphor of the belief box and desire box.

Thus, we have a two-part theory. What proposition a token of mentalese expresses depends on the semantic or representational properties of the syntactic or structural constituents and on the way, they are put together. The part two of the language of thought hypothesis requires that one finds out whether this token is located in the belief box or the desire box. If it plays the role that means that it belongs in the belief box, then the subject believes that chocolate is near, whereas if its role places it in the desire box, then she wants chocolate to be near. This theory is importantly different from Common-sense Functionalism. According to functionalism, the role-played determines the propositional object. But according to the language of thought account, functional role primarily settles whether the state is a belief or a desire, not what its propositional object is. Nevertheless, a state's propositional object is supposed to play a role is explaining behaviour. If the syntactic token of the proposition that chocolate is merely is in, say, the desire box, the very physical properties which allow us to identify it as a token which represents that proposition will, together with the fact of being in the desire box, serve to explain the chocolate-seeking behaviour of the agent. The desire for chocolate causes chocolate-seeking behaviour. For Fodor, the state causes chocolate-seeking behaviour because it is a desire for chocolate. The contrast is sometimes expressed in terms of a contrast between inside-out and outside-in stories. The language of thought theorist advances inside-out story that identifies the propositional object independently of the functional role, and sees the sentence of mentalese that encodes the proposition as in part determining the functional role. Functionalists by contrast, are an outside-in story that sees the functional role as determining the propositional object.

One very important question can be raised here that, why are we supposed to believe in the language of thought? The answer on offer is that only this hypothesis can explain what needs to be explained. A number of features of thought are identified, and it is argued that the language of thought hypothesis is the best explanation of them. The features that have been seen as particularly supportive of the hypothesis have varied somewhat over time, and from one presentation of the 'case for it' to another. We will focus on three features of thought: a) systematicity and productivity, b) the similarities in behaviorising from different thoughts, and c) the way thought evolves causally. Language is systematic. Both 'Jill loves Mery' and 'Mery loves Jill' are meaningful sentences of English, and in general, if 'aRb' is a meaningful sentence, then so is 'bRa. What is more, this is no accident. It falls out of the combinatorial structure of our language, from the fact that it is made of parts that we can rearrange to express different propositions which are systematically related one to another. So, with thought, if you can think that 'Jill loves Mary', then you can think that Mary loves Jill, and in general if you can think that aRb then you can think that bRa. Indeed, all the patterns of 'systematicity' that we see in language are mirrored in thought. On the other hand, if brain states have syntax, if mentalese exists, we have a relatively straight forward explanation of the 'systematicity' of thought. The fact that if you can think that aRb then you can think bRa, is explained by the fact that the state encodes the former is rearrangement of the parts of the state that encodes the letter. And, of course, this explanation generalizes to explain more complex cases.

A very similar argument can be made, appealing to 'productivity'. Language is productive. This ability is tied to the way the meaning of sentences depends on their structure to the compositionality of meaning, as it is often put. Thought is also productive. We can think quite new thoughts, and there is no upper limit to the new thoughts. Supporters of the language of thought hypothesis points out that if thought is like language in having a syntax, then we can explain the productivity of thought in the same way as we explain in productivity of language. Thirdly, it is good methodology to posit similar causes of similar effects, and their hypothesis does exactly that. For instance, the thought that coffee is there will share with the thought that coffee is here the bit that represents the coffee. Finally, thoughts 'evolve causally' over time. One thought leads to another by causing it. Moreover, this causal evolution overtime is driven by the propositional objects of the thoughts in question. It is because one thought is a that 'if there is snow outside it is cold outside' thought and the other is that 'there is snow outside' thought that the new thought is that 'it is cold outside' thought. Supporters of the language of thought hypothesis are fond of pointing out, reasonably enough, that their picture of the mind and its workings parallels the picture that applies to computers. It cannot be accused of being an untested bit of armchair theorizing.

In this story mental maps play the role given to mental sentences in the language of thought hypothesis. All theorists who allow those thoughts are states in head that represents reality as being one or another way, that encodes propositions, that have content. According to the supporters of the language of thought hypothesis: productivity and systematicity are both aspects of this open-endedness. The language of thought hypothesis is thus more than undeniable claims, that representation in the brain is structured. This means that, if we are right, then it is unfortunate that its supporters often refer to their theory simply as the representational theory of mind. A theory can be representational without being a language of thought theory. The way maps, pictures, diagrams, graphs and the like represent differs in two important, connected respects from the way sentences represent. First, map gives some (putative) information by giving a lot of information. We can now state an alternative to the language of thought hypothesis. It is that intentional states are maps in the head, rather than sentences in the head. The suggestion is not, of course, that something that looks like a map is in the head, but rather, that the way 'head states' represent is like the way 'maps' represent, in being essentially rich and in lacking representational joints. This is important in connection with the suggestion commonly made by supporters of the language of thought hypothesis that maps are good for representing spatial matters- where rivers and mountains are, and the like-but are no good for representing the movement of the share index.

The issue between the language of thought hypothesis and its rival, the map theory or the Connectionist model is a complex one that will not be settled here. Map theory explains the phenomena cited by supporters of the language of thought. This means that phenomena do not favour the language of thought over the map theory of internal representation. Hence map theorists can tell an essentially similar story to language of thought theories about how thoughts evolve over time as a function of their propositional objects. The difference is that in one case the propositional object is encoded in a sentence-like structure and in the other it is encoded in a map-like structure. In this context supporters of the language of thought sometimes stress the point we noted before that computer works on sentence-like internal representation. Supporters of the language of thought sometimes object that the map alternative we have been describing is more than a hope of an alternative than a genuine competitor. This can be said that map theory says nothing concrete about how the brain works. By contrast, the language of thought hypothesis says literally that our brains work in broadly the way computers work -namely, by manipulating re-identifiable symbols in a rule-governed way. This means that the language of thought provides an elegant structure explanation of how the brain works that explains systematic and so on. However, the map-theory or Connectionism, which is an alternative approach to computation, requires a detailed analysis. This we hope to take up after considering some objections to the language of thought hypothesis.

Objections

There are many objections to the language of thought. First, Fodor was too strict in what would count as successful accommodation of intentionally within a physicalist world picture. Further the objections came from those who do not share Fodor's optimism that our intentional kinds and generalizations can be vindicated by means of reduction into states and patterns of interaction characterized without the use of intentional vocabulary. The problem, with the computational account of rationality does not rest solely in the problem of capturing rationalizing links in purely formal manner. The rejection of such formality is still compatible with seeing rational links as essentially calculations, resting on objective patterns of relations, although these require intentional vocabulary and semantic as well as syntactic content for their articulation. Fodor's computational theory of mind enables us to see how minds might fit into the material world and how minds are related to brains. Minds are not identifiable with nor reducible to brains. Brains realize minds just as computing machines realize particular program.

In proposing a computational representational account of cognitive states and processes, Fodor's primary objective is to establish the autonomy of psychology as a special science. Psychology, for him, is an extension of our folk psychology. To be precise, the computational representational theory that is committed to language of thought hypothesis is an extension of our folk psychology. The first set of problems we will consider for the computational model are internal to the functional materialist paradigm, that is to say, to theorists who perceive a need to provide a naturalizing account of rationality and intentionality. Our line of argument suggests that Fodor has been too strict in what would count as a successful accommodation of intentionality within and physicalist world picture. Fodor's response is to claim that this fails to treat our intentional descriptions in a sufficiently realist way. Without attention to what is going on inside the black box, we can provide no account of the casual efficacy of our intentional states an efficacy that is assumed in our highly successful everyday interactions. If our intentional states are to be genuinely causally effective, then they must demonstrably be anchored in fundamental causal interactions at the physical level.

However, many see the kind of anchorage demanded by the language of thought hypothesis as too strong. It is suggested that some form of intentional realism is defensible without the systematic mapping which a language of thought postulates between intentional interconnection and material ones. Here the main challenge is mounted by those who adopt some form of connectionism as more plausible account of the kind of neural structures that would be needed to form a supervenient base for our intentional state. Fodor's reply to such a proposal is that cognition is productive and systematic that is, we can understand new thoughts and thoughts bear rationalizing links to each other. We argue that their features cannot be accommodated unless the states that instantiate cognition are susceptible to systematic characterization of the kind the language of thought suggests. If they are not so characterizable then it is unintelligible how our intentional representation supervenes on them. If they are so characterizable then connectionist architectures are simply a way of implementing mentalese style computational mechanisms. What then is at issue between two accounts? The issue becomes that of whether the intelligibility of the supervenience claim requires some intermediate level of syntactic description if rationalizing links are to be intelligibly generated from causal ones. When considering this question, it is important to keep in mind the difference between two projects. One is a purely empirical one considering what kind of physical systems can sustain intentional characterization. A second project is more philosophical, although anchored in empirical work. Its goal is to render intelligible how intentional characteristics can supervene on such a physical system. Fodor's account seems to be anchored in this second kind of project, arguing that an intermediate level of syntactic description is needed which can plausibly reduce rationalizing links and itself be instantiated by a physical system. Here Fodor's strategy seems to be more in line with the functionalist project overall, namely to find some intermediate level of description that can plausibly reduce intentional kinds and that can transparently be instantiated by a physical system.

Fodor starts with on account of how to reduce rationalizing links and accommodates these via considerations of syntax. Nonetheless intentional states have a semantic as well as syntax: they are meaningful representations. How are Fodor's computational states to be given semantics? How can the explanatory role, which seems to attach to the semantic content of these states, be accommodated within a picture in which the causal transitions are dedicated by the formal properties of the instantiating state? This is the problem of providing a naturalizing account of what determines the semantic content of our intentional states. At this point it is enough to point out that Fodor's framework sets constrain on what this kind of account can be like. Within the framework of the language of thought, the intentional content of our psychological states has been given a causal role in virtue of its formal properties. Therefore, any aspect of semantic content that we want to claim as causally relevant must be fixed by, or supervene on, such formal, syntactic properties. Furthermore, since such syntactic properties are properties of internal states, brain states, in Fodor's view, then the causally relevant aspects of semantic properties must be dependent on the structure of such internal states. These constrains on the allocation of semantic content, Fodor calls, methodological solipsism. If we cannot find an account of the semantic content of our intentional states that is compatible with Fodor's internal computational model, then it will not provide us with an adequate reduction of intentionality and rationality.

Further objections to the computational model come from those who do not share Fodor's optimism that our intentional kinds and generalizations can be vindicated by means of a reduction into states patterns of interaction characterized without the use of intentional vocabulary. One source of such anti-reductionist arguments is the work of Donals Davidson. Davidson's arguments rest on emphasizing the constitutive role of rationality in the assignment of intentional descriptions. He puts the matter like this: Beliefs, intentions and desires are identified by their objects and these are identified by their logical and semantic properties. If attitudes can be identified at all, then they must be found to be largely consistent with one another (because of their logical properties) and in line with the world (because of their semantic properties) ---------if a creature has propositional attitudes, then that creature is approximately rational. Therefore, even at the level of spelling out the network of conditional dependencies within which our intentional states are placed, we need to use intentional and semantic vocabulary. Beliefs, for example, tend to produce other beliefs for which they provide reasons. They are not causally stable in the presence of environmental conditions, which counts as evidence against them. In combination with desire, they tend to lead to actions that are ways of satisfying those desires. Davidson claims that the reason giving links that serve to define our intentional kinds "have no echo in physical theory". Without the use of intentional vocabulary such links could not captured. The consequence of this is to render it highly implausible that any classification of extensional terms could be projectably co-extensive with, and thus mirror, a classification anchored in a network of rationalizing link. This argument is used against Fodor's theory to suggest that, no set of conditional dependencies, characterized without the use of intentional vocabulary can map those which employ such vocabulary.

The Computationalists will reply that it is just this argument that his theory is designed to undermine. Once these rationalizing links have been formalized, via logic for theoretical reasoning and by decision theory for practical reason, then it is possible to find patterns of purely syntactical transitions that are projectably isomorphic with them. This makes it clear that a central issue in assessing the Davidsonian argument is whether relations such as "provide evidence for", "gives you some point in pursuing", and so on can be captured in ways that abstract from the semantic content of the intentional states involved and utilized only formal features. This seems dubious. Even where rationalizing links rest on relations of deductive and inductive validity, it is unclear that these can be captured in a purely formal way. John McDowell makes the point about deductive rationality, which extends even more powerfully to cases of inductive and practical reasoning. The problems with the computational account of rationality do not rest solely on the problem of capturing rationalizing links in a purely formal manner. The rejection of such formality is still compatible with seeing rational links as essentially calculations, resting on objective patterns of relations, although these require intentional vocabulary and 'semantic' and as well 'syntactic' content for their articulation. To see rationalizing links solely as instances of such general calculations

are, however, to miss something out of the picture. We understand someone's beliefs, responses, actions, by seeing them as appropriate, given their other intentional states. Such appropriateness can sometimes be demonstrated by processes of calculation and sometimes not. Nevertheless, even where such calculations are involved, the understanding that results are consequent on seeing that, given these logical connections, this is the thing to believe in this context. The force of inductive/deductive reasoning is not the mere instantiation of a formal pattern that happens to move us just because we are the kind of creatures who respond causally to these patterns. It is more, because we recognize that these patterns are truth preserving; we acknowledge the appropriateness of the conclusions given by the premises.

References:

1. Block N., *Readings in Philosophy of Psychology*, C.U.P.1980.
2. *Encyclopedia of Philosophy*, Routledge & Kegan Paul, London & New York, 1st Published 1998, by Routledge.
3. Fodor J., *The Language of Thought*, Harvard University Press 1975.
4. Fodor J., *Psychosemantics*, M.I.T. Press, Cambridge,1987.