# LANGUAGE OF THOUGHT

Christoph Schulze

## Introduction

The Language of Thought Hypothesis (LOTH) is a concept in cognitive science which describes mental activity in the brain as a form of language. The hypothesis was developed by Jerry Fodor in his book [1]. It states that the mind works with a language that is similar to regular languages, where an array of "words" together with syntactic and semantic rules make up the meaning of sentences and that these constructs are processed much like in a computer[2]. The origins for this hypothesis are older than Fodor's book however. Leibnitz already postulated the existence of an interpretable language of thought, which he called *lingua mentis* [3].

Another interesting aspect about thought is if our natural languages affect the way we think. By understanding if and how we can affect the way the mind works we might be able to derive the structure of how it operates. Lera Boroditsky [4]experimented with English and Chinese speakers to determine how their mind processes time. This was motivated by the fact that English and Mandarin differ in the way that time is represented in each language (Predominantly horizontal in English and vertical in Mandarin). Her experiments showed that to some extend our natural language can influence the way we think.

The remainder of this report goes into more detail about the LOTH and looks at some of the arguments for and against it. The last section will elaborate on Lera Boroditsky's work how languages affect the way we think.

## The Language of Thought Hypothesis

The language of thought hypothesis is part of a larger canon of theses(See [2] for the full list) and can be summarized as.

- Mental representations are structured
- Parts of these structures are transferable, that means they can appear in different representations
- Mental representations have a compositional semantic. Which means that the meaning of complex representations can be extracted out of its parts.

The language of thought, which is sometimes called "Mentalese" is thought to have a structure much like regular languages. A sentence is composed out of several parts which can have a meaning by themselves but which form a greater meaning when combined. The different parts of one sentence can appear in another one forming completely different meanings. The language of thought hypothesis states that the human brain has a language that works very much like such a regular language.

The goal of the hypothesis is to explain how things like wishes, beliefs, and convictions are realized physically. To this end the Fodor postulated a computational model of the human mind that works very much like a computer. Thoughts are believed to be a series of operations that make use of syntactical aspects of the language during their processing.

The hypothesis does not apply to the whole mind it only claims to describe thought processes that consist out of propositional attitudes like, S believes P or S hopes that P. S in this context is the subject and P the

proposition of the attitude. It does not cover processes like experience, sensory processes, dreaming etc. However, some researchers think that there might be more than one LOT and that those different LOTs would be responsible for handling other thought processes not covered in the LOTH [5].

There is one major difference to regular languages however, and that is the way the language is realized. Normal languages use optic or acoustic mechanisms. The language of thought is only realized in neuronal patterns very much like memory patterns in a computer which can thus far not be made accessible properly.[6]

## Arguments for the Language of Thought Hypothesis

For the first argument Fodor analyzed the existing models for the representation of higher cognition and concluded that the only remotely feasible models so far are all computational and that therefore computational models must at least be treated as valid until a better model can be created.

The Productivity Argument claims that since there can be an infinite number of mental states and humans have only a finite amount of "memory" there must be a way to produce these states from a set of atomic elements. Since the only way to achieve this, according to Fodor, is to use a combinatorial syntax and compositional semantics the human mind must have computational processes based on these concepts to form these states.

The Systematicity Argument goes along similar lines. It argues that propositional attitudes are systematic. This can best be explained with an example:

- "Juliette loves Romeo"
- "Romeo loves Juliette"

The two sentences contain the same components and the same structure the difference in meaning is only determined by the way the components are arranged. Since this, again according to Fodor, also requires a combinatorial syntax and compositional semantic it excludes other possibilities besides the LOTH.

## Arguments against the Language of Thought Hypothesis

Daniell Dennett argues that there can be propositional attitudes without a corresponding explicit mental state for this proposition. As an example he mentions a chess program and the propositional attitude "I think I should get my queen out early" which describes a strategy in chess. He claims that even so chess programs have explicit representations of a lot of aspects of the game they do not have representations for the above mentioned propositional attitudes. Based on the chess example Dennet tries to show that in sufficiently complex systems there can be cases where the illusion of an explicit mental state for propositional attitudes like the one mentioned above appear without actually being in the system. Along the same lines he also argues that there are cases in computers for example of explicit state representations that don't belong to any propositional attitudes. As an example he mentions low level machine operations in computers.

Proponents of Eliminative Materialism like Churchland [7] argue that there are no discrete mental states in Neuroscience like the ones described in the LOTH. Instead they argue that the neurons with their action potentials are continuous and therefore operations in the brain cannot be realized as discrete mental states.

Connectivists proposed an alternative hypothesis in which the mind of a person is made up of a neural network and not of a language processing system. This was a challenge because Fodor's first argument stated that there are no other models that can explain cognitive functions like LOT. Some Connectivists and proponents of LOT eventually combined the two arguments and showed that it was possible that the LOT could be implemented on a neural network.

## How languages affect thought

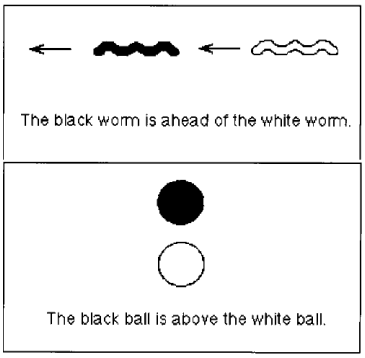


**Figure 1: Primer Questions from [4]**

Lera Boroditsky [4] ran several experiments to determine if human language can affect the way we think. The question that she tried to answer was "Does the language you speak affect how you think about the world". In her experiments she used native English speakers and native Mandarin speakers since both languages have a different way of how to describe time. In English time is predominantly described as if it were horizontal (E.g. "good times *ahead* of us", "hardships *behind* us"). In contrast to that Mandarin predominantly describes time in a vertical fashion (E.g. using "*up*" and "*down*" to talk about the previous or next month). In the experiment the participants would first see a primer question (See Figure 1) and then they would have to answer a question about time (E.g. "March comes before April"). The results showed that the English speakers could answer questions faster if they got a horizontal primer and that this was reversed for Mandarin speakers.

She concluded that languages can affect the way we think but has to restrict this to abstract domains like time because of work done by other researches which did similar experiments that indicate that some aspects, like the perceptions of color, are already fixed before learning languages and that therefore languages most likely can only influence domains that are not so reliant on sensory experience.

## References

[1] J. A. Fodor, *The Language of Thought*. Harvard University Press, 1975, p. 214.

[2] M. Aydede, "The Language of Thought Hypothesis."

[3] M. Kulstad and L. Carlin, "Leibniz's Philosophy of Mind."

[4] L. Boroditsky, "Does language shape thought? Mandarin and English speakers' conceptions of time.," *Cognitive psychology*, vol. 43, no. 1, pp. 1–22, Aug. 2001.

[5] J. Fodor, *The Modularity of Mind*. The MIT Press, 1983.

[6] A. Beckermann, *Can there be a language of thought?* Berlin/New York: Walter de Gruyter, 1997, p. 384.

[7] P. S. Churchland, "Neurophilosophy: Toward a Unified Science of the Mind-Brain," *Philosophical Books*, vol. 28, no. 4, pp. 233–236, Oct. 1987.