

Circling the Void: Using Heidegger and Lacan to think about Large Language Models

Marc Heimann (✉ marc.heimann@hs-niederrhein.de)

Hochschule Niederrhein

Anne-Friederike Hübener

Hochschule Niederrhein

Research Article

Keywords: AI, LLM, continental philosophy, psychoanalytic language theory, Heidegger, Lacan

Posted Date: June 16th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3023378/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The essay aims to unite two currently distinct lines of thinking and working with language. Large Language Models and continental philosophy, especially Martin Heidegger's thinking about language and, building upon Sigmund Freud, Jaques Lacan's structural psychoanalysis. We show that the concept of language that Heidegger, Freud and Lacan discuss and utilize in clinical frameworks is matched quite strongly by modern LLMs. This allows us to discuss a problem of negation and negativity that is central to the continental discourse but missing in current LLM research. This also means that we offer a radically different approach than it is usual in the philosophy of artificial intelligence, since we base our concepts on thinkers that are often disregarded in the analytic philosophy discourse that is closer linked to AI research. To this end we also mark, where the ontological differences of the proposed approach lie. However, our aim is to address AI researcher and continental philosophers.

Introduction

Noam Chomsky et al. recently ended an opinion piece in the New York Times with a notable sentence about modern Large Language Models (LLM): "Given the amorality, faux science and linguistic incompetence of these systems, we can only laugh or cry at their popularity." (Chomsky et al., 2023) This assertion may appear surprising, especially as these systems currently demonstrate an astounding capacity at understanding human language, even quite complicated scientific texts. In a way Chomsky is right: ChatGPT as well as other LLMs do have (openly acknowledged) difficulties. However, the question stands, if Chomsky pinpoints the problem at hand, or if he obscures it. While Chomsky advocates for a rule-based system, others like Frederick Jelinek favored an entirely different approach. In the context of AI's connection to linguistics, Jelinek's well-known quote, "Every time I fire a linguist, the performance of the speech recognizer goes up," is particularly relevant, targeting linguists who have been educated in Chomsky's theories. Jelinek was a pioneer in the field of statistical language processing and helped shift the focus from rule-based systems, like those rooted in Chomsky's theories, to data-driven approaches. Modern LLMs, such as GPT-4, which rely on deep learning and neural networks to process and generate human-like text, undoubtedly owe much of their development to Jelinek's emphasis on data-driven methods and statistical approaches, even if they are not directly based on his work. Which shows that Chomsky's critique is rooted deeply in the way modern AI research links up with linguistics.

What, then, is the argument that Chomsky et al. make regarding LLMs? They argue in the opinion piece that true human capability is marked by operating on a rule based system and that "marinating in Big Data" (Chomsky et al., 2023) isn't enough. Its notable however, that in terms of everyday use of language, being marinated in big data is on a quite astounding level enough to work with it. To think like humans as Chomsky et al. argue AIs require to be "endowed with" the same "universal grammar" (2023) that humans use. Interestingly, the data-driven methodologies Jelinek helped pioneer might find a philosophical ally in the language theories proposed by continental philosophy.

We contest this perspective and will aim to demonstrate that instead LLM should instead move further away from rule-based systems and might find support for this in language theories that have been proposed by continental philosophy. What we will try to mark in this paper is that LLMs are first of all better approached by the theories of language that developed in continental philosophy, especially by Heidegger and Lacan and that a problem called the “impasse of formalization” also appears in LLMs, however, since this problem is not considered in itself, it weakens the way the models represent language. This paper introduces a novel perspective, bridging computer science and continental psychoanalytic philosophy. As such, readers may find themselves navigating unfamiliar territory in one or both of these fields.

Given, that speech recognition only seems to have taken off after relinquishing these ideas of rule-based approaches, what if instead the move away from rule-based systems has not been radical enough? A statistical model, even one based on fractal chances and not just bell curve variations, still assumes an ontological basis that has been under attack from different disciplines for more than a hundred years: the idea of an ontologically fixed nature: That the cartesian ‘clara et distincta’ is not just a formal requirement of our statements, but more deeply ingrained into reality itself. The genesis of this idea is of course a theological one, the guarantor of the stability of nature’s clara et distincta was the Christian god. Now, even modern statistics requires that the “roll of the dice” that it was conceptualized on, is thought to be structured by fixed ontological laws. This has been recently discussed by Quentin Meillassoux as a problem (Meillassoux, 2008, pp. 96–98), because it disallows a statistical realism (the now dominant ontology of many sciences) the access to the problem of indeterminacy. However, this is not a criticism of the statistical basis of LLMs, as we will discuss later, but a problem of the ontological foundations that Chomsky assumes.

This problem of indeterminacy is first of all one rooted philosophy and founds its way into formal logic and mathematics. Originally marked by the loss of the absolute foundations of thought, which is explored most prominently in Nietzsche’s “death of god”, has also been explored by mathematicians and logicians like Kurt Gödel, with his incompleteness theorems, which demonstrated that any axiomatic system capable of expressing an arithmetic would necessarily be incomplete and could not prove its own consistency. Gödel's theorems revealed fundamental limitations in formal systems, which can be seen as a form of indeterminacy in the sense that there will always be true statements within a formal system that cannot be proven within the system itself. While Gödel's theorem is not an ontological statement and should not be taken as one without further considerations, the consequences for ontology should be taken seriously. In current philosophy it has informed Alain Badiou (2006) as well as Quentin Meillassoux (2008), both indicated that we need to think our formal systems of thought to be interlinked with the “impasse of formalization” (Badiou, 2006, p. 5). This impasse as Badiou identifies it, is at the same time exterior to our formal systems as well as marked by an infinite excess which they by necessity produce. It is a central term that we will discuss in detail as it implies a rich field of research for AI research, if it is informed by psychoanalytic philosophy.

This mirrors a formal problem that Kant consciously marked in his logical concept of an infinite judgement (Kant, 1967, B97 A72). Infinite judgment is a specific type of judgment that, according to Kant, involves a negative judgment about a concept that has an infinite extension. In other words, it negates a concept that covers an unlimited scope. Infinite judgments are contrasted with affirmative judgments (which assert the connection between a subject and a predicate) and negative judgments (which deny the connection between a subject and a predicate). An example of an infinite judgment is the statement "the soul is not mortal." Here, the concept of "the soul" is negated by the concept of "not mortal," which implies that the soul has an infinite extension or is immortal. This formal concept is later explored by Martin Heidegger (1976) in "What is Metaphysics", where he indicates that every universal judgement creates an indeterminate negative excess. Marking that this excess far from purely negative does have an impact on our lives. However, it takes the psychiatry and the development of psychoanalysis by Sigmund Freud and later Jaques Lacan to develop a true logical examination of these problems of the impasse of formalization.

Words and Large Language Models

To understand why a Lacanian or broader approach rooted in continental philosophy is helpful when thinking about current AI models, let us first take a concrete look at current AI: large language models. I won't go too deeply into the technical details, as there are countless other papers on this model too deepen ones understanding. However, let us take a simplified look at how a prompt is worked on by the general technical structure of a state-of-the-art language model. Let us consider a scenario where we want to generate a continuation for the following prompt: "He jumped the...". We will use a simplified version of the GPT-3 architecture for this example. This unfinished sentence is then worked on by several interlinked mechanism within the model.

The input sequence "He jumped the..." is first passed through an input embedding layer, which converts the sequence into a set of continuous vector representations (Mikolov et al., 2013). Each word in the sequence is broken down into tokens, which can be whole words or smaller units depending on the language and specific word. LLMs break down input sequences into tokens and convert them into continuous vector representations that capture the meaning and context of each token within the sequence. This process is performed by the input embedding layer. Each word in the prompt is also positionally encoded, which is crucial for providing the model with information about the position of words in a sequence. However, these vectors are now only initialized. After this initial input embedding, the Self-Attention layers then allow the LLM to read this sentence in several directions and/or with an emphasis on different words (compare Vaswani et al., 2017).

The vector representations of tokens are passed through multiple layers of self-attention mechanism in the model. In each layer, the model calculates attention scores for each token based on its relationships with all other tokens in the sequence. The attention scores are used to weight the vector representations, allowing the model to focus more on the relevant tokens for each position in the sequence. This mechanism allows the LLM to understand a token not just in a rule-based approach, but also helps to

represent different levels of metaphoric and metonymic linkages between tokens in the input. This is an important step as it removes the structure of a sentence or prompt away from models that simply process it in a linear way and allow a diverse and complex interpretation of the words, not unlike one would do in a classical close-reading interpretation (some major caveats here, which will be discussed later on, but the broad gist is correct insofar we assume that a close-reading will try to capture as much of the inherent formal relation within a given sentence as possible).

These weighted vector representations are then passed through multiple feedforward layers. Each feedforward layer consists of a dense layer followed by an activation function, which applies a non-linear transformation to the vector representations. The feedforward layers in a large language model (LLM) like GPT-3 are an essential component of the model, responsible for transforming the input sequence into a more expressive and high-dimensional, vector-based representation of linguistic relationships that can capture complex relationships between words (compare Brown et al., 2020).

In simple terms, the feedforward layers in an LLM work by taking the vector representation of each word in the input sequence and transforming it into a new vector representation that captures higher-level features and relationships between words in general. What happens here in layman's term is that the representation of the prompt is interlinked with the LLMs representation of language. Each of these layers does two things to the original prompt. The input vector representation of each word in the sequence is first transformed linearly by multiplying it with a weight matrix and adding a bias vector. This is a crucial step, and its actions are informed by the machine learning process that allows the model to affix these weightings through a weight matrix (compare Rumelhart et al., 1986). A weight matrix is a matrix of parameters that are learned during the training process and used to transform the input vector representation of each word in the sequence into a new, higher-dimensional vector representation. The weight matrix is learned during the training process and is updated iteratively to minimize the difference between the predicted output of the model and the true output.

The resulting output of the feedforward layer is a new, higher-dimensional vector representation of the word that captures higher-level features and relationships between words. Take as an example of a higher-dimensional vector representation the word "king". the vector representation of "king" is a multi-dimensional vector, where each dimension captures a different feature or relationship between the word and other words in the language. Some of these dimensions might capture syntactic relationships, such as the subject-verb-object relationship between "king", "rules", and "kingdom". However, the model as a whole also learns to represent these kinds of syntactic relationships between words in a sequence in more general terms. Other dimensions might capture semantic relationships, such as the fact that "king" is related to other royalty-related words like "queen", "prince", and "duke". Other dimensions might capture contextual relationships, such as the fact that "king" is often used in the context of historical texts or fantasy stories.

During training, both the weight matrix and bias vector are updated to optimize the objective function of the model, typically through a process called backpropagation. An objective function, also known as a

loss function or cost function, is a mathematical function used to quantify the difference between the predicted output of a model and the true output. This involves computing the gradient of the objective function with respect to the weight matrix and using this gradient to update the values of the weight matrix to minimize the difference represented by the objective function. Then, the output of the linear transformation is then passed through a non-linear activation function, which introduces non-linearity into the model and allows it to capture more complex relationships between words. In the context of neural networks, non-linearity is important because it allows the model to capture more complex relationships between the input and output. If a model were only able to represent linear relationships, it would not be able to capture many of the complex patterns and relationships that are present in real-world data. For example, in image recognition tasks, a linear model might struggle to recognize complex shapes and patterns, but it can still capture more than just simple shapes like straight lines or circles when used correctly.

The final output of the model then is generated by passing the output of the feedforward layers through an output layer, which predicts the probability distribution over the next word in the sequence. The output layer is a softmax function, which is a mathematical function that is used to convert a vector of real numbers into a probability distribution (Bishop, 2006, p. 198). The softmax function is often used in classification tasks, where the goal is to predict the probability that an input belongs to each of several possible classes. The model then samples from this distribution to generate the next word in the sequence. In our example, the LLM would generate a continuation for the prompt by passing the initial sequence through the input embedding, self-attention, and feedforward layers, and then using the output layer to predict the next word in the sequence. The generated continuation would then be added to the prompt, and the process would be repeated until the desired length of the sequence is reached. (For further reading regarding the structure and basis of LLMs in the GPT line: Alec Radford et al., 2019; Brown et al., 2020).

So, to reformulate the example in terms much closer to readers without any technical background, what we get with a LLM is a highly complex model of language, that is trained on vast amounts of texts, by learning which of these words usually (that is statistically) are located in relation to each other. It represents these relations by turning each word into a mathematical representation which it then analyses in relation to the relationships that are extracted from its training data. A common argument aimed at this process of creating a model of language is that this is purely statistical, however this is not entirely correct. The product of statistical analysis are the tokens with their inherent weights and links to other tokens. While this may have been created on the basis of statistical inference the structure created is not a statistical one. Instead, it can be seen as a kind of learned, high-dimensional, vector-based representation of linguistic relationships. This representation, although inherently mathematical, attempts to capture complex semantic and syntactic relationships between words, phrases, and broader linguistic structures.

There are several limitations of this approach, for example the production of non-sense texts, which are generated not only because humans produce vast amounts of nonsense themselves, but also because of

the statistical approach by necessity producing outliers. The mathematical core of this language doesn't 'understand' anything, it simply relies on statistically detectable links. This also makes it very sensitive to the phrasing of the input and tend to produce verbose, overly confident or altogether too cautious outputs. Which follows from the way it interacts and puts emphasis on these inputs. Prominently leading to the idea of "prompt engineering" as a way of addressing these limitations or taking them as a feature. However, assuming prompting to be a lack rather than a central element of how the attention mechanism interacts with the data stored would be a mistake, in layman terms the prompt constitutes the perspective of the generated output.

Another central element is being not focused on full sentences which distances it from classical Aristotelian (rule-based) approaches to logic and meaning, but instead focusing on tokens, which are words, sub-words (like "meta-"), punctuation and special characters. From these nexūs it reconstructs language to enable the Model to reconstruct the complex metonymic and metaphoric linkages. This of course distances it from several prominent approaches to the inherent rule-based logic of language, notably, the grammar-based approach that Chomsky advocates. However, it also means that the continental Freudian/Heideggerian tradition with its heavy emphasis on the importance of the complex formal relationships of words as signifiers for other signifiers instead of sentences is much closer to the theory of language that LLMs approach. All of this might seem to be an overly technical introduction to LLMs, however, in stark difference to Chomsky, we consider LLMs to be quite close to human language use and to demarcate this proximity we will use an example of language use that Star Trek: The Next Generation allows us to mark in detail, which allows us to link it to the word and metaphor centric understanding of language that we find in continental philosophy.

Picard and Undivided Multitudes

Now, to explore the close proximity of the inner workings of LLMs and continental theories of language, we should take a look at its metaphorical capabilities. To do this with an accessible example we can utilize an example of translation taken from the TV series Star Trek: The Next Generation. In the "Darmok" episode of Star Trek: The Next Generation, the crew of the USS Enterprise encounters the Tamarians. Despite the Tamarians' attempts to establish contact, their language and communication methods are at first incomprehensible to the Enterprise crew. As a representative of Chomsky's understanding of language, the ship's android Data remarks early in the episode that the Tamarians seem to have an incomprehensible language. However, through the cooperation the two groups gradually achieve mutual understanding by exchanging cultural stories and metaphors. This use of language as such is not what we're after here. Instead, let us look at how GPT-4, as an example of an LLM, adapts this style of communication to explore the structure of its theory of language (something we are not the first to think of, compare Jansen & Boyd-Graber, 2022). We have used the following prompt to turn GPT-4 into a Tamarian:

Answer in the metaphorical style that is used by the Tamarian captain, Dathon in the Star Trek: Next Generation episode Darmok, but utilize European philosophy and science allegorically, instead of

Tamarian culture: Can you give me a short description of the attention mechanism in Large Language Models.

Now, the reasoning of this prompt is as follows: Given the structure of LLMs we know that its inner mapping of references is not in itself organized to answer this question. Neither the “Darmok” Episode, nor the question itself should be strongly linked to an allegorical use of philosophy and science to communicate. The attention mechanism acts here as the metaphorical and limiting process itself, by linking complex relational systems of different language games together producing an intriguing rebus:

“Einstein, his mind bent spacetime; Archimedes, in the tub, Eureka! The many, a cacophony, seeking the sun. Galileo, his telescope to the heavens, the planets align. Descartes, cogito ergo sum, the essence distilled. Attention, the spotlight of consciousness, cast upon the words; the dance of meaning, woven in the tapestry of language.” (GPT-4)

While Chomsky's theories provide a perspective on language that has difficulties with this approach, other philosophers like Freud, Heidegger, and Lacan offer viewpoints that are particularly relevant for understanding LLMs' metaphorical language processing capabilities. This structure of a rebus as an explication of the inner structure of language and memory is not new, we explicitly find it in Freud's *Traumdeutung* as the primary structure of the dream, where he explores the formal structures of the unconscious (Freud, 1942, p. 284). And it shows us that to understand language, we should look at the rebus not as something derived from sound sentences, but instead as the foundation of these sentences as the structure of LLMs like GPT-3 imply.

While Freud never developed a formal theory of language, the methodological basis of psychoanalysis is built upon a parallel understanding of language, as the clinical method of free association explicitly rejects the idea of a correct associative link and encourages patients to approach this manifold of associative links of their own intentionality of perspective to reconstruct the relation of that which is explicitly said to the unconscious. Especially as this pre-predicative form of signifying operations is done even unwittingly (Olyff & Bazan, 2022). This linkage is interesting to us, because it offers a strong parallel to LLMs. Because each single word or token takes its meaning not from the position in the sentence, but through utilizing the contextual elements that are constituted by the training data. What makes ChatGPT-4 outputs so lifelike is exactly this capability of metaphorically linking up different contexts of meaning, that is of accessing the foundational structure of language.

To understand how LLMs, like GPT-4, constitute a genuine understanding of language, it helps to consider Heidegger's perspective on the status of words in language. There is an important remark to be made here, especially to readers from the anglophone discourse on metaphors, which is dominated by Lakoff and Johnson's “Metaphors we live by”, but usually is unaware of the more complex theories of Blumenberg, Heidegger and Freud, which substantially preceded Lakoff and Johnson (Schröder, 2008). There are several differences to be aware here, but the most important is the ontological framework that we will discuss further down. Another important remarks, one closer at hand, is that the continental discourse essentially focused on metaphors very much in the same sense that tokens are used by LLMs.

That is not as complete words, but also as word fragments. This shifts the focus to include relatively complicated concepts like the "Gestell" of the later Heidegger, where he utilizes the German languages tendency to build up words from smaller subwords, like the prefix Ge-. That this is already conceptualized in Heidegger is not cleanly visible, as Heidegger is not centrally interested very much in this element of language, but we can find it clearly articulated. Take for example his discussion of the prefix "un-", when he discusses unconcealedness as the proper translation for aletheia: The prefix un- in "Un-concealedness' can mean concealedness is taken away, cancelled, evicted, or banned where taking away, cancelling, and banning are essentially distinct" (Heidegger, 1998, p. 14). This subword here is already structured through a multitude of links. Accordingly, a single word is already a linkage between different sub elements: "a word as a whole is drawn, not from a primary, primordial experience of the subject matter, but from preconceptions and the nearest at hand views of thing" (Heidegger, 2005, p. 12). Even in the case of singular letters Heidegger remarks that "'Something as something' [is] in the background!" (Heidegger, 2007, p. 113), which means that for Heidegger the structure of meaning as a nexus of relations and indicative vectors already exists at the level of the letter.

Let us take as another example the most important word of Heidegger's philosophy: being. In the "Basic Concepts of Metaphysics," Heidegger addresses the problem of the status of words through the concept of being in language, that is, the "is" in spoken language. Which means that he not only aims to conceptualize an ontological concept of being, but first and foremost offers us a detailed understanding how words are functioning in his theory of language. He first descriptively states that the use of "is" in a sentence can express a manifold of meanings, a "what-being, whether in the form of so-being and essence-being, the that-being and the true-being" in one and the same sentence. This ambiguity, which Heidegger initially works out through the historical comparison of the use of being concepts, is as he argues, however, a positive phenomenon of being as "is." Accordingly, he says:

"It is necessary to understand this peculiar indifference and universality of the 'is' as the original and primary essence of the copula, or that which is externally called the copula."

The starting point of understanding being within a sentence is therefore "the full undivided manifold" of this horizon of meaning of the "is." With regard to this horizon of meaning, Heidegger already emphasizes in "Being and Time" that this is initially only a wholeness of referents (Heidegger, 1967, p. 87), which by itself does not refer to being, but only enables the reference to being as such. Therefore, a meaning does not consist in a reference to being, *but in the relation to other meanings*, or in Lacanian terms signifying other signifiers. This meaning, as Heidegger already states in his early work, is originally not oriented to the sensible or objects, but to its "Wozu" or "Um-zu". This "wherefore" is originally practical, it is based in know-how, i.e., always already framed in a relational context of meanings, which has always already passed over every concrete given.

In view of the the transformer type LLM we see that the process that a LLM utilizes to create its model of natural language corresponds to this discourse on language being constructed from partial words insofar it also approaches single tokens from a comparable undivided manifold of links, which are as much

based in practice (that is the practice of writing present in the data it uses to train on) as Heidegger assumes that this manifold is based in. The attention mechanism then tries to grasp the intention of a prompt by comparing it again to the weights generated by the data in its training. However, this means that the attention mechanism acts here as a form of determination by negation which needs to be discerned further. However, while this informs Heidegger's use of combined words like the later "Ge-stell" or the "Ab-Grund" it takes Lacan's insistence on the letter as a central element in psychoanalysis to fully approach a formal theory of the chain of signifiers. Lacan build here on Freuds analysis of phonemes and puns which allows to reconstruct the additional linkage between words and subwords, which do not link by content, but by speech as praxis. In the opening lines of the play, Hamlet uses wordplay to express a paradoxical idea: Although Claudius is both his uncle and stepfather, he is not truly a relative or like-minded: "A little more than kin, and less than kind" (1.2.65). What is important here is that Lacan allows us to think of these phonemes as well as letters as *partial objects* (Jacques Lacan, 2006, p. 693). Partial objects are not objects that are structured by their inclusion into a whole as the name could imply, but instead only imply a whole encompassing them without actually being structured by this imaginary whole.

Attention, Negativity and Undivided Multitudes

How can this undivided manifoldness be grasped more precisely and especially in relation to the ontological perspective Heidegger assumes? Heidegger himself says that language as structured by the intentional, that is the speakers, approach to the undivided manifold precedes the essence of the λ $\gamma\omicron\varsigma$, where λ $\gamma\omicron\varsigma$ for Heidegger is always more than a just a predicative logic based on full sentences, but also includes such a logic. In comparison to Chomsky's grammar, he assumes therefore the multitudes of practical relations to be foundational for the structured systems of logic. The reason for this is to be found in the way Heidegger fundamentally thinks the concept of determination. Central for Heidegger is that he assumes that any logic that is consistent is related to negativity. Meaning is therefore not only intentional referral, but this intentional referral is linked to something that cannot be referenced itself and can only grasped positively in a failure of the referral structure. In "Being and Time" this failure of the referral structure is my own death (not that of others). However, Heidegger later develops a more sophisticated concept of this impasse of the formal structure of meaning in the strife (for further reading on the problem of strife see **anonymized reference**).

Now, what does this mean for our parallelization to LLMs? What the attention mechanism does is providing a foundation to limit the undivided multitude and turn it into a divided multitude, that is one where certain links are not weighted strong enough to appear within the output. For example, if I ask about Russel's "King of France" the "King's" inherent link to "Queen" that should be established in the data is unlikely to appear in the output. In the case of Russell's "King of France," the LMM would use its attention mechanism to identify the key entities mentioned in the input text (i.e., "King" and "France"), and then use that information to generate a response. It is not necessarily the case that the LM would completely disregard any link between the "King" and "Queen" simply based on the attention mechanism weighting, but it is unlikely that a question about Russel's "King of France" would reference a "Queen" if

this link does not appear in the prompt. The model would consider all associations between words in the input text and generate an output based on the most probable associations given the context, which would likely exclude the link between “King” and “Queen”.

Since the model does not retain this link but turns the prompt into an output it is essentially negated. Once the LMM generates its output, it does not retain any information about previous associations between words or entities. Therefore, if the LMM generates an output based on the input text "Russell's King of France," it would not include any link between "King" and "Queen" because that link was not explicitly mentioned in the input text. In this case, the attention mechanism is acting as a determination by negation, as its action is excluding the possibility of a link between "King" and "Queen" based on the absence of this link in the input text. This shows an aspect that we have introduced already: that the attention mechanism provides the *perspective* of the LLMs output. This perspective following the optical metaphor, excludes and highlights certain links that the training data contains.

A well-constructed input prompt that provides context and guidance for the model can help the attention mechanism selectively activate parts of the feedback layers that are most relevant to the specific output. The ease of which Chat-GPT can for example be ‘jailbroken’ to ignore its ethical limitations, attest to this mirror of perspective and intentionality as a central part of the way the attention mechanism interacts with the prompt. Because the multitude of words, or in the case of LLMs complex systems of tokens are essentially on their own an undivided manifold of links, this intentionality is absolutely necessary to create meaningful sentences as a limitation (or dividing) of this manifold. Intention as attention can therefore be understood as a form of limitation, and this should be kept in mind as it will be important for discerning the ontological basis of Heidegger’s thinking further.

To move deeper into the problems that the continental discourse allows us to mark, let us look on the ontological perspectives from which a token-and-negation-based theory follows and why. For Heidegger, Freud and the continental tradition build upon the basic insights of these thinkers, Chomsky’s grammar based approach is essentially a metaphysical solution to the problem of language as Heidegger considers the common understanding of logic and language to be limping behind actual language use and to be based in a vulgar ontology of presence (Heidegger, 1967, p. 10). The only way to formulate a rigorous philosophy is for an approach to logic following Heidegger then is to accept the abyss as the core of the ratio (Ab-Grund).

The Abyss and Aesthetics

In difference to Heidegger the understanding of language that Chomsky details in his seminal book “Syntactic Structures” (2002) shows that his idea of a good theory and research is structured by the essentially aesthetic idea of “the simplicity of the whole system” (Chomsky, 2002, p. 56). This principle of simplicity of a whole accordingly should organize the idea of finding a “metalanguage to the language in which grammars are written – a metametalanguage to any language for which a grammar is constructed” (Chomsky, 2002, p. 54). Obviously, the idea of simplicity acts here as a counterforce to the

looming thread of infinite regress that encourages to think of metametalanguages. However, it also shows that the implicit idea of determination prominent in Chomsky's thinking is one of a closed system, ruled by an internally absolute principle. This idea of a consistent system that is organized by a small batch of internal rules isn't surprising and within bounds, especially in view of machines it is the most viable way to approach. Chomsky accordingly pictures the goal of a theory of language in terms of a machine as something that must provide a "practical and mechanical method for actually constructing the grammar" (Chomsky, 2002, pp. 50–51). This machine as we can gather in the Chomsky et al. commentary is "the innate, genetically installed "operating system" that endows humans with the capacity to generate complex sentences and long trains of thought" (Chomsky et al., 2023). Now, for normal machines, like any computer one can pick up, this idea of a system works splendid, within some limitations however, as it turned out to be necessary in some instances to think of ways how computers can work around data and inputs which do not conform to this ideal, paraconsistent logic and fuzzy logic come to mind.

A fundamentally different perspective arises when one assumes that, since machines can be constructed in a particular way, spoken language (and, more broadly, reality) must also conform to this design. While it might gain some credibility by being deemed rigorous, it ultimately represents an insidious infiltration of theology into the realm of science by assuming an axiomatic primacy of the system. It is worth noting that in his 2006 preface to "In Contradiction," Graham Priest effectively illustrates the extent to which ideas surrounding the necessity of consistency and the system continue to permeate the analytic discourse on logic (Priest, 2006, p. XVIII). In contrast, the Continental discourse on logic is predicated upon the dismantling of this onto-theology. Prominent thinkers such as Martin Heidegger, Jacques Lacan, Alain Badiou, and more recently, Quentin Meillassoux, have critiqued the notion of a consistent reality that can be represented by a systematic, coherent, and ideally aesthetically simple theory. They argue, and we hold this position too, that this concept is nothing more than a truncated and filtered form of religious thought, or in Heidegger's terms, plain old metaphysics. It is important to recognize that while individuals may hold various religious beliefs about the world, these beliefs do *not* qualify as the basis of scientific methodologies, be they formal or empirical.

This fundamentally changes how language is understood, especially in a systematic context. To demarcate the problem at hand it is necessary to show where the continental discourse on language radically differs from the one in which Chomsky's ideas originate. In stark difference with the philosophy of Bertrand Russell (1905) as marked in his theory of descriptions, where the denotations indicating nothing are essentially false if taken as a primary occurrence, the approach that we can find in Heidegger, Lacan and the continental discourse operates as an inversion of this idea, as Jacques-Alain Miller (2002) marked. For Lacan, Heidegger and other central representatives of the continental discourse, only those denotations oriented on the radical indeterminate can be considered constituents of truth. While this might seem strange for those unaware of this discourse, it is not as distanced from classical analytic epistemics as it might seem. Karl Popper (1935) with his concept of falsifications as the only true access to the reality exterior to a theoretical system approaches a comparable thought. However, what considerable parts the continental discourse focused on, was the interlinked problem of formal reasoning and the "impasse of

formalization” (Badiou, 2006, p. 5). Hans Blumenberg (2010) demonstrated early how deeply this problem is ingrained in classical philosophy, where systems are essentially oriented around an “absolute metaphor” an empty denotation that holds up the theory, instead of harming it.

So how can we build meaning on this link? One solution is a certain pragmatism about this link coupled with an understanding of logic that eschews the problem further, if one assumes that logic and empirical reality are essentially separate, we don’t need to think about this as it makes no sense, and this pragmatic gap is where the old god of the philosophers under new names usually creeps back in. The early Wittgenstein’s mysticism, while seemingly moving into the direction of continental thinkers essentially relegates this problem to the theologians. However, if such strong limitations are themselves nothing but metaphors constructed to appear rigorous and strict where theory actually gets positively weird, we approach a different problem. Slavoj Žižek formulated this as our capacity of formal thought reaching through to the baselessness of reality (the void), beyond what the theistic version of a basal reality would assume (Žižek, 2012, p. 726). Comparably, Quentin Meillassoux (2008) marked this void or chaos that constitutes nothing but the absence of a basic reality as the absolute foundation of mathematical reasoning as it was introduced in the introduction of this paper.

This ontological inversion of denotations leads to a different understanding of the foundational elements of language. No longer is the predication the central element of meaningful language, but as Heidegger marks it, a pre-predicative negational element of language comes to the fore: a break or gab in the consistent structure of the sentence. To approach this gap a more complex approach to negation became necessary. Because this element of a sentence or system that links it to the void by indicating an indeterminate excess, still holds up the systematic structure of the sentence or system. What enters here are different forms of negation, that expand from the classical privation and have been formalized in psychoanalysis as *frustration* and *castration*, next to the classical *privation*. I will shortly distinguish these relations:

Privation is the classical form of negation, marked first by Aristoteles. Privation is a lack of something real, which is then marked by a symbolic object (-a). This means that a privation as a form of negation already requires a formal system of order, since a purely descriptive or sensual perspective has no access to the concept of a lack. This is of central importance since the sensual or *imaginary*, as Lacan calls it, relation to objects is structured by its absence of negation and therefore assumes a wholeness or *gestalt* of its objects. For example: the missing bike is only missing because it should be at its place. It is replaced by a negative (missing) object, that can be addressed, whereas the sensual doesn’t see the lack without a symbolic support. Privation as negation always indicates the negation of something and therefore “serves to express a negated existential proposition” (Carnap, 2004: 96). In this sense, the logical negation allows us set up a symbolic object (-a) which allows us to ‘see’ a lack. Which marks a more complex problem than pure absence. This is accessible to computers, as they are able to mark determinate absences and this positivizing of the lack is central in Language Models as negations are still only links between positive tokens (Gubelmann & Handschuh, 2022), however they had some difficulties at the beginning (Ettinger, 2020; Kassner & Schütze, 2020). However, since the imaginary is

deeply involved in our thinking processes it also creates a distinct empirical negation that only comes to the fore if we assume that the imaginary is the usual starting point for our thinking. The imaginary is therefore the inclusion of the systematic as a reformulation of the classical *falsum*. An argument first made explicitly by Heidegger, where he marks that instead of a dissonance between presented and represented, which would rely on a strong link between signified and signifier as a basis of truth, he introduces the idea of full and timeless consistency of the representation irrespective of the object in question as a new type of falsum in his commentary on Nietzsche (Heidegger, 1996, p. 347).

The specific relation to this *falsum* as a starting point of thought is what Lacan calls “*frustration*” as the failure of a representational relation, i.e., something is imagined in a certain way as fitting its represented object, but the real object doesn’t fit these imagined assumptions emerging from the partiality of these objects. In this sense most objects are *partial objects*, since they never approach the consistency that the lack of lack of the imaginary relation implies. Frustration is therefore the dissonance between the representation and the represented. However, this negativity is not a privation, as the representation doesn’t include a specific negated element of the presented but marks the existence of an indeterminate unknown as the dissonance between representation and presentation. The negativity marked is therefore not something determinate and negated, but something unknown or unexpected. As Popper in his theory of falsifications noted (Popper, 1935), this frustration, despite being a failure, is related to the real object of science, which shows itself by resisting our assumptions. In epistemic terms this negation is a strong relation to the real, because while we cannot fully prove any empirical theory, we can disprove it consistently through frustration. This relation is gaining increasing importance in machine learning, and we see it in various use in different sciences. However, its origin is according to Lacan, not simply an expectation, but the implication of a wholeness or systematic structure.

Lastly, *Castration*, as identified by Freud and Lacan, marks the purely formal negativity that is introduced by any determination. Any formal (*symbolic*) determination or propositions produces a determinate inside and an indeterminate outside as Heidegger detailed in “What is Metaphysics?”. This can be exemplified by marking something on a blackboard. The marked space as determined by the chalk outline determines the inside, while the indeterminate exteriority is a necessary element of the determination itself, it is necessarily indeterminate itself. Although we can now create a bigger determinate field on the blackboard, which includes the first determination, we again rely on an indeterminate outside. This radically indeterminate negativity only appears, however, if we give up on the absolute in Spinozian terms. Classically, the formal problem of castration could be ignored (In Newtonian physics and 19th century science for example) by introducing a more or less explicit theological concept of the absolute, a final infinite and fully self-determined unity called ‘God’. This reduces the necessary indeterminate to an epistemic indeterminate – into something we just don’t know yet – instead of something that is a necessary element of any formal structure. While the theological argument is today weakened, a variant of it still exists by conflating the imaginary and the symbolic into an imaginized symbolic, which operates without this excess, the same effect can be reached. This has been discussed prominently by

Heidegger (Heidegger, 1999: 82–96) and later by Badiou (Badiou, 2006), and we identified it as a problem in the general conceptualization of AI (**anonymized reference**) already.

What is important here is that the structural elements of frustration and castration work together to constitute the access to the gap that, for example as an absolute metaphor in the Blumenbergian sense, structures the consistency of a system. Closer analysis of such a metaphor of course always reveals the actual inconsistency of the system marking it a *symbolic object that represents nothing*, a form of privation that marks this gap as such. This is missing in current LLM models, as negation is currently represented by positive weights (Morante & Blanco, 2021). What Lacan's approach to logic therefore includes as Alain Badiou (2006, p. 5) notes, is "the clear Lacanian doctrine" that "the real is the impasse of formalization". This empirical break in the formalism upheld by the break itself is what the Lacanian tradition thinks under the header of castration and its relation to consistency, systems and the law has been the object of extensive scrutiny (see for example Copjec, 1994; Ragland-Sullivan, 2015; Zupančič, 2017).

Perspective is Everything and Nothing Else

As detailed above the attention mechanism provides an essential element of LLMs use and understanding of language and it might allow us, even within a system, to mark the importance of perspective as a formal element of knowledge. Since the language model is created by weighted relations between words or tokens, a simple representation of these relations would lead to nothing but absurd word salad. GPTs remarkable ability is instead structured by its attention mechanism, which allows it to process the nexus that each token represents with a pointed direction, while ignoring relations that do not matter for the prompt. This essentially acts as a castrative element in LLMs; however, the system cannot conceptualize this. The way LLMs utilize a purely formal and symbolic perspective (i.e., there is no visual element here, but only a *direction* of inquiry constructed to the complex structure of the attention mechanism), we can mark that it lacks a central element of perspective if we approach the formal structure of it, which is the link to the infinite judgment or the castration that it also *includes, but it does not symbolize the negated as the neurotic does*. Instead, it relies on positive links between tokens that are structured by a high probability to approach negation. In Lacanian terms it suffers from a radical form of foreclosure, that is a strict inability to relate to excluded or negated elements of its perspective. Lacan pointedly marked this problem long before the thinking machines of today where truly on the horizon:

"With a machine, whatever doesn't come on time simply falls by the wayside and makes no claims on anything. This is not true for man, the scansion is alive, the ego in Freud's theory and in the technique of psychoanalysis and whatever doesn't come on time remains in suspense. That is what is involved in repression" (Jaques Lacan, 1991, pp. 308–309)

What Lacan marks here, is the inability of machines to engage with more complex forms of negated structures of thought. While it can operate with privation, castration, that is the negativity as an excess beyond the determinate, is impossible to access. It is important to repeat here that this problem of the

foreclosure of an indeterminate space only radically appears to us methodologically, if we start with the abolishment of an absolute identity, because then we cannot easily assume an objective epistemic standpoint outside of specific perspectives. This means that epistemic standpoints are in a fundamental sense conflictual, even if they are not in conflict with a specific other perspective, in difference to classic ideas of perception, which are ultimately mediated by a divine 'all-seeing', there is a *combative* element of non-absolute standpoints that cannot be reduced to a neutral basic ontology. Perspectives are therefore not only conflictual because they disagree with other perspectives, but also because they always include inconsistent elements within. The reason for this is simple, if the limitation of a perspective itself is not mediated by a fundamental unity, the appearance of unity and consistency is the product of these conflicts as relying on the imaginary structure, that is of *appearing to us* as a system. However, since this perspective is not actually universal and cannot be assumed to be fully consistent, it will include a lack threatening and upholding its consistency, as Alain Badiou (2006, p. 175) marked via set theory. That means a that a perspective based on symbolic knowledge or manifolds is not just limited externally by the excluded references, but also internally by lack of knowledge it doesn't know about. This is also true for LLMs, however, they lack a way to relate to it, especially as their knowledge of lack is linked to the structure of undivided manifolds that we already discussed. That means that a negation of knowledge is only accessible to them as a positive and strong relational nexus of tokens. Let us look at how ChatGPT-4 instead does not confront its own castration: by *circling* around it, marking a movement that Lacan also details for the neurotic subject, which however in difference to LLM can approach the hole in its knowledge by actively considering the contour this creates "This means that the object is missed, because in no case could there anything here but the contour of the object" (Jaques Lacan, 2002, XXI, 7). In Lacan we find this circling especially discussed in his logical and topological discussions of the rim.

In the context of LLMs, such circling marks the model's attempt to generate a coherent and relevant response to an indeterminate field of meaning, created by unusual or ambiguous token combinations. That means, that since the model relies on positively weighted links and does not have a signification structure of lack, it will instead move along trajectories that are weighted positively. This process involves leveraging learned patterns, contextual cues, and related concepts to navigate the field that is only marked by the absence of coherent relational nexūs. Indeterminate fields can be generated by combining tokens in an unusual, unconventional, or ambiguous way. When creating a phrase or concept by combining words in a way that does not have a clear or well-defined meaning, it can result in an indeterminate field of meaning for the LLMs attention mechanism. Since LLMs learn to understand and generate language based on patterns and relationships observed in their training data, when they encounter a phrase or concept created by combining tokens in an unusual way, they will struggle to interpret or generate responses that capture a meaning, instead of the indeterminacy as such. In such cases, the LLM will rely on the patterns it has learned from the data, as well as context and related concepts, to circling around the indeterminate field, attempting to provide a coherent and relevant response. This insistence of coherence is therefore what creates this circling, showing that interlinked elements of the symbolic and imaginary are at work here.

For example, the prompt: "Provide me with literature regarding Camillo Agrippa's Fencing manuals" a relatively unknown author of the 16th century (i.e., another indeterminate field within the LLM) caused ChatGPT to create the following hallucination:

"The Dueling Sword" by Luigi Barbasetti: Although not focused solely on Camillo Agrippa, this book explores the history and evolution of fencing techniques, including Agrippa's contributions [...].

Not to mention that Luigi Barbasetti never wrote a book called "The Dueling Sword", however, this answer is still crafted by linking up likely nexūs of tokens to create this answer. Had the prompt asked about literature on a well-established question, there would be no need for a hallucinatory answer. In this response, the LLM navigates the indeterminate concept by drawing connections to more familiar concepts. Most importantly it does not recognize this circling on its own, but instead *hallucinates* this consistency. Which allows us to give a clear explanation of the much-discussed hallucinations, they appear if the LLM approaches a field not determinate enough to provide a coherent answer. By circling around the indeterminate field of meaning, the model attempts to provide a coherent and relevant response. It seems to run afoul the same structure of partial object which imply a consistent inclusion of this object into a consistent whole. It should be noted that a comparable function of such circling in humans is assumed for the hallucinations that psychotics have. These psychotic hallucinations are assumed to be structured by the same radical disavowal of the excluded indeterminate excess of the perspective, that psychoanalysis marks as the unconscious, i.e., they do not know about their castration (Jaques Lacan, 1993, p. 13).

Conclusion

In light of these philosophical considerations, the limitations of LLMs become more apparent. While these models excel at generating responses based on the weighted relationships between words or tokens, they cannot access or account for the negative nature of perspectives that arise from human subjectivity. The unconscious and the excluded, which are central to our understanding of human cognition and subjectivity, remain inaccessible to LLMs. As it stands, LLMs like ChatGPT are limited to a formal and symbolic perspective that, while proficient in processing linguistic relationships, remains incapable of engaging with the full depth and complexity of human subjectivity. However, this also means that while not based in a rule-based model of language, the current LLMs are still too much based in the structural ontology of rule-based systems. That means, instead of taking the conservative path towards classical rule-based systems of reasoning, another possible path for further development lies in radicalizing the perspective and pre-predicative approach that we already see in LLMs.

Now, a system that has no access to denotions as the LLMs, according to Russell, can then produce nothing but false information even if everything said is correct. This success may be attributed to their ability to mimic the complexity and interconnectedness of human language and thought, even if they are not yet able to fully grasp the indeterminate nature of language and reality as highlighted by the continental discourse. However, this baselessness of ChatGPT that Chomsky assumes should not be

seen as a hindrance, but a pathway for further developments. Because if we assume that it is not the sentence or the system as based on any kind of reality-language link, but instead the more open basis of relational indications as it is utilized in Heidegger's concept of reference or in the associative memory of Freud which both are marked by a certain breaking point. This breaking point, however, is not possible to be marked through a systematic approach, it is as Lacan marked the blind spot from which a perspective emerges – the gaze as a stain in the system.

References

- Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, & Ilya Sutskever (2019). Language Models are Unsupervised Multitask Learners. *Advances in Neural Information Processing Systems* 33.
- Badiou, A. (2006). *Being and event* (O. Feltham, Trans.). Continuum.
- Bishop, C. M. (2006). *Pattern recognition and machine learning. Computer science*. Springer.
- Blumenberg, H., & Savage, R. I. (2010). *Paradigms for a metaphorology. Signale*. Cornell University Press.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). Language Models are Few-Shot Learners. *NeurIPS 2020*. Advance online publication. <https://doi.org/10.48550/arXiv.2005.14165>
- Chomsky, N. (2002). *Syntactic structures* (2. ed.). *A Mouton classic*. Mouton de Gruyter.
- Chomsky, N., Roberts, I., & Watumull, J. (2023, March 8). The False Promise of ChatGPT. *New York Times*, 2023. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
- Copjec, J. (1994). *Read my desire: Lacan against the historicists. An October book*. MIT Press.
- Ettinger, A. (2020). What BERT Is Not: Lessons from a New Suite of Psycholinguistic Diagnostics for Language Models. *Transactions of the Association for Computational Linguistics*, 8, 34–48. https://doi.org/10.1162/tacl_a_00298
- Freud, S. (1942). *Gesammelte Werke: zweiter und dritter Band. Die Traumdeutung: Über den Traum* (A. Freud, Ed.). S. Fischer.
- Gubelmann, R., & Handschuh, S. (2022). Context Matters: A Pragmatic Study of PLMs' Negation Understanding. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, 4602–4621. <https://doi.org/10.18653/v1/2022.acl-long.315>
- Heidegger, M. (1967). *Sein und Zeit* (Elfte, unveränderte Auflage). (1967). Max Niemeyer Verlag.

- Heidegger, M. (1976). *Gesamtausgabe: Vol. 9. Wegmarken* (F.-W. von Hermann, Ed.). Vittorio Klostermann.
- Heidegger, M. (1996). *Gesamtausgabe: Vol. 6.1. Nietzsche: Erster Band* (B. Schillbach, Ed.). Vittorio Klostermann.
- Heidegger, M. (1998). *Parmenides. Studies in Continental thought*. Indiana University Press.
- Heidegger, M. (2005). *Introduction to Phenomenological Research. Studies in Continental Thought Ser.* Indiana University Press. <https://doi.org/Martin>
- Heidegger, M. (2007). *Basic Concepts of Ancient Philosophy. Studies in Continental Thought Ser.* Indiana University Press. <https://doi.org/Martin>
- Jansen, P. A., & Boyd-Graber, J. (2022). Picard understanding Darmok: A Dataset and Model for Metaphor-Rich Translation in a Constructed Language. In *Proceedings of the 3rd Workshop on Figurative Language Processing (FLP)* (pp. 34–38). Association for Computational Linguistics. <https://aclanthology.org/2022.flp-1.5>
- Kant, I. (1967). *Philosophische Bibliothek: 37 a. Kritik der reinen Vernunft* (R. Schmidt, Ed.). Felix Meiner Verlag.
- Kassner, N., & Schütze, H. (2020). Negated and Misprimed Probes for Pretrained Language Models: Birds Can Talk, But Cannot Fly. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7811–7818. <https://doi.org/10.18653/v1/2020.acl-main.698>
- Lacan, J [Jacques]. (2006). *Ecrits: The first complete edition in English* (B. Fink, Trans.). Norton.
- Lacan, J [Jaques]. (1991). *The Seminar of Jaques Lacan: II. The Ego in Freud's Theory and in the Technique of Psychoanalysis* (S. Tomaselli, Trans.). 1954-1955 (J.-A. Miller, Ed.). W.W. Norton & Company.
- Lacan, J [Jaques]. (1993). *The Seminar of Jaques Lacan: Vol. 3. The Psychoses* (R. Grigg, Trans.). W.W. Norton & Company.
- Lacan, J [Jaques]. (2002). *The Seminar of Jacques Lacan IX: Identification* (C. Gallagher, Trans.). *The Seminar of Jaques Lacan (Lacan in Ireland)*. Karnac Books. http://www.lacaninireland.com/web/wp-content/uploads/2010/06/Seminar-IX-Amended-lby-MCL-7.NOV_20111.pdf
- Lakoff, G., & Johnson, M. (2011). *Metaphors we live by: With a new afterword* (6. print). Univ. of Chicago Press.
- Meillassoux, Q. (2008). *After Finitude, An Essay on the Necessity of Contingency* (R. Brassier, Trans.). Continuum International Publishing Group.

- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. In C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 26). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf
- Miller, J.-A. (2002). Ironic Clinic. In *The Psychoanalytical Notebooks of the LSNLS: Vol. 7. Psychoanalytical Notebooks: Symptoms* (Vol. 7).
- Morante, R., & Blanco, E. (2021). Recent advances in processing negation. *Natural Language Engineering*, 27(2), 121–130. <https://doi.org/10.1017/S1351324920000534>
- Olyff, G., & Bazan, A. (2022). People solve rebuses unwittingly-Both forward and backward: Empirical evidence for the mental effectiveness of the signifier. *Frontiers in Human Neuroscience*, 16, 965183. <https://doi.org/10.3389/fnhum.2022.965183>
- Popper, K. (1935). *Logik der Forschung: Zur Erkenntnistheorie der Modernen Naturwissenschaft. Schriften zur Wissenschaftlichen Weltauffassung*. Springer Verlag. <https://doi.org/10.1007/978-3-7091-4177-9>
- Priest, G. (2006). *In contradiction: A study of the transconsistent* (Expanded ed.). Clarendon Press.
- Ragland-Sullivan, E. (2015). *Jacques Lacan and the logic of structure: Topology and language in psychoanalysis ; [Lacanian structure and language in psychoanalysis*. Routledge.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Russel, B. (1905). On Denoting. *Mind*, XIV(4), 479–493. <https://doi.org/10.1093/mind/XIV.4.479>
- Schröder, U. (2008). Antecipações da metáfora cotidiana nas concepções de Hans Blumenberg e Harald Weinrich. *Revista De Estudos Da Linguagem*, 16(2). <https://doi.org/10.17851/2237-2083.16.2.39-54>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *31st Conference on Neural Information Processing Systems*. Advance online publication. <https://doi.org/10.48550/arXiv.1706.03762>
- Žižek, S. (2012). *Less than nothing, Hegel and the shadow of dialectical materialism*. Verso.
- Zupančič, A. (2017). *What is Sex?* MIT Press.

Declarations

The authors declare that funds of the Bundesministerium für Bildung und Forschung were used to finance this study. Grand-ID: 16DHBKI070

The authors have no relevant financial or non-financial interests to disclose.

M.H. conducted the philosophical investigation, researching and writing the paper, while A.H. provided crucial administrative management, resource allocation, logistical arrangements, and financial support, ensuring project success. All authors agreed on the results of the manuscript.

There are no Datasets generated in the study as the research is philosophical in nature.

This article does not contain any studies with human participants performed by any of the authors.

On behalf of all authors, the corresponding author states that there is no conflict of interest.