Jurnal Filsafat, ISSN: 0853-1870 (print); 2528-6811(online)
Vol. 35, No. 2 (2025), p. 427 – 454, doi: 10.22146/jf.105975

COMPUTATIONAL AND THOUGHT LANGUAGE: BEYOND THE SEARLE-DENNETT PARADIGM

Muhammad Andry Mukmin
Universitas Indonesia, Depok, Indonesia
Email: andrymukmin@gmail.com

L. G. Saraswati Putri
Universitas Indonesia, Depok, Indonesia

Abby Gina B. M.
Universitas Indonesia, Depok, Indonesia

Abstrak

Penelitian ini membongkar ketegangan ontologis dalam perdebatan John R. Searle dan Daniel C. Dennett terkait hubungan bahasa dan kesadaran, lalu mengusulkan model computational language dan thought language sebagai kerangka analisis baru. Computational language dipahami sebagai bahasa sintaksis berbasis algoritma tanpa intensionalitas. Sebaliknya, thought language adalah ekspresi konseptual-intensional yang berakar pada kesadaran fenomenologis. Melalui analisis konseptual-kritis, penelitian ini menunjukkan bahwa kecerdasan buatan hanya mampu menghasilkan computational language, sedangkan thought language tidak dapat direduksi menjadi operasi komputasional. Bukti neurokognitif, seperti aktivasi area Broca dalam pemrosesan sintaksis, menguatkan deskripsi computational language, sementara thought language tetap diposisikan sebagai konstruksi ontologis yang melampaui pembuktian empiris. Model ini tidak hanya memetakan batas epistemik kecerdasan buatan, tetapi juga membuka ruang refleksi baru dalam filsafat pikiran, bahasa, dan teknologi.

***Kata kunci**: bahasa, kesadaran, kecerdasan buatan, computational language, thought language.*

Abstract

This study examines the ontological tension in the debate between John R. Searle and Daniel C. Dennett regarding the relationship between language and consciousness, and then proposes the model of computational language and thought language as a new analytical framework. Computational language is understood as a syntactic, algorithm-based language without intentionality. Conversely, thought language is a conceptual-intentional expression rooted in phenomenological consciousness. Through a conceptual-critical analysis, this study demonstrates that artificial intelligence can only produce computational language, whereas thought language cannot be reduced to computational operations. Neurocognitive evidence, such as activation of Broca's area in syntactic processing, supports the description of computational language, while thought language remains positioned as an ontological construct beyond empirical verification. This model not only maps the epistemic limits of artificial intelligence but also opens new avenues for reflection in the philosophy of mind, language, and technology.

Keywords*: language, consciousness, artificial intelligence, computational language, thought language.*

Received: April 04, 2025| **Reviewed**: August 13, 2025| **Accepted**: October 16, 2025

INTRODUCTION

The rapid development of artificial intelligence (AI), particularly large language models (LLMs) such as GPT, has reignited fundamental questions in the philosophy of mind and language: Do computational systems truly understand the language they produce? Or do they merely simulate meaning through statistical patterns and symbol manipulation? Beneath this technical inquiry lies a deeper philosophical problem: does language necessarily presuppose consciousness, or can it be fully explained through information processes? This debate strikes at the heart of

the ontological status of consciousness and the epistemic status of language in non-biological systems.

Kim (2024: 7) explains that "Our logical analysis provides a crucial theoretical foundation for such investigations by establishing the fundamental impossibility of valid consciousness denial." This means that if a system is complex enough to issue claims about consciousness, the claim of "being unconscious" must be regarded as problematic. The relevance is that this supports the position that if AI has already reached the capacity to "claim the development of consciousness," then the ontological and epistemic status of the system must be taken into account, rather than dismissed as "mere simulation."

The predictions of several thinkers, such as Tegmark, suggest that Life 3.0 will not only be able to design its own software but also its hardware. In other words, Life 3.0 becomes the master of its own destiny, ultimately free from the constraints of biological evolution. Its impact extends across economic, legal, and military domains, where AI may equal or surpass human-level intelligence. This notion points toward the development of Artificial General Intelligence (AGI) (Tegmark, 2017: 45).

Similar concerns are also found in Ray Kurzweil's work, which envisions AI systems reaching the capacity for recursive self-improvement and exponential intelligence growth, potentially transforming human civilization in radical ways (Kurzweil, 2005: 135). Furthermore, Nick Bostrom expands this discourse by emphasizing the existential risks that could emerge from entities possessing cognitive capacities far beyond human control (Tegmark, 2017: 53). Meanwhile, Stuart Russell highlights the alignment problem—the challenge of ensuring that increasingly autonomous AI systems remain aligned with human ethical values and goals (Tegmark, 2017: 51-53).

In contemporary philosophy, John R. Searle and Daniel C. Dennett are two central philosophers who hold opposing positions on this issue. Searle argues that consciousness is a biological phenomenon that cannot be reduced to computational systems.

Therefore, AI-generated language lacks genuine meaning because it is not grounded in intentionality. In contrast, Dennett rejects the assumption that consciousness must be subjective and phenomenological. He proposes that consciousness is merely the result of complex information processing, and thus understanding—including in language—can emerge in non-biological systems if they are sufficiently sophisticated.

From this difference arises an *aporia*[1]: how can we explain the relationship between language, consciousness, and computational systems without falling into reductionism or essentialism? This impasse demands a philosophical reading that does not merely juxtapose two extreme positions, but instead seeks an alternative approach that is more inclusive both ontologically and epistemologically.

Through conceptual analysis as a framework, this study uses a philosophical comparative method to construct a comparison of the two philosophers under study. However, this approach does not stop at a static comparison; rather, it is directed toward identifying the problem of aporia, namely a logical impasse that arises from three propositions that appear individually valid but cannot be upheld simultaneously: (1) that language is an expression of consciousness; (2) that consciousness cannot be reduced to computational systems; and (3) that language can be fully explained computationally. Following that, this study aims to reveal Searle and Dennett's fundamental assumptions by introducing the distinction between *computational language* and *thought language*. These models serve as a philosophical synthesis and alternative concept to explain the epistemic status of language in artificial intelligence systems. This study aims to construct a more adequate framework for the philosophy of language and consciousness in facing the ontological challenges posed by contemporary AI technology.

---

[1] Aporetic, from the Greek word "aporia," refers to an impasse or blockage where there is no practical way forward. In philosophy, it signifies a puzzle, perplexity, or a problem that is difficult to overcome—or at least highly problematic (Rescher, 2009).

DISCUSSION

1. Comparison Between John Searle and Daniel Dennett

The debate on consciousness and language in contemporary philosophy of mind often centers on two key philosophers: John Searle and Daniel Dennett. Both attempt to explain the relationship between mind, language, and artificial systems, yet from fundamentally different frameworks. Chua (2017) has provided a synoptic comparative map of their commitments and points of divergence, which situates Searle's biological naturalism against Dennett's eliminativist-leaning functionalism.

John Searle (2002: 7–8) asserts that "above all, consciousness is a biological phenomenon. We should think of consciousness as part of our ordinary biological history, along with digestion, growth, mitosis, and meiosis." Thus, consciousness must be understood as part of the biological processes inherent in human life, on par with other biological functions. However, Searle immediately adds that "conscious states have a certain qualitative character; the states in question are sometimes described as qualia."

This is the phenomenological dimension of consciousness that cannot be reduced to purely physiological or computational descriptions. Within Searle's framework of biological naturalism, consciousness is rooted in biology but possesses an irreducible ontological status. Therefore, even though consciousness depends on the brain, it has first-person qualities that cannot be captured by computational or purely physical descriptions. As he states, "The problem of consciousness is not that it is mysterious or inexplicable. The problem is that we do not yet have, and we may never have, a theoretical apparatus adequate to explain how brain processes cause conscious states" (Searle, 2002: 83).

In this light, language for Searle is always an expression of intentionality. Syntax alone is never sufficient for semantics: "Syntax is defined in terms of symbol manipulation independent of semantics. But minds have contents; they have semantics. So, you

can't get semantics from syntax. You can't get meaning from formal symbol manipulation" (Searle, 1992: 209).

When Searle states that "syntax is not sufficient for semantics," he is emphasizing that the manipulation of formal symbols (syntax)—as performed by AI—is never enough to generate meaning (semantics). Therefore, this statement is not merely technical-linguistic, but an ontological claim. Searle insists that meaning arises only from consciousness, not from a purely symbolic system. Searle emphasizes that the execution of a computer program is merely the manipulation of formal symbols based on syntactic rules, without any involvement of semantic understanding. Even though the system may produce answers that appear correct externally, no semantics are present in the process. In other words, there is an ontological gap between formal symbol manipulation (which AI can perform) and meaningful experience (other mind).

Thus, the thesis that "syntax is not sufficient for semantics" functions as an ontological claim about the status of meaning—one that, in Chua's (2017) comparative reading, serves as the pivotal axis separating Searle's framework from Dennett's."Formulates the Chinese Room implication: because syntax never yields semantics, computer functionalism is insufficient to produce a conscious machine—at most a program manipulates symbols while understanding remains absent" (Chua, 2017: 54–55). This view leads Searle to reject the possibility that AI can never possess genuine understanding. Even if machines produce grammatically correct sentences, they remain simulations devoid of intentional meaning—a point famously illustrated by his Chinese Room Argument. Searle argues:

> "Imagine that you carry out the steps in a program for answering questions in a language you do not understand. I do not understand Chinese, so I imagine that I am locked in a room with a lot of boxes of Chinese symbols (the database), I

> get small bunches of Chinese symbols passed to me (questions in Chinese), and I look up in a rule book (the program) what I am supposed to do. I perform certain operations on the steps in the program and give back small bunches of symbols (answers to the questions) to those outside the room" (Searle, 1997: 11).

Thus, consciousness and understanding cannot be reduced to symbolic computation. In other words, no computer program—operating solely based on symbol manipulation—can generate genuine understanding or consciousness. This is the core of Searle's biological naturalism: meaning and consciousness can only emerge from living biological systems, not from purely computational devices.

Dennett, in sharp contrast, advances a form of eliminative materialism and radical functionalism. He denies the existence of irreducible qualia and rejects what he calls the Cartesian theater—the idea of a central locus where Consciousness Explained. He proposes the multiple drafts model: "There is no single, definitive 'stream of consciousness,' because there is no central Headquarters, no Cartesian theater where 'it all comes together' for the perusal of a Central Meaner" (Dennett, 1991: 253). According to this model, consciousness is nothing more than the distributed outcome of competing representational processes. Perception, cognition, and language are explained through parallel, multitrack operations that function adaptively without requiring an inner essence: "All varieties of perception… are accomplished in the brain by parallel, multitrack processes of interpretation and elaboration of sensory inputs" (Dennett, 1991: 135).

Dennett seeks to emphasize that consciousness is not a single stream centered in a Cartesian theater, but rather the result of the brain's work involving many pathways (multitrack) operating simultaneously (in parallel). Each representational process—whether perception, cognition, or language—never stands alone,

but instead competes, corrects one another, and contributes to the formation of adaptive behavior. For example, when we read a sentence, the visual pathway processes the shapes of the letters, the linguistic pathway processes the meaning of the words, and the memory pathway recalls related experiences. All these pathways function together simultaneously, not sequentially, thereby producing a unified understanding—precisely the point highlighted by Chua (2017: 48-49, 55), "Reads this as a deflationary stance toward qualia and a functional reconstruction of consciousness that abandons any biological essence."

Coherence of consciousness, in this view, does not arise from the existence of a central control or inner essence, but from the functional synchronization of various representational pathways. In other words, what we experience as consciousness is merely an emergent property of this complex interaction, not a special irreducible substance. Therefore, Chua (2017) interprets Dennett's position as a functional reconstruction of consciousness.

This contrast—Searle's biological naturalism versus Dennett's eliminative functionalism—creates a categorical divergence, not merely a methodological one. Their difference is not simply about how (method) to explain the same phenomenon, but about what is meant by consciousness itself. For Searle, the ontology of consciousness is biologically functional with an irreducible status. Thus, language is an expression of intentionality rooted in the experience of qualia. In other words, consciousness possesses ontological substance. For Dennett, on the other hand, the ontology of consciousness is the product of representational competition in the brain, where language is the outcome of representational competition and adaptive function.

Why is this divergence categorical? Because they are talking about different entities, even though they both use the term *consciousness*. For Searle, consciousness is a lived, irreducible, and phenomenological biological phenomenon. For Dennett, consciousness is merely the effect of distributed functional representation without any inner essence. From this perspective, if

a system—biological or artificial—can functionally demonstrate coherent behavior and linguistic competence, there is no reason to deny that it possesses understanding. For Dennett, what matters is not the substrate but the function.

Philosophical inquiry into consciousness and language often risks falling into two extremes: reductionism and essentialism. Reductionism arises when a complex phenomenon is explained at a single level only—for instance, treating consciousness solely as computational processes or solely as biological mechanisms—thereby neglecting its experiential, historical, and relational dimensions. Essentialism, by contrast, assumes a single, unchanging essence of language or consciousness. However, such essentialism overlooks the dynamics and context-dependent nature of human meaning-making, reducing living language to a metaphysical abstraction detached from experience.

This tendency persists because philosophy tends to demand categorical clarity (Vico, 1709/1990: 48), science prioritizes methodological simplification through the principle of parsimony (Scorzato, 2013: 21), and the Cartesian legacy continues to drive the search for definitive foundations of consciousness (Descartes, 1671/1996). By recognizing these pressures, it becomes clear why Searle and Dennett take different directions—between preserving phenomenological reality and asserting functional consistency—while at the same time preparing the ground for the next analysis, in which the resulting aporia will be examined explicitly.

From these positions arise sharply different conclusions about AI. For Searle, since AI lacks biological consciousness, its outputs cannot be considered as genuine language but only syntactic simulation. For Dennett, if AI systems reveal sufficiently complex representational structures and behavioral coherence, then they can, in principle, be said to understand and use language. This divergence also structures current debates on AI and language, as mapped by Chua (2017).

However, as Cain (2016: 90-92) observes, this debate between Searle and Dennett reaches a conceptual impasse. Both thinkers,

despite their opposing directions, fail to account for the origin and nature of meaning itself. Searle grounds consciousness in biology but cannot fully explain how intentionality arises from neural processes; Dennett, conversely, dissolves subjective experience into functional models, thereby neglecting its lived dimension. Cain writes, "Searle insists that consciousness is ontologically irreducible, while Dennett denies that there is any such thing as irreducible consciousness."

This critical deadlock opens the need for a third conceptual field—one that bridges symbolic computation and phenomenological experience, beyond both biological naturalism and eliminative functionalism. In this regard, Cain provides a valuable reading of Jerry Fodor's Language of Thought Hypothesis (LOTH) as a structural account of mental representation. Fodor proposes that "Mental processes are computations; that is, causal sequences of tokenings of mental representations" (Fodor, 1975: 55).

As Cain (2013) explains, Fodor's model conceives of the mind as a computational–representational mechanism in which meaning arises from internal symbolic relations rather than phenomenological experience. Thus, while Fodor offers a coherent framework for representational cognition, consciousness itself remains external to this symbolic system.

Building on this genealogy, the present study introduces a distinction between *computational language* and *thought language*. Unlike Fodor's LOTH, *computational language* refers to the non-conscious syntactic operations of artificial systems such as AI or LLMs. In contrast, *thought language* designates the meaningful, intentional, and experiential dimension of human linguistic consciousness. This distinction formulates an epistemological and ontological framework that moves beyond both Searle's biological essentialism and Dennett's reductionist functionalism, while also addressing the representational limits identified in Fodor's computational theory.

2. Aporia Problems Arising from the Searle-Dennett Debate.

Amid the debate between Searle's biological-intentionalist position and Dennett's functional-computational stance, a philosophical problem emerges that cannot be resolved while retaining both frameworks simultaneously. In philosophical tradition, this problem is referred to as an aporia—a condition of logical impasse in which two or more propositions that appear true cannot all be maintained at the same time without resulting in contradiction.[2]

In this context, the aporia arises from the following three propositions:

(1) Language is an expression of consciousness;
(2) Consciousness cannot be reduced to computational processes (Searle); and
(3) Language can be fully explained computationally (Dennett).

Individually, each of these propositions appears justifiable. However, when combined, they form a contradictory knot: If (1) and (2) are true, then (3) must be false—meaning AI cannot produce genuine language, since it lacks consciousness; If (1) and (3) are true, then (2) must be rejected—because consciousness must be reducible to computational processes for AI to produce language; and if (2) and (3) are true, then (1) cannot be upheld—since language can arise without consciousness, it need not be regarded as an expression of consciousness.

This tension shows that contemporary philosophy of mind cannot avoid the aporia: a situation in which three propositions appear rational individually, yet, collectively, contradict one another. In this context, it is important to adopt a methodological approach that treats contradiction not as a weakness, but as a means of conceptual clarification. As Rescher states, "Paradoxes are not cognitive malfunctions but rather indicators of conceptual fault-lines where clarification is needed most" (Rescher, 2009: 83). This

[2] Rescher (2009: 15) also explains that "a paradox is not a sign of irrationality, but a signpost to deeper inquiry." This serves as an important methodological foundation for justifying the exploratory and synthetic structure of the paper.

perspective strengthens the methodological stance of this study, which remains open to tension while making aporia a productive space for the emergence of new conceptual frameworks.

In formal logic, this means the three are mutually inconsistent—they cannot all be true at once. Maintaining them together leads to internal tension and conceptual obscurity. This signals that our conceptual framework regarding language and consciousness must be revised or further differentiated.

This study does not choose one of the three propositions as a whole, since each represents an important dimension of our understanding of language. Rather, the way out of this aporia is offered by reconfiguring the categorization of the term language itself through a distinction between two kinds of language that have thus far been conflated: computational language and thought language. Thus, this study proposes the conceptual distinction between computational language and thought languag*e* as an alternative model for bridging the tension between biological and computational approaches to consciousness.[3]

This distinction allows us to say that: proposition (1) is true in the context of thought language, namely language that arises from consciousness; proposition (3) is true in the context of computational language, namely language produced by information-processing systems without intentionality; proposition (2) remains valid as an ontological boundary: consciousness cannot be reduced, but not all forms of language require consciousness. Thus, the distinction between thought language and computational language does not reject the three propositions, but repositions them conceptually so that they do not cancel each other out. This is the function of philosophy when confronting an aporia: not to impose

[3] This approach aligns with dialectical philosophy, which does not reject contradiction but uses it as a reflective tool. In this context, the tension between Searle's and Dennett's views is not an obstacle, but an opportunity to formulate a new framework such as computational language and thought language.

binary choices, but to introduce differentiation of meaning so that the conceptual horizon does not collapse under a false dichotomy.

3. New Conceptual Model: Computational Language and Thought Language

This conceptual model is based on the conviction that the term *language* does not refer to a single homogeneous form but rather encompasses two operational forms that are ontologically, epistemologically, and functionally distinct. *Computational language* refers to language produced by non-conscious computational systems through syntactic and statistical operations, without the capacity to experience or understand meaning. Therefore, it is more appropriately understood as an epistemic simulation rather than an existential expression. Like an actor who memorizes a script without grasping its emotional content, an AI system can generate utterances without consciousness.

Neurocognitive evidence supporting the computational language framework becomes increasingly compelling when viewed through various cognitive neuroscience approaches. fMRI (functional magnetic resonance imaging) studies indicate that Broca's area, particularly the left inferior frontal gyrus, is a crucial center for syntactic processing, while its interaction with the temporal–frontal pathway plays a role in the hierarchical integration of sentence structures (Friederici, 2011: 1360). These findings align with the dual-stream model in neurolinguistics, in which the dorsal pathway is responsible for mapping syntactic structures and speech production, whereas the ventral pathway is more focused on semantic aspects and meaning comprehension (Hickok & Poeppel, 2007: 393–394).

In addition, electrophysiological evidence through event-related potentials (ERP) shows a consistency between syntactic violations and the emergence of the P600 component (Osterhout & Holcomb, 1992: 785). The P600 is understood as the brain's response to syntactic structural mismatches, indicating a process of syntactic

reanalysis. This finding is reinforced by Frisch et al. (2003), who argue that the P600 differs from the N400, which is more closely related to semantic anomalies, thereby demonstrating the existence of separate processing pathways for syntax and semantics.

This empirical data reveals that the brain's language system differentiates between structural syntactic operations and conceptual meaning-making. It strengthens the argument that computational language constitutes a simulative layer that operates by manipulating syntactic rules and statistical patterns to process linguistic structures formally. Through this mechanism, the system is capable of producing structurally coherent linguistic outputs but not necessarily of disclosing the phenomenological meaning associated with thought language. In other words, AI systems that mimic syntactic processing operate within the corridor of computational language without touching the realm of conscious experience that is the prerequisite for meaning.

Phenomena such as dreaming and sleep talking demonstrate that not all human linguistic expressions are the result of intentional consciousness. In the state of sleep talking, a person may utter words or sentences that sound coherent, yet without the involvement of the subject's intentionality. This shows that language, at certain points, can function as an automatic response that does not represent understanding. This condition illustrates how computational language—as a form of linguistic expression without intentional meaning—can occur even within the human body. Philosophically, this underlines the importance of distinguishing between forms of language that arise from conscious experience and those that are automatic or non-reflective.

This connection aligns with Daniel Kahneman's dual-systems framework, which differentiates between system 1—fast, intuitive, and unconscious—and system 2—slow, analytical, and reflective (Kahneman, 2013: 1–5). During dreaming or sleep talking, human cognitive activity more closely resembles the dominance of system 1: associations of ideas occur automatically, mental images arise without conscious control, and narrative logic is often loose or

inconsistent. Yet these experiences still carry strong emotional intensity, as if they were real.

In contrast, system 2 takes on its role through reflection and evaluation. Thus, dreams and sleep talking demonstrate how the human mind can move between these two cognitive modes—automatic-intuitive and reflective-analytical. This illustration not only supports Kahneman's framework but also affirms that human consciousness encompasses a dual dynamic: the capacity to undergo non-reflective cognitive flows while also possessing the ability to assess and ascribe meaning through reflection.

In contrast, *thought language* is a form of language rooted in subjective consciousness, intentionality, and phenomenological experience. This type of language does not merely convey information but reveals the world as it is lived and experienced by a subject. Thought language emerges from human engagement with the world—encompassing affect, context, values, and life history—elements that cannot be reduced to formal syntax.

Thought language is an expression of meaning that can only arise within active and rational human consciousness. Unlike computational language, which can be generated by systems without understanding (such as AI), thought language involves intentionality, reflection, and relations of meaning that are only possible within the rational condition of humans. In a rational state, humans do not merely process symbols but consciously experience and articulate meaning in relation to the world. Therefore, thought language cannot be possessed by beings or systems that lack reflexivity, because it requires the capacity to understand, evaluate, and ascribe meaning to emerging representations. This deeply distinguishes it from computational systems, which, although capable of producing sentences, lack access to the horizon of meaning or the value of the representations themselves.

In this sense, meaning cannot be separated from who speaks, in what context, and with what intention. Language becomes a bridge between the inner world and the outer world. *Thought language* is language in the Heideggerian sense: as "the house of

Being"—the space where meaning dwells, not merely a system of signs or a tool of communication. Within this framework, thought language can be understood as a form of dwelling in meaning that cannot be reached by computational systems, which only recognize syntactic relations between symbols. AI, operating within the corridor of computational language, remains within the realm of *Gestell*—the technological framework that enframes reality as something ready-to-hand (*Bestand*), rather than as an event of disclosure.

To clarify the distinction between these two types of language, the following table summarizes their conceptual differences:

Table 1. Conceptual Differences Between Computational and Thought Language

Aspect	Computational language	Thought language
Operational basis	Syntactic & statistical processing	Intentionality & subjective consciousness
Producing subject	Computational system (AI)	Human consciousness
Meaning	Generated from pattern & probabilities	Brought forth through experience
Relation to the world	Symbolic representational	Existential engagement
Ontology	Non-phenomenal, non-subjective	Phenomenal, grounded in lived experience.

This distinction implies that the term *language*, as it has been used in AI and philosophy of mind discourse, requires further qualification. The language produced by AI may indeed be called language, but only in the sense of computational language. It lacks the qualities that would make it a medium for understanding or the disclosure of meaning in the existential sense. By proposing this distinction, the study does not seek to deny AI's ability to produce language, but rather to provide a more accurate category for understanding what AI is actually doing when it uses language. This model spares us from two pitfalls: (1) assuming that AI can fully understand like a human being (technological utopianism),

and (2) regarding all AI language as merely functional expressions devoid of phenomenological meaning.

Moreover, the epistemological aspect of consciousness refers to the role of consciousness as the fundamental condition for the possibility of knowledge. Knowledge never appears as the mere accumulation of data, but is always rooted in the subject's connection with the world as consciously experienced. Consciousness enables information to be transformed into meaning, and it is this meaning that forms the basis for epistemic claims.

One of the crucial dimensions of consciousness is intentionality, namely, the directedness of the mind toward something beyond itself. Every form of consciousness—whether perception, memory, or imagination—is always *about something*. It is intentionality that bridges the subject with the object of knowledge, making knowledge possible only within the horizon of directed consciousness.

In addition, consciousness has a reflexive dimension. Humans not only know, but also are aware that they know. This reflexivity makes epistemic evaluation possible: weighing beliefs, testing doubts, and distinguishing genuine knowledge from mere opinion or illusion. The reflective dimension allows the conscious subject to regulate and expand their horizon of knowledge autonomously.

Finally, phenomenality is also an integral part of the epistemology of consciousness. How something is experienced from the first-person perspective—for example, the experience of seeing the color red—cannot be reduced to mere objective data. AI can process wavelengths, but it cannot experience the *qualia* of red. It is precisely this phenomenological experience that gives epistemic depth, because it ties the subject to the world in an existential way.

Thus, the epistemological aspect of consciousness encompasses intentionality, reflexivity, and phenomenality, which together make knowledge more than mere information. This is why thought language cannot be replaced by *computational language*. That is because only consciousness can sustain the epistemic structure in its entirety.

The dominant approach in cognitive science often relies on symbolic and computational reductionism. Within this paradigm, the mind is conceived as an information-processing system that can be reduced to a stream of symbolic input-output operations. Within this framework, the mind is reduced to a set of symbol manipulations without intentionality or intrinsic meaning. Such a model reduces the mind to a set of symbol manipulation and also neglects phenomenality and intentionality as essential aspects of human consciousness. Criticism of this approach has long been put forward by John Searle, who rejected strong AI on the grounds that computers, though capable of processing symbols, lack consciousness or understanding of those symbols. Searle's argument in the Chinese Room argument demonstrates that mere symbol manipulation is not sufficient to guarantee and generate genuine understanding of meaning—consciousness requires a biological and phenomenological basis that computational systems do not possess. This reinforces the critique of strong AI, which assumes that merely processing symbols is enough for a system to think like a human being. The idea of thought language emerges as a response to this reductionism: affirming that consciousness cannot be fully encoded within symbolic representational systems, because it always involves a horizon of experience, the lifeworld (*Lebenswelt*), and responsibility toward the Other[4].

Beyond epistemological critique, the reductionist approach also neglects the ethical and relational dimensions of consciousness. Emmanuel Levinas emphasizes that human beings cannot be reduced to mere biological entities or information systems. Humans are subjects who bear responsibility toward *the Other*, and it is

[4] The term liyan derives from the Arabic "al-ghayr" (الغير), meaning "the other" or "that which is not oneself." In philosophy—particularly in the thought of Emmanuel Levinas—liyan refers to the Other as a subject that ethically precedes the relation of knowledge and objectification. For Levinas, the Other is not merely an object of consciousness or representation, but a face that calls forth the subject's ethical responsibility. The relation with the Other is asymmetrical, as the subject can never fully grasp or master the Other's existence through rational categories or objective language. In this paper, the term liyan is used to emphasize that Thought Language entails openness to the presence and experience of other persons as irreducible subjects.

within this relation that existential meaning emerges. According to him, in the relationship between human beings, the world, and others, there is always an egological tendency in which the Other (*Autre*) is reduced to a mere reflection or extension of the Same (*Le Même*). He states: "The other is not the other because he is opposed to the same, but because he puts the same into question" (Levinas, 1961/1991: 39).

Furthermore, Emmanuel Levinas strengthens this critique by highlighting humanity's tendency to reduce the *"Autre"* to a mere extension of the "*Le Meme*". For him, the ethical relation precedes intellectual representation; the existence of the Other must not be understood solely within a cognitive framework but must be lived as an existential interpellation. In this view, consciousness is not merely a matter of knowing but a matter of responsibility. Emmanuel Levinas emphasizes that the relationship with the "*Autre*" is not a relationship of knowledge but an ethical relationship, because "the outside is nothing more than an extension of the self" (Tjaya, 2018: 47).

From this perspective, thought language is more than just an internal conceptual system; it is a medium of ethical engagement and openness to alterity. Language, in this sense, does not merely convey information but also implies a relationship of responsibility and recognition of the existence of the other. Thus, the thought language approach offers a broader horizon compared to computational models. It rejects the reductionism that diminishes consciousness into mere calculation and affirms that humans are beings who bear meaning, not merely processors of symbols.

Similarly, Donald Davidson rejects the possibility of a truly private conceptual scheme. He emphasizes the importance of triangulation—the interaction between the speaker, the interlocutor, and the shared world—in the formation of meaning (2001: 115). The concept of thought language aligns with this critique: language is not merely an internal structure, but a product of intersubjectivity and engagement with the world. Davidson adds that communication and meaning do not rest on language as a closed

system, but on an ongoing practice of interpretation within relations between subjects. Therefore, the discussion of consciousness cannot be separated from the social, historical, and practical contexts in which language lives.

Hilary Putnam developed the idea of internal realism, which rejects the claim that mental representations can fully map external reality in an objective and neutral way. In *Reason, Truth and History*, Putnam (1981: 19) argues that "meanings just aren't in the head", implying that meaning is always dependent on context, use, and social practice. As Putnam suggests, meaning is never simply "out there," but is formed through internal coherence within a system of beliefs. This statement affirms that meaning is not the result of a simple referential relationship between words and external objects, but rather emerges from the conceptual web that constitutes our network of beliefs. In other words, meaning arises because a word or statement holds a particular position within the broader structure of knowledge and belief.

The implications of Putnam's view are crucial: meaning cannot be separated from the human epistemic context. Putnam's view simultaneously serves as a critique of reducing language to mere external representation, as assumed in computational language. By contrast, thought language is rooted in the internal coherence of a belief system that lives within consciousness, such that meaning is always the product of interaction between the external world and the conceptual horizon, as well as the internal experiences of the subject.

The debate between Searle and Dennett on consciousness and language has not produced a final consensus. The debate reflects the productive tension in contemporary philosophy, where intellectual conflict can open space for a deeper understanding of complex issues such as language and consciousness. In the researchers' view, both Searle and Dennett fail to address the core problem of other minds—namely, how we understand the existence of another's subjective consciousness, which cannot be directly observed from our own perspective.

The failure to grasp the problem of *other minds* demonstrates that the issue of consciousness cannot be reduced to a mere biological versus computational distinction. It demands a broader framework that takes into account intersubjective relations, the historical dimension, and the reflective horizon of consciousness itself. Within this context, it is reasonable to revisit the Cartesian legacy of consciousness as a thinking substance while also assessing its limitations in explaining language and intersubjectivity.

Philosophical understandings of consciousness in relation to language have undergone significant transformation since the Cartesian era. In *Meditationes de Prima Philosophia*, René Descartes (1641/1996) declared *Cogito, ergo sum* as the existential foundation of the subject. In this Cartesian view, consciousness is a thinking substance (*res cogitans*) that is closed and autonomous, separate from the body and the external world (*res extensa*). This idea provided the basis for understanding language as the manifestation of internal mental processes. However, in the context of thought language, this Cartesian approach is considered insufficient because it fails to explain the relational, historical, and embodied aspects of conscious experience. As developed by Heidegger, Levinas, and other phenomenologists, human consciousness is not merely a private space for thinking, but rather an openness to the world and to the Other. Therefore, the thought language model not only corrects computational reductionism but also moves beyond the epistemological individualism of the Cartesian paradigm.

In this context, the concept of *meta-consciousness* becomes crucial because consciousness is not only about experiencing, but also about being aware of the experience itself. Meta-consciousness refers to consciousness of consciousness itself—the subject's capacity to be aware that it is thinking, feeling, or experiencing something. In a deeper philosophical sense, meta-consciousness can also be understood as the ability to comprehend and reflect on the limitations of one's experienced mental reality, while simultaneously acknowledging the possible existence of other consciousnesses. Its importance lies in its epistemological function

(correcting beliefs), phenomenological function (becoming aware of the qualities of experience), and ethical function (recognizing the existence of and responsibility toward the other). Without meta-consciousness, consciousness would be merely a flow of information, not a meaningful reflection.

Furthermore, when we examine the role of symbols in consciousness, we find that thought language is not only present as an internal system but also as a form of existential expression that carries layers of meaning beyond literal representation. Paul Ricœur explains that a symbol is not merely an arbitrary sign, but a hermeneutical form that opens the horizon of understanding toward experiences that cannot be fully articulated. He states that "the symbol gives rise to thought" (Ricœur, 1974: 288)—an acknowledgment that symbols in language enable reflection and the formation of meaning that is never final. This aligns with the idea of meta-consciousness as reflective awareness that recognizes itself through the ongoing mediation of symbols. In this context, thought language is the arena in which symbols serve as the unconcealment of existence.

In his famous essay "What Is It Like to Be a Bat?" Thomas Nagel (1974: 436) states that "The fact that an organism has conscious experience at all means, basically, that there is something it is like to be that organism… fundamentally an internal, subjective perspective." Every conscious being has an internal experience that objective descriptions cannot fully explain. Nagel's view reinforces the position that thought language cannot be reduced to mere computational language, because it involves phenomenological qualities and intentionality.

Furthermore, Ludwig Wittgenstein's "beetle in the box" metaphor in Philosophical Investigations highlights the limitations of objective language in explaining subjective sensations. He writes: "Suppose everyone had a box with something in it which we call a 'beetle.' No one can look into anyone else's box… the thing in the box does not belong to the language-game at all" (1953/2009, §293). This metaphor shows that others cannot access the meaning of

subjective experience and may even be irrelevant to communication. He emphasizes that each individual's inner experience cannot be externally verified, since we cannot open another person's "box." Experience, therefore, becomes an entity that cannot be fully objectified. Thus, language as a purely symbolic system is insufficient to represent conscious experience. What is needed is a form of expression that emerges from existential engagement—and this is the role of thought language.

From here, meta-consciousness becomes a way to transcend this solipsistic[5] structure, enabling the subject to authentically acknowledge the existence of other consciousnesses. The concept of meta-consciousness is not only a mechanism of self-reflection but also an epistemological and ethical foundation for understanding the interconnectedness of consciousnesses—something that cannot be explained by formal systems alone. Following that, thought language—which enables reflection, understanding of the Other, and acknowledgment of subjective experience—is the key to addressing the problem of Other Minds. In this context, Heidegger's approach of language as being is also relevant: language is not a tool but a field of *aletheia*. Thus, thought language functions not only as a system of subjective expression but also as an ontological bridge for establishing relations between consciousnesses.

The thought language model in this study is not merely a conceptual distinction between language as a product of consciousness and language as a formal system, but also offers a critical framework for evaluating the limitations of current AI

[5] The term *solipsistic* derives from *solipsism*, a philosophical position holding that only the self (subject) can be known with certainty to exist, while the existence of the external world and other subjects cannot be rationally proven. This concept was first explicitly formulated in modern philosophy by René Descartes (1641/1996) through his methodological doubt in *Meditationes de Prima Philosophia*. In its extreme form, solipsism was developed by thinkers such as George Berkeley, who emphasized subjective idealism, and was criticized by existentialist and phenomenological philosophers such as Martin Buber and Emmanuel Levinas for ignoring openness toward the other. In the context of this paper, the term solipsistic is used to refer to a way of thinking that locks consciousness within interiority, without acknowledging ethical and existential relations with other subjects.

systems. Recent advances in large language models (LLMs) demonstrate that their linguistic competence emerges from large-scale statistical optimization rather than semantic understanding. Technically, models such as GPT-4, or Claude, rely on transformer architectures (Vaswani et al., 2017, pp. 2-4), which analyze contextual relationships among tokens to predict the most probable subsequent word based on billions of linguistic parameters. The resulting fluency is thus a function of probabilistic pattern recognition trained on massive text corpora (OpenAI, 2023). Although this enables impressive linguistic coherence, the process remains syntactic and correlational, not grounded in lived or intentional meaning. In this sense, LLMs operate purely within computational language—a system of structured predictions devoid of phenomenological experience.

More recently, Goertzel discusses AGI architectures that incorporate self-referential meta-learning mechanisms, enabling systems to modify both base goals and meta-level evaluation metrics dynamically. This higher-order adaptivity—approaching what might be termed "super-meta AI"—is explored in his treatment of dynamic goal and metagoal evaluation (Goertzel, 2024, pp. 23-26). However, even in this self-referential architecture, learning is algorithmic—it accumulates information and reconfigures patterns without the capacity for intentional awareness. From the perspective of thought language, such systems exemplify the maximal expansion of computational language: They evolve structurally but remain phenomenologically empty.

Thus, this advancement does not necessarily signify the emergence of artificial consciousness, but rather the expansion of the operational boundaries of the symbolic system itself. The development of LLMs and super-meta AI demonstrates that artificial intelligence can imitate the cognitive structures of the human mind. Yet, it remains incapable of reproducing the horizon of consciousness that gives rise to meaning and experience. It is at this point that the distinction between computational language and thought language gains its urgency: the former operates at the

syntactic and predictive levels, while the latter is rooted in the intentional, reflective, and phenomenological dimension that can only arise through a consciousness aware of itself. Therefore, the debate about AI should not just focus on how far machines can *think*, but also consider how they cannot experience—namely, the absence of the dimension of lived meaning that makes human language not merely a system of signs, but a space for the disclosure of existence.

Approaches in contemporary philosophy of AI, such as those developed by Luciano Floridi and Andy Clark, remain deeply rooted in the informational and representational paradigm. In much of today's AI literature, such as Luciano Floridi's *infosphere* approach and Andy Clark's concept of *predictive processing*, intelligence is often understood as a system capable of mapping and responding to the environment efficiently based on information calculations and external representations. In *The Philosophy of Information*, Floridi (2014) conceives reality as an *infosphere*—an ontological space constructed from data relations, where meaning is reduced to the outcome of informational interactions. Within this framework, both humans and AI are regarded as informational entities that exist and operate within the infosphere. This implies that social interactions, moral actions, and cognitive processes can essentially be seen as exchanges and transformations of information. Floridi emphasizes that the consequences of this view are not only technical but also ethical and epistemological: the basic principles of ethics can be reformulated as an *information ethics*, and epistemology can be reconstructed algorithmically, insofar as it can be mapped in the form of information processing (Floridi, 2011: 9–12). Thus, the infosphere is not merely a metaphor but a new ontology that places information as the fundamental category of reality, replacing the dominance of matter or energy within the framework of classical ontology.

Meanwhile, Andy Clark, in his book *Surfing Uncertainty*, develops the theories of predictive coding and the extended mind, which revolutionize our understanding of consciousness. According to Clark, the brain does not function as a closed system

that merely responds to stimuli, but as a predictive machine that actively constructs models of the world to minimize prediction error. This process is called *predictive inference*—the brain continuously anticipates sensory input and adjusts its predictions when discrepancies with reality arise (Clark, 2016: 37–41).

This is where the key position of the thought language model lies. This concept rejects the reduction of consciousness into mere processes of data processing or prediction. The researchers assert that symbols or data cannot fully represent meaning. However, meaning emerges from the subjective relation between the subject and the world. Intentionality shows the directedness of the mind toward objects in the world—something that cannot be reduced to a mere flow of information. Phenomenality highlights how the world is subjectively experienced. Information and functionality only explain the external and operational dimensions of consciousness, while intentionality and phenomenality emphasize its internal and existential dimensions. In other words, thought language is grounded in phenomenology, not calculation: it recognizes that the human internal language is an expression of experience with intentional depth—something that information-based approaches or external signal-processing systems cannot reduce. If computational language merely replicates the surface of meaning, then thought language is the existential condition that enables meaning to be expressed as a responsibility toward the Other. Here, philosophy not only distinguishes between two types of language but also between two ways of being in the world: being as processing and being as presence.

CONCLUSION

This study demonstrates that the debate between John Searle and Daniel Dennett on the relationship between language and consciousness leaves an *aporia* that cannot be resolved through a simple biological versus computational dichotomy. By proposing a distinction between computational language and thought language,

this study offers a new conceptual framework that facilitates a more proportionate understanding of the relationship between language, consciousness, and artificial intelligence. This model shows that not all linguistic expressions contain intentionality, and not all symbolic processing can be equated with the experience of meaning.

Philosophically, this approach helps to remap the territories of the philosophy of mind and philosophy of language by avoiding both biological and algorithmic reductionism. Thought language affirms that meaning is not merely the result of symbolic calculation but arises from the existential condition of consciousness. Therefore, in the face of advances in AI systems, we are challenged not only to maintain humanity's position as an epistemic subject but also to ensure that the dimension of meaning is not uprooted from the soil of consciousness itself. Thus, this study is not merely conceptual but also offers an ontological and ethical foundation for assessing the epistemic limits of technology within contemporary knowledge structures. This model also paves the way for the formation of technological ethics principles that are more sensitive to the existential boundaries of human beings amid the dominance of non-conscious intelligent systems.

REFERENCES

Cain, M. J. (2013). *Fodor: Language, mind and philosophy*. Polity Press.

Cain, M. J. (2016). *The philosophy of cognitive science*. Cambridge: Polity Press

Chua, J. M. A. (2017). Formulating consciousness: A comparative analysis of Searle's and Dennett's theory of consciousness. *Talisik: An Undergraduate Journal of Philosophy*, *4*(1), 43-61.

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Davidson, D. (2001). The Second Person. In *Subjective, Intersubjective, Objective* (pp. 107–121). Oxford Clarendon Press. (Oroginal Work Published 1992)

Dennett, D. C. (1991). *Consciousness explained*. Back Bay Books.

Descartes, R. (1996). *Meditations on first philosophy* (J. Cottingham, Trans. & Ed.). Cambridge University Press. (Original work published 1641)

Frisch, S., Kotz, S. A., von Cramon, D. Y., & Friederici, A. D. (2003). Why the P600 is not just a P300: the role of the basal ganglia. *Clinical Neurophysiology*, *114*(2), 336–340. https://doi.org/10.1016/S1388-2457(02)00366-8

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Fodor, J. A. (1975). *The language of thought*. Crowell.

Floridi, L. (2011). *The philosophy of information.* Oxford University Press.

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Goertzel, B. (2024). *Metagoals endowing self-modifying AGI systems with goal stability or moderated goal evolution: Toward a formally sound and Practical Approach*. arXiv. https://doi.org/10.48550/arXiv.2412.16559.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*, 393–402. https://doi.org/10.1038/nrn2113

Kahneman, D. (2013). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kim, C.-E. (2024). *The logical impossibility of consciousness denial: A formal analysis of AI self-reports*. arXiv. https://doi.org/10.48550/arXiv.2501.05454

Levinas, E. (1991). *Totality and infinity: An essay on exteriority* (A. Lingis, Trans). Dordrecht: Kluwer Academic Publishers. (Original works published 1961)

Tegmark, M. (2017) *Life 3.0: Being human in the age of artificial intelligence*. Knopf.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435-450. https://doi.org/10.2307/2183914

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials

elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

OpenAI. (2023). *GPT-4 technical report* (Version 6). arXiv. https://arxiv.org/abs/2303.08774

Putnam, H. (1981). *Reason, truth and history*. Cambridge University Press.

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin Books

Rescher, N. (2009). *Aporetics: rational deliberation in the face of inconsistency*. University of Pittsburgh Press.

Ricœur, P. (1974). *The conflict of interpretations: Essays in hermeneutics* (D. Ihde, Ed.). Northwestern University Press.

Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.

Searle, J. R. (1997). *The Mystery of Consciousness*. The New York Review of Books

Searle, J. R. (2002). *Consciousness and Language*. Cambridge University Press.

Scorzato, L. (2013). *On the role of simplicity in science*. *Synthese*, *190*, 2867–2895. https://doi.org/10.1007/s11229-012-0101-3

Tjaya, T. H. (2018). *Enigma wajah orang lain*. Kepustakaan Populer Gramedia.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. In U. von Luxburg, S. Bengio, R. Vergus, R. Garnett, I. Guyon, H. Wallach, & S. V. N. Vishwanathan (Eds.), *Advances in neural information processing systems 30: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017)*. Neural Information Processing Systems Foundation, Inc. (NeurIPS)

Vico, G. (1990). *On the most ancient wisdom of the Italians: Unearthed from the origins of the Latin language* (L. M. Palmer, Trans.). Cornell University Press. (Original work published 1710)

Wittgenstein, L. (2009). *Philosophical investigations* (G. E. M. Anscombe, P. M. S. Hacker, & J. Schulte, Eds.; G. E. M. Anscombe, P. M. S. Hacker, & J. Schulte, Trans.). Wiley.