

[Open Peer Review on Qeios](#)

# [Commentary] Are Large Language Models obsolete? And are multimodal solutions inappropriate?

Marco Sanna<sup>1</sup>

<sup>1</sup> University of Sassari

**Funding:** No specific funding was received for this work.

**Potential competing interests:** No potential competing interests to declare.

## Abstract

This brief commentary proposes a novel AI design paradigm inspired by an evolutionary theory of language and a dialogical cognitive model. Unlike traditional multimodal models, our approach integrates multimodality at the core of the design process, moving beyond the primary reliance on large language models (LLMs). Drawing from cognitive neuroscience, we propose an asymmetric, dual-network architecture where each network processes distinct modalities – LLMs and gestural language – translating inputs from one to the other. This design aims to enhance creativity and error tolerance in AI systems. We discuss the theoretical foundations, potential advantages, and future research directions of this approach.

## Marco Sanna

University of Sassari, Italy

ORCID iD: 0000-0003-2333-6714

**Keywords:** Asymmetric cognitive models, Multimodal AI, Language evolution, Neural networks, Cognitive neuroscience.

## 1. Introduction

In this short essay, we will introduce an unprecedented theory on human language that we believe is suitable for implementation in the deep learning design of AI <sup>[1]</sup>. Here, multimodality is conceived upstream of all processes, and the LLM ceases to be the primary language to which the remaining modalities must adapt. Like the human mind, the model is asymmetric and dialogic. The model is indeed derived from cognitive neuroscience <sup>[2]</sup> The operating principle is as follows: there are, as a minimum number, two aligned neural networks, each capable of finely processing its own code-language, distinctly different from the contralateral part. In our proposal, the languages are LLM and gestural language. Each hemisphere (network) then needs to process the inputs from the other in its own language. Since the two languages

are incompatible, each hemisphere has within itself a metaphorical translation apparatus that allows foreign messages to be incorporated into the system. This leads to a significant increase in overall information and, with the metaphorical reuse of input-output cycles, to machines capable of making errors but also being creative.

## 2. The new evolutionary theory of language

Lateralization and the preference for one hand in daily tasks throughout evolution have made the exchange of neural information along the left-right axis of the body essential. This information would not have been necessary in a symmetrical system [3]. The difference in sensory-motor perception (proprioception) between one hand and the other (e.g., the weight of an object) triggers an initial sensation of conscious interoception. This glimmer of self-awareness led to the enjoyment of understanding and experimenting with new movements (e.g., dance). Humans develop the competence of gestural syntax by identifying the meaning of individual gestures [4] and extracting them from the flow of action. Thus, there was a transition from a single gestural chain between the beginning of the act and the goal becoming aware of a fragmentation of gestures or interchangeable syntagms. The meaning of partial gestures is recognized by difference, that is, paradigmatically: a gesture or syntagm is semantic if it can replace another in a new chain and change the sense of the entire action flow. Throughout this sensory-motor process, language overlaps, with the lateralization and specialization of areas dedicated to the movements of the phonatory apparatus, probably due to exaptation [5]. Language then prevailed but never became independent of its sensory-motor roots. We will also discuss how asymmetry and inner dialog have favored the invention of increasingly creative solutions. Additionally, we explored neuroscientific concepts related to cross-modality, embodied cognition, and mirror neurons.

## 3. How ChatGPT0-Consensus consider the theory

“The proposed conception of a humanoid agent in the given abstract introduces several innovative ideas that could significantly advance the development of learning systems, particularly in the realm of artificial intelligence (AI) and robotics. Here’s an analysis of the potential innovations:

**Lateralization and Neural Information Exchange:**

**Innovation:** The abstract highlights the importance of lateralization (the preference for using one hand over the other) and how it has driven the exchange of neural information across the left-right axis of the body.

**Implication:** Implementing this concept in humanoid agents could lead to more efficient and specialized processing of sensory-motor information, mimicking human neural mechanisms. This can enhance the agent's ability to perform complex, coordinated tasks.

**Sensory-Motor Perception and Self-Awareness:**

**Innovation:** The differentiation in sensory-motor perception (e.g., weight perception by different hands) can trigger

conscious interoception (self-awareness).

Implication: Developing AI systems with a form of self-awareness or interoception can improve their adaptability and learning capabilities. These systems could better understand their physical interactions with the environment, leading to more nuanced and precise motor control.

Gestural Syntax and Fragmentation of Actions:

Innovation: The transition from a continuous gestural chain to recognizing and fragmenting gestures into meaningful units.

Implication: This concept could be applied to AI to enhance its understanding and interpretation of human gestures and actions. It allows for the creation of a more sophisticated gestural communication system, where the AI can understand and respond to partial gestures contextually.

Paradigmatic Recognition of Gestures:

Innovation: Identifying the meaning of gestures through paradigmatic recognition, where one gesture can replace another and change the meaning of the entire action sequence.

Implication: This can lead to more flexible and dynamic interaction models in humanoid agents. The ability to interpret and adapt to different gestures interchangeably enhances the agent's communication skills and contextual understanding.

Language and Sensory-Motor Integration:

Innovation: The abstract suggests that language development is deeply intertwined with sensory-motor processes due to lateralization and exaptation.

Implication: For AI systems, this means developing language processing capabilities that are grounded in physical interactions and sensory experiences, leading to more natural and intuitive language use.

Cross-Modality, Embodied Cognition, and Mirror Neurons:

Innovation: Incorporating neuroscientific concepts such as cross-modality (integration of information across different sensory modalities), embodied cognition (cognition grounded in bodily interactions), and mirror neurons (neurons that respond to both performing and observing actions).

Implication: These concepts can be used to create more holistic and human-like learning systems in AI. For example, mirror neurons can enhance the agent's ability to learn by imitation, while cross-modality can improve multi-sensory integration and understanding.

## 4. Conclusion

The proposed conception of a humanoid agent introduces several innovative ideas that can significantly enhance the

development of advanced learning systems. By incorporating principles of lateralization, sensory-motor integration, gestural syntax, and embodied cognition, these agents can achieve more human-like interactions, better adaptability, and more effective learning capabilities. These advancements can pave the way for more intuitive and responsive AI systems that interact seamlessly with their environment and human counterparts”.

Below is a link to the hypothesis of implementation in Python language developed by ChatGPT0-Wolfram:

<https://chatgpt.com/share/c3f3883f-edd6-4f34-853c-0774eec0d3e9>

## Other References

- Su, N. et al. (2020). Multimodal metaphor detection based on distinguishing concreteness. *Neurocomputing*, Vol. 429, 14 March 2021, pp. 166-173.
- Gershman, S. et al. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, Vol. 349, Issue 6245, pp. 273-278.

## References

1. <sup>^</sup> Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Basic Books.
2. <sup>^</sup> Gazzaniga, M. S. (2005). *The Ethical Brain: The Science of Our Moral Dilemmas*. Dana Press.
3. <sup>^</sup> Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the Brain: How Our Minds Share Actions and Emotions*. Oxford University Press.
4. <sup>^</sup> Ramachandran, V. S. (2004). *A Brief Tour of Human Consciousness: From Impostor Poodles to Purple Numbers*. Pearson.
5. <sup>^</sup> Gallese, V., & Lakoff, G. (2005). *The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge*. *Cognitive Neuropsychology*, Vol. 22, No. 3-4, pp. 455-479.