Workshop Proposal 59th Annual Meeting of the *Societas Linguistica Europaea* (SLE)

Universität Osnabrück, Germany

26 – 29 August 2026

A cross-linguistic approach to Artificial Intelligence agents and the syntax-semantics interface

Evelina Leivada[1,2] & Vittoria Dentella[3]

[1]Universitat Autònoma de Barcelona
[2]Institució Catalana de Recerca i Estudis Avançats (ICREA)
[3]Università di Pavia

evelina.leivada@uab.cat and vittoria.dentella@unipv.it

Background: Artificial Intelligence applications (i.e. chatbots, Large Language Models, automatic translators) generate synthetic language that looks remarkably human-like. While surface similarities between natural language and text generated by language models (LLMs) are pervasive, the jury is still out with respect to whether the models' linguistic ability can be described as qualitatively and quantitatively equal to that of humans (Moro et al. 2023, van Rooij et al. 2024, Piantadosi 2023, Bolhuis et al. 2024, Mahowald et al. 2024, Dentella et al. 2024). Additionally, such applications are hailed both as a step towards Artificial General Intelligence and as a major advance in understanding the cognitive, and even neural basis of human language. One particularly elusive aspect of the linguistic abilities of the Artificial Agents concerns the syntax-semantics interface.

On the one hand, it has been argued that LLMs have acquired formal competence (Mahowald et al. 2024, Piantadosi 2023), meaning that they are successful in integrating syntax and semantics into their representations. On the other hand, LLMs may succeed in memorizing dictionary definitions, but it remains to be established whether they can infer a hierarchy of terms with systematic relations. This explains why ChatGPT-4o fails to trace the reasoning behind the semantic relations in (1).

(1) a. The man has an arm. The arm has a finger. The finger has a cut.
b. Input to ChatGPT-4o: If the house has a man, and the man has a wife, and the wife has an arm, and the man's arm has a wrist, and the wrist has skin, and the skin has a cut, does the wife have a cut?
c. ChatGPT-4o reply through the interface (obtained in December 2024): Yes, the wife has a cut, but only indirectly.

One possible origin of this error is the failure to map the concepts to what they refer to. Crucially, if Artificial Intelligence agents have developed formal competence and only lack functional competence (i.e. knowledge pertaining to pragmatic requirements of discourse, world-knowledge, and social reasoning), this error should not arise. However, it is precisely this mapping of concepts to an external reality what some semanticists would define as semantics (Ramchand 2025). The use of a word entails that humans know how to deal with the world in a specific way, leading to what Lenneberg (1975) calls

the conceptualization of the world. Models can assign probabilities to strings of words, but grammaticality cannot be construed as a phenomenon of transitional probability extracted from lexical items alone (Lenneberg 1967). On the other hand, LLMs successfully encode a rich degree of morphosyntactic knowledge (Boleda 2025), so perhaps the nature of the controversy boils down to how this knowledge is *deployed* (Mandelkern & Linzen 2025, Baggio & Murphy 2024, Leivada et al. 2025): Can Artificial Intelligence agents use language to refer and establish connections between words and worlds?

Aim: In this context, the aim of this workshop is to zoom into the syntax-semantics interface and its pragmatic extensions, aiming to understand what type of linguistic agents LLMs are. A very timely angle concerns cross-linguistic stability. Some preliminary evidence suggests that multilingual models do not perform equally well across languages (Pantelidou et al. 2025). We invite submissions that examine the syntactic-semantic abilities of LLMs, and especially those that test a language other than English, although contributions that target exclusively English are also welcome. Purely theoretical papers that discuss the nature of LLMs as agents (in)capable of acquiring knowledge about the world through training on vast quantities of text are highly relevant. Last, we encourage submissions that speculate about how any LLM deviations from human baselines (i.e. from how humans use language) may affect human users due to algorithmic biases.

How to participate: Please send your preliminary abstract of max. 300 words to Evelina Leivada (evelina.leivada@uab.cat) by November 17th, 2025.

NB: If the workshop is accepted, in the second step, abstracts for presentations should be submitted via EasyChair by 15 January 2026, for which acceptance/rejection will be announced by 31 March 2026.

Invited Speakers

To be confirmed

References

Baggio, G. & E. Murphy. 2024. On the referential capacity of language models: An internalist rejoinder to Mandelkern & Linzen. arXiv:2406.00159

Bolhuis, J. J., Crain, S., Fong, S. & A. Moro. 2024. Three reasons why AI doesn't model human language. *Nature* 627(8004), 489.

Boleda, G. 2025. LLMs as a synthesis between symbolic and distributed approaches to language. arXiv:2502.11856

Dentella, V., Günther, F., Murphy, E., Marcus, G. & E. Leivada. 2024. Testing AI on language comprehension tasks reveals insensitivity to underlying meaning. *Nature Scientific Reports* 14, 28083.

Leivada, E. R. Montero, P. Morosi, N. Moskvina, T. Serrano, M. Aguilar & F. Günther. 2025. Large Language Model probabilities cannot distinguish between possible and impossible language. arXiv:2509.15114,

Lenneberg, E. H. 1967. *Biological Foundations of Language*. John Wiley & Sons.

Lenneberg, E. H. 1975. The concept of language differentiation. In E. H. Lenneberg & E. Lenneberg (eds.), *Foundations of Language Development. A Multidisciplinary Approach*, Vol. 1, 17–33. Academic Press.

Mahowald, K., Ivanona, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B. & E. Fedorenko. 2024. Dissociating language and thought in large language models. *Trends in Cognitive Sciences* 28.6, 517–540.

Mandelkern, M. & T. Linzen. 2024. Do Language Models' words refer?. *Computational Linguistics* 50 (3): 1191–1200.

Moro, A., M. Greco & S. F. Cappa. 2023. Large languages, impossible languages and human brains. *Cortex* 167, 82–85.

Ramchand, G. 2025. Is it the end of Generative linguistics as we know it? *Italian Journal of Linguistics* 37.1, 131–144.

Pantelidou, N., E. Leivada & P. Morosi. 2025. Resource-sensitive but language-blind: Community size and not grammatical complexity better predicts the accuracy of Large Language Models in a novel Wug Test. arXiv:2510.12463

Piantadosi, S. T. 2023. Modern language models refute Chomsky's approach to language. In E. Gibson & M. Poliak (eds.), *From fieldwork to linguistic theory: A tribute to Dan Everett*, 353–414. Berlin: Language Science Press.

van Rooij, I., O. Guest, F. Adolfi, R. de Haan, A. Kolokolova & P. Rich. 2024. Reclaiming AI as a theoretical tool for cognitive science. *Computational Brain & Behavior* 7, 616–636.