

Minds

Big questions for linguistics in the age of AI

Ad Backus,⁴ Michael Cohen,⁴ Neil Cohn,⁴ Myrthe Faber,⁴
Emiel Krahmer,⁴ Schuyler Laparle,⁴ Emar Maier,¹
Emiel van Miltenburg,⁴ Floris Roelofsen,² Eleonora Sciubba,⁴
Merel Scholman,³ Dimitar Shterionov,⁴ Maureen Sie,⁴
Frédéric Tomas,⁴ Eva Vanmassenhove,⁴ Noortje Venhuizen⁴
& Connie de Vos⁴

¹ University of Groningen | ² University of Amsterdam | ³ Utrecht University | ⁴ Tilburg University

It is an interesting time to be a linguist. The advent of large language models (LLMs) like GPT-4 has raised fundamental questions about language and its nature, such as whether artificial systems are able to “use” language in a similar way to humans. While such issues are at the core of ongoing scientific and societal debates, the role of linguistics in the development of these technologies has been surprisingly limited. LLMs are suddenly widespread because of the opportunities they provide, but also come with several risks. Schools are therefore regulating their use, and some countries—like Italy—even prohibited them. LLMs can be wrong, for example, in the sense that they can “hallucinate” and come up with structurally correct, but false statements (Ji et al. 2023). This behavior is due to the way in which these models learn to use language; not by assigning meaning to form but rather by learning statistical regularities about how words and sentences typically co-occur. As a result, the ‘meaning’ that these models assign to language is generally not grounded in any experience with the world.¹

In other words, LLMs can be described as ‘stochastic parrots’ that semirandomly reproduce linguistic forms without actually being related to the outside world (Bender et al. 2021).² We believe linguists can pick up a much larger role in these discussions clarifying how LLMs could be adapted to become more similar

1. But see systems such as DALL-E <https://openai.com/product/dall-e-2>.

2. But see Video symposium: Debate: Do Language Models Need Sensory Grounding for Meaning and Understanding? <https://www.youtube.com/watch?v=x1o964woozk>

to the language we use as humans. We do not reflect on the philosophical question of what language essentially is, though such reflections play a role, but rather identify 5 elements LLMs fall short of. More specifically, we argue that linguistic models and representations should centralize MINDS: Multimodality, Interoperability, Nonopacity, Diversity, and Sociality. We argue that these aspects of human language constitute the main challenges to linguistics as a social science and that elucidating them would require a concerted effort from the field itself, but also from affiliated domains such as philosophy, anthropology, sociology, and psychology.

Multimodality

There is a long-standing bias in linguistic theory toward the written mode, one that continues to be reflected in the development of LLMs to some extent. More recently, however, we have pushed our theories of language toward embodiment and multimodality. This reframes face-to-face interaction as the canonical form of communication, rather than as the messy result of language production in real life. If this reframing is taken seriously by the field, multimodality becomes central to the way we think and theorize about language. As communicators interact, they take advantage of every semiotic mode available to them (Ferrara & Hodge 2018). This includes what we typically think of as ‘conventionalized’ systems, like sign, speech, and graphics, but also contextually-bound systems like gesture, bodily orientation, and reference to the real space in which the interaction occurs (Goodwin 2000; Perniss 2018). Our theories must not only accommodate multimodality, but be built around it (Dingemanse et al. 2023; Cohn & Schilperoord 2022).

For LLMs in particular, the turn toward multimodality impacts the minimal requirements for data input and output (Chrupała 2022). Crucially, it must create output that is not only ‘realistic’ in a general sense, but is also situated, that is to say grounded in a specific multimodal and interactional context. This could potentially be achieved through reinforcement learning with feedback from human Interactants (Corti & Gillespie 2016).

Interoperability

The success of LLMs has challenged many traditional assumptions about language, including the idea that language can only be studied by discerning its different components/levels of analysis (modularity; Jackendoff 2002). LLMs take as

input large amounts of (written) text and use statistical modeling techniques to develop abstractions over this input, which can be considered a type of ‘meaning’ representation (in the sense of the distributional hypothesis, meaning as use; Landauer & Dumais 1997). Critically, however, these systems seem to eschew the need for different, modular levels of linguistic analysis that have traditionally been part and parcel of linguistics, such as the trichotomy between syntax, semantics and pragmatics (Contreras Kallens et al. 2023).

A similar trend can be observed in cognitive and neuroscientific research that has challenged the idea that these levels are implemented as completely separate modules. More likely they interact and are co-dependent (Elman 2009). This suggests that there is a need for multimodal and interoperable representations that allow for the exchange between and integration of overlapping sources of information. These representations also need to be interoperable outside of the symbolic domain—that is, they need to allow for understanding the mapping of forms to the real world. And this is exactly what is missing from the “meaning” representations that LLMs construct: they are void of many of the aspects that linguists have considered to be crucial for models of understanding, such as grounding in the world, the body (embodiment), and interpersonal communication (Bender & Koller 2020). A big question in linguistics therefore lies in addressing how we can define models and representations in which the traditional levels of linguistic analysis interact and combine in attributing meaning to the linguistic signal, both in cognitive architectures and in LLMs.

Non-opacity

Opacity threatens the interpretability of LLMs as well as models of meaning. Interpreting LLMs is a complex case for multiple reasons. Publicly available conversational agents do not share openly the grounds on which they have been built, leading to a lack of transparency and accountability of the biases at play. Similarly, AI is prone to the black-box problem, as the algorithmic path taken between an input and an output is generally inscrutable. This ethical matter is important, as the use of artificial intelligence has been suggested as an advantageous course of action if one considers the correct guesses compared to the negative guesses for marketing purposes or medical diagnostics, for instance (Castelvecchi 2016; Holm 2019). However, more recently, an entire research field has begun to investigate the use of explainable AI, intending to reveal the inner workings of the models used for decision making and classification purposes.³ The content produced by

3. See e.g. <https://blackboxnlp.github.io/>

LLMs can be seen as the result of a probabilistic system in which the next word in a text is selected based on the context in which it occurs.⁴ As such, LLMs are aligning words according to expected patterns, but do not deal well with infrequent language phenomena, mathematical queries put into words, and logical fallacies (Mahowald et al. 2023). What is needed is a differentiated perspective of “understanding” in human and AI systems.

Diversity

Because they are built on “our” language, LLMs reflect the same biases or even increase them, because they (over)generalize biased input. Regarding linguistics in particular, the abovementioned stochastic biases imply that the lexical subtleties and overall diversity found in language generated by these models will be poor. Concerns have been raised about how human biases present in the input of a model can lead to direct, indirect, or intersectional discrimination (Abid et al. 2021; Venkit et al. 2023), and engender AI biases after the learning phase (Wirtz et al. 2019). Such discrimination is visible in the development of sexist stereotypes when AI is trained on uncontrolled open access data, where men are described as maestros and women as homemakers (Caliskan et al. 2017). Similarly, LLMs are persistently based on majority languages and ignore the vast variety of linguistic systems around the world (Bender 2011). For this reason, they may increase existing inequalities engendered by the technological advantages they bring to those who use majority languages and those who do not. In an ideal world, these technologies would rather help bridge such communicative gaps to support inclusivity.

Sociality

Above and beyond building better LLMs, a generalized theory of human behavior, and of language in particular, needs to take into account our sociality (Hu et al. 2022). This includes the fact that meaning arises from situated conversations between persons who share common ground to varying degrees and also actively seek to increase it through their interactions. Their communicative success in doing so is affected by their shared communicative histories, metalinguistic awareness, as well as the deictic affordances provided by the physical space they share. In addition to the practical implications, we should discuss ethically

4. See e.g. <https://twitter.com/tallinzen/status/1643961153087758343>

whether we can implement these technologies in sensitive contexts such as health-care. To this end, these systems would need to be geared toward constructing intersubjectivity with individual patients rather than sending de-contextualised information (Verhagen 2008).

To conclude, the advance of LLMs challenges us to rethink what language is and capitalise on this process to redefine what the characteristics of a general linguistic theory should be; how intense the cooperation with other disciplines should be; and, figure out to what extent they are different disciplines. Huge opportunities await if we can unify the subdisciplines and integrate them into a new type of linguistics: theories with wider reach and greater empirical support, better integration in the wider field of Humanities and Social Sciences, and the development of LLMs and other new technologies. This is because many linguistic behaviors, such as individual and group differences in language use and change, language acquisition, psychology of habits, identity construction and group formation, all support the notion that experience matters, that there is a lot of statistical learning, and that people are not robots or computers (Baria & Cross 2021). What is needed is research that directly addresses the link between behavior and representation in both directions, with an open call to colleagues in other disciplines to join us in interdisciplinary projects as well as to each other (in our hyperspecialized linguistics corners), and with full realization of the contemporary world.

Funding

Open Access publication of this article was funded through a Transformative Agreement with Tilburg University.

LOT Essay Prize



This essay won shared second prize in the 2023 essay competition sponsored by LOT.

Acknowledgements

A first draft of this essay was written by an interdisciplinary mix of members of all departments in the Tilburg School for Humanities and Digital Sciences, including the Departments of Communication and Cognition, Culture Studies, Cognitive Science and Artificial Intelligence and Philosophy coordinated by Dr. Connie de Vos.

References

- doi Abid, Abubakar, Maheen Farooqi & James Zou. 2021. "Persistent Anti-Muslim Bias in Large Language Models". *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21)*, 298–306.
- Baria, A. T., & Keith Cross. 2021. "The brain is a computer is a brain: neuroscience's internal debate and the social significance of the Computational Metaphor". arXiv preprint arXiv:2107.14042.
- doi Bender, Emily M. 2011. "On achieving and evaluating language-independence in NLP". *Linguistic Issues in Language Technology* 6.
- doi Bender, Emily M., Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 610–623.
- doi Bender, Emily M., & Alexander Koller. 2020. "Climbing towards NLU: On meaning, form, and understanding in the age of data". *Proceedings of the 58th annual meeting of the association for computational linguistics*, 5185–5198.
- doi Caliskan, Aylin, Joanna J. Bryson & Arvind Narayanan. 2017. "Semantics Derived Automatically from Language Corpora Contain Human-like Biases". *Science* 356: 183–186.
- doi Castelvechi, Davide. 2016. "Can we open the black box of AI?". *Nature News*, 538(7623), 20.
- doi Chrupała, Grzegorz. 2022. "Visually grounded models of spoken language: A survey of datasets, architectures and evaluation techniques". *Journal of Artificial Intelligence Research* 73: 673–707.
- doi Cohn, Neil & Joost Schilperoord. 2022. "Reimagining language". *Cognitive Science* 46(7): 1–6, e13164.
- doi Contreras Kallens, Pablo, Ross Deans Kristensen-McLachlan & Morten H. Christiansen. 2023. "Large language models demonstrate the potential of statistical learning in language". *Cognitive Science* 47(3), e13256.
- doi Corti, Kevin, & Alex Gillespie. 2016. "Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human". *Computers in Human Behavior* 58: 431–442.
- doi Dingemans, Mark, Andreas Liesenfeld, Marlou Rasenberg, Saul Albert, Felix K. Ameka, Abeba Birhane, Dimitris Bolis, Justine Cassell, Rebecca Clift, Elena Cuffari, Hanne De Jaeger, Catarina Dutilh Novaes, N. J. Enfield, Riccardo Fusaroli, Eleni Gregoromichelaki, Edwin Hutchins, Ivana Konvalinka, Damian Milton, Joanna Rączaszek-Leonardi, ... Martina Wiltschko. 2023. "Beyond single-mindedness: A figure-ground reversal for the cognitive sciences". *Cognitive Science* 47(1): e13230.
- doi Elman, Jeffrey L. 2009. "On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon". *Cognitive Science* 33(4): 547–582.
- doi Ferrara, Lindsay, & Gabrielle Hodge. 2018. "Language as description, indication, and depiction". *Frontiers in Psychology* 9: 716.
- doi Goodwin, Charles. 2000. "Action and embodiment within situated human interaction". *Journal of Pragmatics* 32(10): 1489–1522.
- doi Holm, Elizabeth A. 2019. "In defense of the black box". *Science* 364(6435): 26–27.

- Hu, Jennifer, Sammy Floyd, Olessia Jouravlev, Evelina Fedorenko, & Edward Gibson. 2022. “A fine-grained comparison of pragmatic language understanding in humans and language models”. *arXiv:2212.06801*.
-  Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
-  Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Yejin Bang, Wenliang Dai, Andrea Madotto, & Pascale Fung. 2023. “Survey of hallucination in natural language generation”. *ACM Computing Surveys* 55(12): 1–38.
-  Landauer, Thomas K., & Susan T. Dumais. 1997. “A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge”. *Psychological Review* 104(2): 211–240.
- Mahowald, Kyle, Anna A. Ivanova, Idan A. Blank, Nancy Kanwisher, Joshua B. Tenenbaum, & Evelina Fedorenko. 2023. “Dissociating language and thought in large language models: a cognitive perspective”. *arXiv preprint arXiv:2301.06627*.
-  Perniss, Pamela. 2018. “Why we should study multimodal language”. *Frontiers in Psychology* 9: 1109.
- Venkit, Pranav Narayanan, Sanjana Gautam, Ruchi Panchanadikar, Ting-Hao Huang, & Shomir Wilson. 2023. “Nationality Bias in Text Generation”. *arXiv preprint arXiv:2302.02463*.
-  Verhagen, Arie. 2008. “Intersubjectivity and the architecture of the language system”. In *The shared mind: Perspectives on intersubjectivity*, ed. by Jordan Zlatev, Timothy P. Racine, Chris Sinha & Esa Itkonen, 307–331. Amsterdam/Philadelphia: John Benjamins.
-  Wirtz, Bernd W., Jan C. Weyerer, & Carolin Geyer. 2019. “Artificial intelligence and the public sector – applications and challenges”. *International Journal of Public Administration* 42(7), 596–615.

Address for correspondence

Ad Backus
 Tilburg School of Humanities and Digital Sciences
 Postbus 90153
 5000 LE Tilburg
 The Netherlands
 a.m.backus@tilburguniversity.edu

Co-author information

Michael Cohen
Tilburg School of Humanities and Digital
Sciences

M.Cohen@tilburguniversity.edu

Neil Cohn
Tilburg School of Humanities and Digital
Sciences

N.Cohn@tilburguniversity.edu

Myrthe Faber
Tilburg School of Humanities and Digital
Sciences

M.Faber@tilburguniversity.edu

Emiel Krahrmer
Tilburg School of Humanities and Digital
Sciences

E.J.Krahrmer@tilburguniversity.edu

Schuyler Laparle
Tilburg School of Humanities and Digital
Sciences

S.M.Laparle@tilburguniversity.edu

Emar Maier
University of Groningen
Faculty of Philosophy

e.maier@rug.nl

Emiel van Miltenburg
Tilburg School of Humanities and Digital
Sciences

C.W.J.vanMiltenburg@tilburguniversity.edu

Floris Roelofsen
University of Amsterdam
Faculty of Science, ILLC

f.roelofsen@uva.nl

Eleonora Sciubba
Tilburg School of Humanities and Digital
Sciences

M.E.Sciubba@tilburguniversity.edu

Merel Scholman
Utrecht University
Faculty of Humanities

m.c.j.scholman@uu.nl

Dimitar Shterionov
Tilburg School of Humanities and Digital
Sciences

D.Shterionov@tilburguniversity.edu

Maureen Sie
Tilburg School of Humanities and Digital
Sciences

M.M.S.K.Sie@tilburguniversity.edu

Frédéric Tomas
Tilburg School of Humanities and Digital
Sciences

F.J.Y.Tomas@tilburguniversity.edu

Eva Vanmassenhove
Tilburg School of Humanities and Digital
Sciences

E.O.J.Vanmassenhove@tilburguniversity.edu

Noortje Venhuizen
Tilburg School of Humanities and Digital
Sciences

N.J.Venhuizen@tilburguniversity.edu

Connie de Vos
Tilburg School of Humanities and Digital
Sciences

C.L.G.deVos@tilburguniversity.edu

Publication history

Date received: 1 May 2023

Date accepted: 15 May 2023