JRAI Journal of the Royal Anthropological Institute



# Introduction: Towards a linguistic anthropology of AI

Webb Keane *University of Michigan*

Constantine V. Nakassis *University of Chicago*

This essay introduces the themed cluster of articles, 'Towards a linguistic anthropology of AI'. The advent of artificial intelligence (AI), especially in large language models capable of producing coherent discourse mimicking conversational interaction, is exerting unprecedented pressure on prevailing concepts of language, personhood, and the human. This provides a challenge and opportunity for reflection, particularly for linguistic anthropology, that sub-field centred on language as medium/practice of meaning-making in social life. Moving beyond AI as communicating machine to explore communicative processes that structure AI as ideological imaginary and technological infrastructure, the articles in this cluster draw on linguistic anthropology to examine different facets of AI as people take it up, such as the features that invite users to endow chatbots with personhood (Keane), attempts to use AI to unlock animal communication (Handman), and research seeking linguistic biomarkers predictive of psychopathology (Semel). To contextualize the articles, this introduction offers some basic semiotic distinctions developed within linguistic anthropology needed to parse AI, with a focus on the concept of semiotic ideology. We then situate AI within a set of ideological genealogies – from Enlightenment visions of universal languages to the representationalism of modern linguistics – before turning to themes of transduction and semiosis beyond the human.

## Introduction

The advent of artificial intelligence (AI)-driven programs capable of producing coherent discourse is exerting unprecedented pressure on many prevailing concepts of language, and by extension, of personhood, and even of those features commonly taken to distinguish the human. As we write, anyone with computer power in hand has access to chatbots that seem able to pass Alan Turing's famous 'Imitation Game', the so-called Turing Test. AI-powered devices can respond to their users' prompts in ways that are often jarringly indistinguishable from (non-face-to-face) human conversation.

*Journal of the Royal Anthropological Institute (N.S.)* **00**, *1-16*

Everywhere one turns, one is seemingly confronted by the transformations – present, future, actual, imagined – being brought about by AI.

But not everything is as novel as might seem when first encountered. The apparent 'newness' of new technology itself requires critical scrutiny (Gershon 2010). One of the virtues of the anthropological approach is an ability to draw on an ethnographic breadth and historical depth that can provide this needed scrutiny. Given the extraordinary, even utopian, hype and equally fervent moral panic surrounding developments in AI, an anthropological perspective offers some sobering, if not always reassuring, recontextualization and recalibration. The ethnographic sensibility, we argue, works to resituate such hype and panic within the actual practices, material conditions, social formations, and histories that make AI and related technologies possible and consequential. It is here that we can separate the hype from the realities of how these new technologies are transforming the worlds in which they are being introduced, as well as the material effects of such hype and panic.

Linguistic anthropology is especially well situated to do this work, precisely because many of the advances in AI technologies that have most captured the public imagination have turned on social interaction by means of written or spoken words. Indeed, the Turing Test is, above all, a test of the ability to engage in conversation – a semiotic skill. By *semiotic*, we refer to the processes by which people make meaning, communicate with one another, as well as with non-human animals and other entities, and interpret and make sense of the world around them by means of perceptible signs. Semiotic processes include not just speech and art, but also, for instance, the reading of medical symptoms, forensic inferences, and divination. It is AI's seeming mastery of aspects of this semiotic skill, especially in producing effective discourse in social interactions, that most astonishes the ordinary user and most effectively fosters the illusion that there is 'someone there' in their devices (Keane 2024; Kohavi & Weichselbraun 2026).

But linguistic anthropological approaches do not just aim to explain how AI technologies serve as participants within social interactions. A focus on communicative processes, broadly construed, is also critical for situating those interactions within the semiotic and ideological worlds in which they unfold, are imagined to unfold, and are designed to unfold. That is, a linguistic anthropology of AI is focused not only on the linguistic capacities of AI technologies but also on the semiotic processes that structure those very technologies and the interactions and social and political contexts in which they are embedded.

With this in mind, we have assembled this cluster of articles to bring linguistic anthropological approaches to bear on some recent developments in AI-related technology.[1] The aim is not to be exhaustive, of course. For example, not all AI technologies involve naturalistic language-focused interaction with human users (see, e.g., D'Armenio, Dondero, Deliège & Sarti 2025 and Dondero 2025 on image GenAI; Wilf 2023 on music and AI; also see Ilana Gershon's commentary in this issue). Rather, our purpose is to highlight certain general semiotic and ethnographic issues that these technologies raise. It is to ask how language-focused AI technologies open up new critical vistas on classic anthropological and linguistic questions, such as what is language, what is it to be a speaker, what is a person, and what defines the human – and the non-human? To get at these questions, the authors in this cluster work at the intersection of concrete practices involving AI, the concepts they presuppose, and the larger claims to which they give rise.

Importantly, each contributor is engaged in what Beth Semel, in her article in this issue, calls a 'repatriation project'. By this she refers to linguistic anthropology's long struggle against the tendency in mainstream linguistics and cognitive science to treat language as an autonomous, self-contained system, situated within the individual mind, often identified with the brain, that can be understood completely apart from the material and social domains within which it is used and of which it speaks. This repatriation project aims to bring language back into the social world of its users, not as an add-on, but as one of its fundamental and irreducible dimensions. Such a repatriation is all the more needed in an era of AI and in its study.

## Language beyond linguistics

The articles in this thematic cluster draw on certain key concepts from semiotically informed anthropological approaches to language. Such approaches are based on a model of the sign that goes beyond the dominant conceptions of language underlying modern linguistics that are rooted in the traditions of Ferdinand de Saussure, Gottlob Frege, and Noam Chomsky. What those traditions share is a representationalist view of language. Such a view assumes that language is, in essence, a decontextualized 'system' of arbitrary signs, a 'code'; and further, that the main function of that code is to produce propositions that refer to real or imagined states of affairs in the world and make statements about them. On this representationalist view, the purpose of words is to let us utter sentences like, 'The cat is on the mat'. This view has little to say about language as action (things like insults, commands, allusions, prayers, or jokes), nor about speech registers (like slang or legalese), discourse genres (like lectures or tweets), nor other ways in which language marks social identities (like regionalisms or accents). Although the representationalist view of language has been criticized by many, from Bronisław Malinowski to Ludwig Wittgenstein, J.L. Austin to Dell Hymes, it still holds sway in much of the thinking behind AI's design and public reception.

By contrast to this representationalist view, linguistic anthropologists have long focused on meaning not as a reflection of the world but rather as something that, by being a part of it, shapes and affects it. This view sees meaning not as the result of grammatical combinations of elements from a static system, but as emergent from a process or activity of using, negotiating, and circulating signs. That is, it sees meaning as constructed within socially embedded and interactionally contextualized processes, what is sometimes called *pragmatics*, in contrast to codified systems of conventional linguistic rules described by syntax or semantics. To get at how this works, linguistic anthropologists have drawn, selectively, from the semiotics of the American philosopher and logician Charles Sanders Peirce. This approach stresses three crucial features of sign processes.

First, as formulated by Peirce (1956), a sign always has a *material component*, whether that be sound, shape, colour, or any other perceptible quality of the 'sign vehicle'. For language this immediately situates it in a world of physical affordances and constraints, such as the biomarkers of depression that the researchers in Semel's contribution to this issue are seeking in the acoustic subtleties of speakers' utterances. It means that semiosis is never a purely 'internal' cognitive phenomenon but always has a 'public' dimension.

Second, while recognizing the special role that 'arbitrary' signs – what semiotics calls *symbols*, defined by a linguistic rule, play in certain systems of signification like language (as Saussurean structuralism stresses), those material properties also afford *non-arbitrary* modes of signifying, which structuralism largely overlooks. These modes

are *iconicity* (the ways a sign is connected to its object by resemblance) and *indexicality* (the ways a sign is connected to its object by causal or proximate connection). Iconicity is dominant in signs like artworks like paintings that bear, or stipulate, some likeness to their objects, but also in poetry (where sound mimics sense), diagrams (which schematize the relations of parts within a model), and metaphor (that draw resemblance at abstract levels). Indexicality involves a dynamic connection of causality, co-presence, or contiguity between the sign and its object. Examples include the smoke that indicates fire, an index finger pointing to some object in its surroundings, and a regional accent revealing the speaker's place of origin. Indexical signs are highly context-sensitive. Both iconicity and indexicality depend on particular frames, or 'ideologies', as we discuss below.

In any empirical instance, a sign is likely to draw on both iconicity and indexicality in co-ordination. Consider the passport. It authorizes the bearer's passage across borders by forging an indexical link between document and bearer. It does so by virtue of the photograph's iconic resemblance to the bearer. That iconicity is guaranteed indexically, that is, the photo resembles the photographic subject because the image was presumably caused by an initial act of exposure of a camera lens to the person in question. If the passport draws on the affordances of the photograph, it still requires further authorization in the signature, itself indexical of the physical presence of the signatory at the moment of signing, and the conventional symbols – above all, words – that frame photo and signature under the authority of the state.

Even within purely linguistic discourse, iconicity and indexicality are active. Consider, as an example, American presidential candidate Dwight Eisenhower's memorable campaign slogan, 'I like Ike'. As famously analysed by Roman Jakobson (1960), the components (*I*, *like*, *Ike*) of this highly condensed poetic structure resemble and iconically echo each other through rhyme and alliteration. Simultaneously, they 'point to' (index) the imputed speaker of the utterance via the personal pronoun *I* – and often as well as by that speaker displaying those words on a political pin literally attached over their heart. The act (*liking*) and actor (*I*) are anchored both in the immediacy indexed by the present tense of the verb, and through their iconic replication in the very object of their affection (*Ike*).[2] Additionally, the use of the nickname 'Ike' points to the imagined closeness of the affectionate subject to the person referred to (Eisenhower).

A corollary of the relatively non-arbitrary dimension of semiosis is that linguistic signs should be understood as highly distinctive components of a wider, multi-dimensional field of semiotic processes, including iconic and indexical features (in our example above, pins apprehended visually, bodies, political messaging, signatures and seals, photographs, names). This is as true for 'natural' language as for discourse produced by 'artificial' intelligence. Indeed, what is truly revolutionary in AI-produced discourse is less its 'arbitrary' or symbolic component, that which seems most governed by linguistic rules, than the iconic and indexical relations among the swathes of text it produces.[3] At its most rudimentary, large language model (LLM) training selects most likely next words on the basis of their 'fit' with the given stretch of discourse. The selection 'fits' because the result would resemble (be iconic of) already-known, similar conjunctions of words, which thereby provide the textual context that indexes where the selected word should go. As a result, its output feels coherent: any given stretch of text iconically 'fits' with the other signs that it is indexically juxtaposed with, what linguistic anthropologists call its 'co-text' (Silverstein & Urban 1996). It is, in effect, an *indexical icon* of text produced by the human intentions that generated the

source texts that LLMs draw on. When ChatGPT produces the words 'I want to be free' and 'I'm in love with you', as it did for one journalist (Roose 2023), the text is uncanny not only because of its meaningful 'content'. What Roose finds unsettling is the way it is iconic of (it mimics) words spoken by a human being while seeming to point to (index) an individual speaker ('I') when he knows that there is no one there.

Any understanding of AI devices requires attending to all three dimensions of the sign process (iconic, indexical, symbolic). It is the co-ordination among all these dimensions of semiosis that brings signs into the world which their users inhabit, experience, know, and imagine.

The third feature of the sign is that in order to be a sign, it must always signify something *to someone*. That is, the fundamentally social and perspectival, which is also to say ideological, character of signification cannot in principle be excluded from analysis. As Keane's contribution in this issue suggests, this feature calls for attention to the interactive and reflexive dynamics in play in any given stretch of discourse.

This ideological component is critical because it shapes what people will or will not consider to be significant and why. For this reason, linguistic anthropologists attend to *language ideologies* (Gal & Irvine 2019; Schieffelin, Woolard & Kroskrity 1998; Silverstein 1979), or, expanding the concept to signs beyond language as strictly understood, *semiotic ideologies* (Keane 2018). These concepts refer to the role played by sign users' underlying, mostly tacit, understandings of how signs function. In this usage, 'ideology' importantly does not mean 'false consciousness' nor does it have a pejorative sense or imply that there is some objective analytic position 'outside' of ideology (Leone, Keane & Nakassis 2025). Rather, it designates how meaning is a function of users' conceptions about signs and the worlds they traffic with.

Ideology, in this sense, points to the fundamentally reflexive character of human sign use – that a sign functions only when it is taken *to be* a sign by someone. This is no idle question, because what people assume signs can or cannot do, what aspects of reality they do or do not express, and in what ways they fit with each other or not (e.g., belonging to the right or appropriate genre or speech register given the context) profoundly shape the way they are used, interpreted, and indeed, even fought over. As Handman points out in her article in this issue, whether you think that the mutually intelligible languages now called 'Bosnian' and 'Serbian' do or do not require translation carries an entire history of civil war into everyday interactions. In like manner, during the Protestant Reformation, disagreements about whether the Eucharist *did* anything or merely *symbolized* something could become deadly (Keane 2007). These distinctions turn on differences among semiotic ideologies. As the contributors to this cluster show, similar questions are implicated in recent AI technologies.

Semiotic ideology, thus, plays a key analytic role in all the articles of this cluster. Consider Beth Semel's article in this issue, which looks at laboratory researchers seeking acoustical markers of clinical depression as part of a project to create AI technologies to screen for depression. For their purposes, the relevant features of speech are not found in the systems of morphology, semantics, syntax, or even phonology that regiment a given language. Attuned to dimensions of speech that are not fully codified by a given linguistic structure, they seek indexes of medical trouble, signs that are associated with, and thus point to, depression (presumably) through causal chains, perhaps grounded in the limbic system of the brain. Identifying such biomarkers takes enormous labour on the part of scientists, which is itself guided by how they navigate, and put into

play, various ideologies about language, the mind, and much more. As Semel shows, researchers confront in this work a tension between the representationalist view of language that tends to dominate the Western imaginary and a materialist model called for by their everyday research practices. Indeed, as the researchers came to realize through their research, contrary to the famous diagram in Saussure's lectures of two heads connected by arrows meant to depict speech, verbal interaction does not consist of one individual encoding an inner thought and sending it over to another individual to decode into their own inner thought. It always involves social processes that co-construct meaning. Efforts to isolate biomarkers run into the confounding problem of other people in events of social interaction: the research subjects in the study Semel writes about do not always play along with the research protocols, even when they want to oblige (and some resist even that).

The experimental protocol itself has roots in semiotic ideologies. For example, some of the men in the research project Semel discusses appeared to be uncomfortable with speaking in an unusually high pitch (part of the experiment meant to elicit particular markers), perhaps because it was threatening to their masculinity. If high pitch is an indexical icon of femininity – that is, both resembling and, purportedly, produced by female bodies and by extension, an otherwise imperceptible inner feminine character – this semiotic connection depends on ideological support. This required the researchers to undertake the delicate work of trying to get their research subjects to *disconnect* high pitch from 'femininity' in order to ensure that all participants, regardless of their gender, produce standardizable voice data to serve the researchers' machine learning model. This is an example of how computationally modelling language requires ideological modelling of assumptions about language and semiosis – but also about wellness and pathology, race, gender, and the like – that are embedded and naturalized in the infrastructure of an AI system.

In her contribution, Courtney Handman reports on the emerging field of AI-assisted animal communication. She shows that to expand the very definition of 'language' to animals, as some advocates propose, requires going beyond linguistic structure as usually understood and seeks instead evidence of a deep, possibly universal, underlying 'semantics', that is, meanings shared by animals and humans that AI is purportedly able to discover and perhaps eventually operationalize through being trained on datasets of animal communicative behaviour. What these advocates seek in these patterns, however, are indexical and iconic signs, the means by which, they suppose, animals communicate within their environments but which they hope AI technology can 'decode' and 'translate' into propositional human language. 'When humans, non-human animals, and machines all seem to be talking', Handman writes, definitions of language are reconfigured. This is language ideology. Indeed, even to claim that non-human animal communication is 'language' is a profoundly ideological assertion, rooted in complex projections and distinctions, and their blurring, between the human and the non-human cast in the shadow of the idea of a universal language that can translate between any communication system seamlessly (on which, more below). And here, too, although such AI projects challenge dominant concepts of language and the sign, ultimately the image of language involved remains precisely that of a coded form of communication to which animal behaviours can be translated as statements or propositions. It is this image that guides the effort to transform the indexical and iconic signals produced by animals into something that looks more like 'language' as the researchers think of it, namely, a system of symbolic (i.e. rule-governed) signs.

Semiotic ideologies, as already noted, implicate the political. To hold that what prevents humans from mistreating animals is largely a lack of communication, as some of these advocates seem to suggest, is, Handman proposes, to link a semiotic ideology about the effects of communication to some basic, underlying assumptions of liberal political theory. Running through the empirical phenomena that Handman discusses, thus, is an underlying ideological association between the possession of 'language' and the exercise of agency, an association shared by AI boosters and, surprisingly, colonial missionaries. It is as if to hear someone speak meaningfully is to have evidence that they are purposeful agents who may thus have rights and require representation. Conversely, to *not* communicate meaningfully is to fail to be a subject with rights or worthy of recognition. Semiotic ideology is never just about signs as such. It is also about authority and power.

Keane's contribution in this issue unpacks the semiotic work and ingrained habits of social interaction that people bring to AI chatbots in order to understand how ordinary users end up endowing AI with authority and even quasi-divine insights and powers. This tendency draws on certain fundamental properties of language, such as the ability for words to function independently of the context in which they were originally produced. Users seeking advice and insight interact with chatbots that produce verbalizations whose apparent authority derives in part from the enigmatic and unlocatable nature of their sources. These interactions can fall into the same patterns and semiotic ideologies we see in divination and oracular consultation. Moreover, the way chatbots produce texts in quasi-dialogic form taps into users' life-long conversational habits, making it easy to imagine that chatbots possess interiority and intentions – that they are quasi- or meta-humans. This 'likeness', or iconicity, of AI-produced and human-produced texts, like all iconic relations, depends not just on finding resemblance but on *eliminating* some properties as not relevant. To see that a red truck and a blue VW bug are both 'cars', as Captcha asks us, requires ignoring both their colour and their size. And were they both red, we would have to overlook that dimension of similarity in favour of their identity as 'cars'.

Both Semel and Handman in this issue observe that to focus on pattern recognition over huge databases, at which AI excels, invites researchers and AI boosters alike to seek modes of signification that exceed the interpretive capacities of mere humans and raise questions about the ideologies not only of AI programmers and designers, but those built into the technology itself. Indeed, as Keane (this issue) suggests, the distancing of human subjectivity from what takes place in the 'hidden layers' of AI that render its workings inscrutable to its own designers is precisely part of what prompts all sorts of theological or quasi-theological hopes and fears that are attributed to these technologies. This very inscrutability, however, draws on a host of semiotic ideologies for how to interact with divine beings that are familiar across the historical and ethnographic record. Other aspects, however, are more particular to our capitalist present. If AI is taken to be like a divinity, Keane asks, what does that mean for the powers society confers on the corporations that produce and profit from it?

In sum, all three articles draw out how language-based AI technologies put pressure on the very concept of language as studied in dominant traditions, even as they also reproduce that very concept as an aspiration or dream or oracular revelation. This requires us to draw on an expanded semiotics, one which centrally accounts for the multifunctional, ideological dimensions of communication. As these papers show,

semiotic ideologies form a crucial link between semiotic practices and their social and political conditions, but not a deterministic one. In particular, all three articles in this cluster touch on the role that financial interests and the demands of publicity discourse play in the boosterism, simplifications, and high hopes placed on AI's potential. On the one hand, it is important to retain a secure analytical grasp on the difference between actual practices and their popular depictions. On the other hand, ideologies in popular discourse are consequential and need to be taken into account in any ethnographic work as part of the phenomenon of AI.

This encounter with AI may also prompt us to take a more critical appraisal of what language might mean for anthropologists. The success of LLMs demonstrates that coherent communication is not necessarily reducible to abstract deductive rules, as an older generation of linguists attempted to model it, or as first-generation AI did (and failed). This is what linguistic anthropologists have long argued by stressing the importance of context, pragmatics, and ideology. AI also challenges the common (if not universal) assumption that coherent communication with words is distinctively human. If it turns out humans are not the only ones with language then what might language mean beyond the human? And what might language look like that is neither rule-based *nor* fully contextualized but instead probabilistic and narrowly co-textual, as Gershon suggests in her commentary in this issue? In asking this we are not engaging in a recuperative, reactive humanism. We are not trying to redefine language in order to differentiate 'real' human language from 'simulations' by machines – the 'imitation' part of Turing's 'Imitation Game'. The goal is to clarify the dimensions and processes of semiosis that produce language effects, to ask how and what these effects *do*, and to reflect on what this implies about how best to conceptualize language itself.

## AI and the dream of a universal, perfect language

As we noted above, one contribution anthropology can make is to put the semiotic ideologies underlying AI into a larger cultural and historical context. AI technologies embed assumptions with deep roots in the post-Enlightenment West. One such context in which we might situate certain current AI technologies and the discourses that surround them is the dream of universal languages (Eco 1995; Okrent 2009). The aspiration to discover, or fashion, a perfect, universal language has a long theological history in the Western world. This profoundly religious history attempts both to overcome the curse of Babel – that multilingualism leads to strife – and to return to an original Adamic language where the word named the truth of things by transparently revealing their very essence in its form.

Such projects have taken diverse forms over time. But they share a yearning for universal access to truth through a language that everyone can and should use. They also seek to make language, as a tool for making true or false statements, fully transparent and rational. As such, they offer a salve for certain anxieties and problems. One common refrain is that such artificial languages (whether this is the invented language that the seventeenth-century philosopher, John Wilkins called his 'real character' or the 'international auxiliary language' Esperanto invented by L.L. Zamenhof in the nineteenth century) will make up for the shortcomings and failures of natural, human language (Okrent 2009). These failings include the opacity of other minds and the incommensurability of cultures; the ambiguities in language that sow misunderstanding, confusion, and disagreement; and language's excessive variability, caprice, and illogicality. These themes can be found, for instance, in John

Locke's *Essay concerning human understanding*, which is very much a charter for contemporary ideologies of language and the political (Aarsleff 1982; Bauman & Briggs 2003).

The dream of a perfect language can be especially appealing at times of global strife, such as the Thirty Years' War and the First World War. Today's crises perhaps call for their own dreams of a universal code, a new language that would solve the problems of the day, to reconcile differences among humans or, in the case of Handman's utopians (this issue), between humans, animals, and the environment.

Gottfried Leibniz, as a progenitor of forms of formal logic and binary code (the foundation of modern computer languages) is informative here. His own 'rational calculus' secularized the Adamic dream; in it, each semantic atom would be assigned a prime number that could be manipulated through logical combination/multiplication to form more complex terms and propositions (Leibniz 1976). All you would have to do to discover the truth of a proposition or set of propositions, and thus to resolve a dispute, would be to 'sit down together at a table and say to each other (having called, if they so please, a friend) "let us calculate"' (cited by Eco 1995: 281).

Leibniz, as he himself admitted, could never achieve such a system. But his calculus, rooted in the medieval vision of Adamic language, laid the foundation for the mathematization, and later computational modelling, of so-called natural language. This 'universal symbolistic' reformulated language as a logical machine where a proposition, or complex term, is an equation, and its truth value a computation. As Leibniz argued, this equation, and its truth-value, would hold regardless of the assignment of the primitive terms to real-world entities. This is what Leibniz called 'blind thought': the idea that we could make a calculus with a set of symbols without knowing what those symbols denoted. This is precisely the modern syntactician's dream of so-called context-free grammar. Here, the language of God – Reason – is not in the world but is shifted to the 'code' of propositional structure itself, itself taken as revealing the basic components of thought ('Pure Thought' as Gottlob Frege put it 1972 [1879]; cf. Fodor 1975; Pinker 1994). And, of course, the capacity to formulate propositions from basic units with no 'knowledge' or at least 'experience' of what they might refer to is exactly what AI chatbots today do so well.

This view of language was taken up, in a more or less direct way some 150 years later, and conveyed to us today by Gottlob Frege, one of the founders (along with Peirce) of modern logic (Lee 1997). Like Leibniz, Locke, Peirce, and others, Frege also framed his 'conceptual notation' as necessary to counter the inherent irrationality and ambiguity of natural language. Ironically, Frege's logical notation, while designed to counter natural language's lack, came to serve, with later generative linguistics, as a metalanguage to describe natural language's propositional structure, just as forms of philosophical logic have served as the backdrop for the design of computer languages. It is no coincidence, then, that widely circulating, contemporary semiotic ideologies tend to depict so-called natural language as like 'artificial' languages such as computer code, for computer science has operationalized this dream just as generative linguistics has used it to model what it takes to be language's inner essence.

Certainly, today's AI is very different from these utopian projects for universal perfect languages, just as the mathematicization of language in LLMs works on different principles from those found in philosophical logic. But AI emerges from this history as an heir to these projects and as participating in their desires and anxieties. This is not because the image of language has been static across time. As

Semel points out in her article, the Fregean-Chomskian vision is a deductive proof-theoretic model, while new generation AI models are inductive and aleatory (see also Kockelman 2024; Lamoureaux, Castelle & Weichselbraun 2026; Wilf 2023). Even so, AI technologies like LLMs presuppose these earlier ideological projects both by incorporating them into aspects of their very structural logics and, in certain ways and in certain uses, by continuing their search. This dream is pursued in the projects described by Handman and Semel in this issue, where programmers search for an AI to translate plant and animal behaviour into human language (and vice versa), to discover a translator of human biosignals to medical diagnoses, or in the proliferation of translation programs that promise seamless movement across denotational codes. What is different today, however, is that the most widely used LLMs strive for 'naturalness' as their goal rather than being posed as improvements on the flaws or lack of 'natural' language.

## From translation to transduction

If the quest for a perfect language is one of the ideological strands that animate AI technologies and metadiscourses about them, linguistic anthropologists have stressed that there remains a productive gap between ideology, practice, and structure (Gal & Irvine 2019; Keane 2003; Silverstein 1979). And here, again, is why an ethnographic approach is critical. Indeed, what is notable about AI devices is that they not only translate but also *transduce* (Gal 2023; Silverstein 2003), operating not just on semantics but pragmatics as well. *Translation* focuses on the commensuration of source and target texts based on their semantic structure, like rendering the words of one language into another while maintaining their denotational meaning. The assumption is that both source and target texts involve the 'same' kind of thing: propositional meaning conveyed via the symbolic dimension of language. By contrast, *transduction* works beyond the symbolic. Linguistic anthropologists borrow the term from processes like those carried out by a hydroelectric generator, which converts one form of organized energy, the movement of water, to another, the mechanical movement of the turbine blades around a magnetized coil, in order to produce another form of energy, electricity (Silverstein 2003). Semiotic transduction operates on the iconic and indexical dimensions of signs in use, for instance, by reconstructing the meaning of a tone of voice, a gesture, or choice of speech register in some source text into some approximately commensurate form in a target text, as when one renders a verbal insult in a novel into a vulgar hand gesture in a film. Unlike translation, in transduction the source and target may involve different semiotic modalities altogether. This is just what we see when AI converts a source text such as algorithmic operations over a digital database (ones and zeroes) into a target text like interpretable written speech or a visual image, or vice versa, from a human language (in a training database) into an algorithmically mediated text (embeddings in the model).

It is in transduction that the uncanny effects that Keane (this issue) notes in his article come to the fore. For there are major gaps in how AI chatbots can transduce texts that iconize, index, and symbolically denote that world into the databases that allow them to produce discourse. Indeed, AI models are fundamentally unable to manage those gaps without human labour, which not only produces the texts upon which such models are trained, but which is also required to code and curate such training sets and to check the connection between the discourse produced by such chatbots and the world of referents or lack thereof, as with so-called hallucinations (Lamoureaux *et al.* 2026). As we noted,

this detachment from the world is already baked into the formal logical 'context-free' models of language that long precede AI, which are built on internal logical coherence relations rather than referential relations and truth conditions. The uncanniness that results is further powered by the fluid interactional and pragmatic coherence that the resulting texts display, despite lacking any direct grounding in a world shared with human users to which they seem to refer (Kockelman 2024: 78).

Here, we arrive at a sort of aporia in need of deconstruction. AI's social fluency is based on folding the world *into* its 'universe of discourse', namely, the databases that it has been trained on, that belies the obvious fact that current AI technologies lack an embodied real-time connection to that world. Or rather, they require *our bodies* to produce the universe of discourse on which they train, the sense-based categories which they encode, and the affective 'feel' for interaction which they imitate so well. As both Gershon and Keane put it in their contributions to this issue, AI technologies work on co-text, not con-text (also see Kockelman 2024: 36; Lamoureaux *et al.* 2026). They thus require *socio-linguistic histories* to establish the categories and concepts that their discourse reproduces. Yet if AIs are transducers and translators, they can do so because *we too* are transducers – and because we, too, depend on the bodies and minds and histories of others to make meaning. And they, like us, must do so because meaning is always already distributed and co-ordinated, not something 'in' a machine or 'in' an individual mind.

When we look at AI ethnographically, therefore, we find the erased work of transduction (Gal 2015). To take a simple example from Semel's article (this issue), consider the transduction from visual written form to sonic substance: reading aloud. Semel describes finding herself confused when, serving as a research subject, she was given the written forms 'oh' and 'ohh' and asked to pronounce them. The difference between how to read these written forms aloud is hardly intuitive. The researchers only came to terms with the unavoidable work of interpretation needed to render written form into sound at a point in the research process when they could no longer afford to revise their protocols. It is perhaps because scholars are so habituated to writing that details like this are so easy to miss, with the result that the work of transduction that is required to transcribe spoken into written words easily goes unnoticed. What Semel critically points out is that the labour researchers must carry out to control how the subjects transduce written text into verbal talk will be – like the labour of disarticulating the ideological tie between pitch and femininity – blackboxed once this data is used to train a machine learning model. The wider circulating semiotic ideology that attributes to these computational programs the supposed ability to uncover the hidden, Platonic truths that language holds disregards the ideological modelling that shapes their conditions of possibility. To recover this obscured labour requires an ethnographic eye, one attentive to the communicative structuring of AI-mediated design and interaction.

Sometimes, however, transduction does come into the foreground. No one can miss the role it plays in the efforts to reproduce animal communications in forms available to human sense-making that Handman writes about. Notice that transduction here takes two forms. On the one hand, it is the rendering of non-verbal animal signals into words. On the other hand, it is also the rendering of something taken to be pragmatics into semantic form. That is, to say that a certain sound or gesture means 'give me food' (a command) or 'I am hungry' (a statement) involves several interpretive moves. It renders sound or visually apprehendable movement in space and time into a complete

sentence, endowing it with a semantics and a syntax in a particular language, English (other languages would differently transduce these gestures in non-trivially different ways). Moreover, it entails a construal of just what that sentence is trying to do – in this case 'making a demand' or a 'statement' or even a 'hint' (Ervin-Tripp 1976). We can call this construal *meta-pragmatic*, since that construal takes the pragmatics of the sign as its object (Silverstein 1993). Metapragmatic interpretation – framing just what action a given stretch of speech is meant to be doing pragmatically (Bateson 1972 [1955]) – is so ubiquitous in everyday life that it is easy to miss until something goes awry and brings it to our attention. This is the point of the metapragmatic joke in a story by Ring Lardner that Michael Silverstein liked to quote: '"Shut up!" my father explained.'

This is why, as we've been suggesting, interacting with AI gives us such an uncannily familiar and disturbingly human feeling. It is precisely due to how we experience the capacity LLM-based AI texts have for transducing iconic and indexical significations, *not* because of their ability to generate statements or translate one set of propositions into another. It is their pragmatic fluency, their knack for coming up with discourse that seems appropriate to our genre expectations and interactional norms, that elicits our wonder, surprise, and anxieties.

AI technologies, in short, are not primarily translation machines; they are transducers, and their impact is not just in how they render meanings commensurate across media (whether these transductions render animal barks as human sentences; acoustic biomarkers as diagnoses; or verbal inputs as pictorial outputs). It is also how they transduce indexical and iconic significations in creative and unexpected transformative ways, sometimes humorous, sometimes deadly serious, sometimes transparent to our experience and, other times, opaque to it. As these articles show, what structures all this – these transductions and our experiences of them – is the interactional and ideological work that goes into AI's imagination and design. That AI technologies work as transducers is the result of the interdiscursive *chains* of translation, transduction, and transformation they presuppose and often efface, most notably from human labour to technological form to corporate profit and back again (Pasquinelli 2023). We are reminded in this total semiotic fact of the fetish which also, in Marx's formulation, speaks to us, through the mouths of economists (1977 [1886]: 177), about the 'grotesque ideas' that have grown from its 'wooden brain' (1977 [1886]: 164).

## Semiosis beyond the human

We have argued that understanding the social realities of AI requires an ethnographic focus on the often-effaced forms of human labour and interactional responsiveness that go into their design, meaningfulness, and use. This approach calls for seeing AI models and the human labour implicated in them as fundamentally semiotic. To do so, we go beyond the more familiar symbolic capacities of AI-generated text to foreground the iconic and indexical dimensions of language in use. A semiotic approach makes clear that language is never just a code or a structure, nor is its production of meanings confined to the calculations and cognitive processes of individual brains. The co-ordination of iconic, indexical, and symbolic dimensions of linguistic signs takes place in time, through interactions among sign users, under certain contextual circumstances. Against representationalist portrayals of language as primarily a device for making propositions and a conduit for transmitting information, images still commonplace in much public discourse, we stress the emergent, social character of meaning. The

apparent magic of AI-generated discourse derives in part from the way much of this semiosis and the human labour it involves are rendered invisible.

We stress the often-overlooked centrality of transduction to the functioning of AI texts – and to their limits and aporia – and the particular role of semiotic ideologies in that work and its effacement. Further, as we have suggested, taking such a linguistic anthropological perspective on AI raises a number of important questions that far outstrip AI technologies, among others, of language, personhood, and the human.

Having stressed the iconic and indexical, we also note that LLM-trained AI takes advantage of the distinctive property of language, as a system of symbols, that is, 'arbitrary' signs defined in relation to one another: the capacity to function *independently* of an original context. Drawing on vast quantities of text wrested from its context, LLMs operate via processes of transduction that are typically obscured from the user, by which text becomes digital matrix which in turn reappears in the form of text. Keane in his article in this issue notes that an important property of language is that it can be decentred from its event of use, that is, be *entextualized* (Bauman & Briggs 1990) and turned into a free-standing 'text' 'out of context'. In other words, so-called natural language is designed to be used apart from speaker intentions (Derrida 1988; Keane 1997) – but only to be reinserted, in turn, into new contexts, implicating new intentions.

As a result, despite its roots in human discursive labour, as noted above, AI programs trained on LLMs tend to reinforce what Semel calls the neo-Platonic view of language as something self-contained and autonomous of people, the worlds they inhabit, and the histories with which they are entangled. This, in turn, can give plausibility to the appearance of truth-giving objectivity that algorithms encourage and even the temptation to treat AI as a divination tool or access to the god's eye view. Ironically, AI may even be inviting some users to reinvent the medieval scholastic and Enlightenment technique of seeking truths embedded 'in' language and its propositionality. More generally, the ubiquity of divination techniques shows that pattern-seeking in search of hidden truths has a very long history.

The search for hidden truths often leads the searcher to sources beyond the human. In many respects the linguistic facility of AI invites users to continue this quest. In the process, AI exerts pressure at the limits of language and the human. For example, as Handman (this issue) points out, by detaching 'language' from human speakers and writers, AI forces us to reconsider just what counts as language. And if we take language to be part of what defines humans, as has been common in many traditions, it also forces us to reconsider what counts as human. Yet Handman's discussion also shows that the problem is not entirely new in another way: it also arose when early missionaries encountered colonial peoples whose humanity might have been in doubt (a theme in the Enlightenment ideologies of language noted above; see Bauman & Briggs 2003), just as it may arise when parents foster and communicate with children cognitively unable to speak (Reno 2024; Rutherford 2025), or when family members interact with elders in comatose states (Shohet 2021). Humans in diverse times and places have often pondered and reconfigured the limits, and meaning, of humanity and personhood by reflecting on the limits of language: who can communicate and make sense (Green 2024), who can understand, and who can be recognized in and by language (e.g., by being named; Butler 1997; vom Bruck & Bodenhorn 2006). As Keane notes (this issue), AI technologies hover on the unsettled borders of humanity and personhood.

Having noted there is much that is not new in the semiotic ideologies around AI, it is nonetheless certainly the case that new technologies based on LLMs and other self-learning computational systems are transforming the various social worlds in which they are embedded, and to which they are adjacent. And while we do not deny the rapid pace at which this is happening and the real, tangible changes that are swirling around us, an ethnographic eye towards the semiotic organization of such technologies, from design to implementation, from ideology to use, from structure to effects, helps us understand not only *how* such transformations are being wrought, and what to make of them, but also to sort out how much the newness of AI is old wine filling new skins and how much of it is new wines rupturing the old skins into which they are being poured. At the same time, an ethnographic view affords a critical vantage point to recover what is erased in the utopian or apocalyptic discourse of many of its proponents and opponents: namely, the communicative labour put into and emerging out of encounters with such technologies and the real political economic interests wrapped up in them.

But even more than this, as each of the papers shows, AI can serve as a kind of mirror back on ourselves; or rather, as a kind of opportunity to re-open what, at times, appear as settled questions, questions of the human and the non-human, of language and non-language, and about what it is, in the historical moment of the technological new, that disturbs and entices us so. By sharpening our analytical grasp of how AI's semiotic skills are and are not like those of human and other non-human agents, we can better understand both. For this, we need semiotically informed anthropological perspectives. We require a linguistic anthropology of AI.

## Acknowledgements

The authors thank Dimitrios Theodossopoulos for the invitation and support to put together this thematic cluster of articles, and for his helpful feedback, on this introduction and the theme issue as a whole, through the review process.

## NOTES

[1] See also Lamoureaux *et al.* (2026) for a thorough overview and genealogy of linguistic anthropology and the study of large language models.

[2] Tense is an indexical category (what Jakobson called a 'shifter') insofar as the denotation of the category (it being present or past) depends on ('shifts' based on) the moment of its enunciation, relative to which the denoted event is present (co-eval) or past (i.e., prior to).

[3] See Lamoureaux *et al.* (2026) for use of Peirce's semiotics to describe the computational functions of deep neural networks.

## REFERENCES

Aarsleff, H. 1982. *From Locke to Saussure: essays on the study of language and intellectual history*. Minneapolis: University of Minnesota Press.

Bateson, G. 1972 [1955]. A theory of play and fantasy. In *Steps to an ecology of mind*. New York: Ballantine Books.

Bauman, R. & C. Briggs 1990. Poetics and performance as critical perspectives on language and social life. *Annual Review of Anthropology* **19**, 59-88.

——— & ——— 2003. *Voices of modernity: language ideologies and the politics of inequality*. Cambridge: University Press.

Butler, J. 1997. *Excitable speech: a politics of the performative*. New York: Routledge.

D'Armenio, E., M.G. Dondero, A. Deliège & A. Sarti 2025. For a semiotic approach to generative image AI: on compositional criteria. *Semiotic Review* **9**. *https://doi.org/10.71743/ee5nrx33*.

Derrida, J. 1988. *Limited, Inc*. Evanston, Ill.: Northwestern University Press.

Dondero, M.G. 2025. Enunciative praxis and enregisterment in the domain of generative artificial intelligence. *Semiotic Review* **12**. *https://doi.org/10.71743/bwf80g96*.

Eco, U. 1995. *Search for the perfect language*. Oxford: Blackwell.
Ervin-Tripp, S. 1976. Is Sybil there? The structure of American English directives. *Language in Society* **5**, 25-66.
Fodor, J. 1975. *Language of thought*. New York: Cromwell Company.
Frege, G. 1972 [1879]. *Conceptual notation and related articles*. Oxford: University Press.
Gal, S. 2015. Politics of translation. *Annual review of Anthropology* **44**, 225-40.
——— 2023. Rethinking translation and transduction. In *A new companion to linguistic anthropology* (eds) A. Duranti, R. George & R.C. Riner, 178-93. Hoboken, N.J.: Wiley.
——— & J.T. Irvine 2019. *Signs of difference: language and ideology in social life*. Cambridge: University Press.
Gershon, I. 2010. Media ideologies: an introduction. *Journal of Linguistic Anthropology* **20**, 283-93.
Green, M. 2024. *Making sense: language, ethics, and understanding in deaf Nepal*. Oakland: University of California Press.
Jakobson, R. 1960. Closing statement: linguistics and poetics. In *Style in language* (ed.) T. Sebeok, 350-77. Cambridge, Mass.: MIT Press.
Keane, W. 1997. *Signs of recognition: powers and hazards of representation in an Indonesian society*. Berkeley: University of California Press.
——— 2003. Semiotics and the social analysis of material things. *Language & Communication* **23**, 409-25.
——— 2007. *Christian moderns: freedom and fetish in the mission encounter*. Berkeley: University of California Press.
——— 2018. On semiotic ideology. *Signs and Society* **6**, 64-87.
——— 2024. *Animals, robots, gods: adventures in the moral imagination*. London: Allen Lane.
Kockelman, P. 2024. *Last words: large language models and the AI apocalypse*. Chicago: Prickly Paradigm Press.
Kohavi, N. & A. Weichselbraun 2026. Human tests for machine models: what lies 'Beyond the Imitation Game'? *Journal of Linguistic Anthropology* **36**, e70035.
Lamoureaux, S., M. Castelle & A. Weichselbraun 2026. Language machines: toward a linguistic anthropology of large language models. *Journal of Linguistic Anthropology* **36**, e70033.
Lee, B. 1997. *Talking heads: language, metalanguage, and the semiotics of subjectivity*. Durham, N.C.: Duke University Press.
Leibniz, G. 1976. *Philosophical papers and letters*. Dordrecht: Kluwer Academic.
Leone, M., W. Keane & C.V. Nakassis 2025. Semiotic ideology: a dialogue across semiotic traditions. *Semiotic Review* **12**. *https://doi.org/10.71743/zafmdt34*.
Marx, K. 1977 [1886]. *Capital: a critique of political economy*, Vol. **1**. London: Penguin Books.
Okrent, A. 2009. *In the land of invented languages: Esperanto rock stars, Klingon poets, Loglan lovers, and the mad dreamers who tried to build a perfect language*. New York: Spiegel & Grau.
Pasquinelli, M. 2023. *The eye of the master: a social history of artificial intelligence*. London: Verso Books.
Peirce, C.S. 1956. *Philosophical writings of Peirce* (J. Buchler, ed.). Garden City, N.Y.: Dover Books.
Pinker, S. 1994. Mentalese. In *The language instinct: how the mind creates language*, 55-83. New York: William Morrow Company.
Reno, J.O. 2024. *Home signs: an ethnography of life beyond and beside language*. Chicago: University Press.
Roose, K. 2023. A conversation with Bing's chatbot left me deeply unsettled. *New York Times*, 16 February. *https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html*.
Rutherford, D. 2025. *Beautiful mystery: living in a wordless world*. Durham, N.C.: Duke University Press.
Schieffelin, B.B., K.A. Woolard & P.V. Kroskrity (eds) 1998. *Language ideologies: practice and theory*. Oxford: University Press.
Shohet, M. 2021. *Silence and sacrifice: family stories of care and the limits of love in Vietnam*. Oakland: University of California Press.
Silverstein, M. 1979. Language structure and linguistic ideology. In *The elements: a parasession on linguistic units and levels* (eds) P. Clyne, W. Hanks & C. Hofbauer, 193-247. Chicago: Chicago Linguistic Society.
——— 1993. Metapragmatic discourse and metapragmatic function. In *Reflexive language: reported speech and metapragmatics* (ed.) J. Lucy, 33-58. Cambridge: University Press.
——— 2003. Translation, transduction, transformation: skating 'glossando' on thin semiotic ice. In *Translating cultures: perspectives on translation and anthropology* (eds) P. Rubel & A. Rosman, 75-105. Oxford: Berg.
——— & G. Urban 1996. *Natural histories of discourse*. Chicago: University Press.
vom Bruck, G. & B. Bodenhorn 2006. *The anthropology of names of naming*. Cambridge: University Press.
Wilf, E. 2023. *The inspiration machine: computational creativity in poetry and jazz*. Chicago: University Press.

## Introduction : vers une anthropologie linguistique de l'IA

*Résumé*

Ce texte introduit l'ensemble thématique « Vers une anthropologie linguistique de l'IA ». L'essor de l'intelligence artificielle (IA), en particulier des grands modèles de langage capables de produire un discours cohérent imitant une conversation, exerce une pression sans précédent sur les conceptions dominantes du langage, de la personne et de l'humain. Ce défi est aussi une occasion de réflexion, en particulier pour l'anthropologie linguistique, une discipline centrée sur le langage comme moyen et pratique de création de signification dans la vie sociale. Les articles regroupés ici ne se limitent pas à analyser l'IA en tant que machine de communication – ils explorent les processus communicationnels qui la structurent en tant qu'imaginaire idéologique et infrastructure technologique et font appel à l'anthropologie linguistique pour examiner différents aspects de l'IA à mesure que son adoption se généralise, comme les fonctions qui incitent les utilisateurs à traiter avec les robots conversationnels comme s'il s'agissait de personnes (Keane), les tentatives d'utiliser l'IA pour déchiffrer la communication animale (Handman) et la recherche de biomarqueurs linguistiques prédictifs de troubles de la santé mentale (Semel). Afin de placer les articles dans leur contexte, la présente introduction expose quelques distinctions sémiotiques de base développées en anthropologie linguistique et nécessaires pour analyser l'IA, en mettant l'accent sur le concept d'idéologie sémiotique. Les auteurs situent ensuite l'IA dans une série de généalogies idéologiques, de la vision des langues universelles des Lumières au représentationalisme de la linguistique moderne, avant d'aborder les questions de la transduction et de la sémiosis au-delà de l'humain.

Webb Keane is the George Herbert Mead Distinguished University Professor at the University of Michigan. His books include *Animals, robots, gods: adventures in the moral imagination* (2024), *Ethical life: its natural and social histories* (2016), and *Christian moderns: freedom and fetish in the mission encounter* (2007).

*Department of Anthropology, 101 West Hall, 1085 S. University Ave, University of Michigan, Ann Arbor, MI 48109, USA. wkeane@umich.edu*

Constantine V. Nakassis is Professor of Anthropology at the University of Chicago. He is the author of the books, *Doing style*: *youth and mass mediation in South India* (2016) and *Onscreen/offscreen* (2023).

*Department of Anthropology, The University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, USA. cnakassi@uchicago.edu*