Theories and Methodologies

Inside and outside the Language Machines

AARTHI VADDE

The release of *ChatGPT* in the fall of 2022 kicked the AI hype cycle into overdrive. By overdrive, I mean that *ChatGPT* penetrated public consciousness in a way that its predecessor *GPT-3* had not. I read countless essays by tech journalists in papers like *The New York Times*, *The Atlantic*, and *The New Yorker*. They reported on their uncanny encounters with the chatbot; they purported that the college essay was dead; they riposted that the college essay was very much alive.

In my role as director of undergraduate studies in the English department at Duke University, I fielded emails from university colleagues convinced that *ChatGPT* would make plagiarism detection impossible (they are right about that), and from instructors who suspected student papers had been written by *ChatGPT* (no way to prove it but an opportunity to have an honest conversation). I was asked if the English department should have a policy on AI (I said no; individual instructors should develop their own policies depending on the course). I was given a one-hundred-page "living document" by the Office of Learning Innovation to help me keep pace with this fast-changing technology (parts were genuinely helpful).

All this is to say that I recognized the environment described in Matthew Kirschenbaum and Rita Raley's essay and the feeling it produces: exhaustion borne of predictability and saturation at every level. Large language models (LLMs) work because of sophisticated forms of mathematical prediction and vast amounts of training data. The companies investing in LLMs (Microsoft, Meta, and the like) have already used their tentacular reach to integrate them into the infrastructure of word processing, speech recognition, image creation, and various other machine-mediated modes of communication. The universities reacting to the companies investing in the LLMs respond in scripted and templated ways: task forces, white

AARTHI VADDE is E. Blake Byrne Associate Professor of English at Duke University. Her book in progress is called "We the Platform: Contemporary Literature after Web 2.0." She is the author of *Chimeras of Form: Modernist Internationalism beyond Europe* (Columbia UP, 2016; winner of the 2018 Harry Levin Prize) and coeditor of multiple volumes, most recently the eleventh edition of *The Norton Anthology of English Literature* (2024).

© 2024 The Author(s). Published by Cambridge University Press on behalf of Modern Language Association of America
PMLA 139.3 (2024), doi:10.1632/S0030812924000579

https://doi.org/10.1632/S0030812924000579 Published online by Cambridge University Press

papers, and best practices. While the technology might be new (or newly effective in spheres of symbolic expression), the economic model is woefully familiar. The neoliberal university will capitulate to the gods of data. It will align its mission with the requisite management schemes and workforce arrangements, and it will be up to faculty members, students, and workers to fight for their education and their information.

Kirschenbaum and Raley's high-level diagnosis of the economic and cultural situation of the university seems airtight to me; indeed, it reminds me of what humanists do so well and train our students to do so well: see the intersections of power and knowledge, dissect the biopolitical logics of control behind the probabilistic models of prediction. These are theoretical tasks that take on new force in a world where the cost of producing deepfakes and political propaganda is lower than ever. They are also necessary to guard our "cognitive liberty," a new kind of human right made urgent by the development of neurotechnology and the migration of data-collecting machines into the brain itself (Farahany). Brain data foretells a world after language-based data in which human thought will not need to become human expression in order to be extracted.

To raise such visions of the future is not to contemplate a postlinguistic society but to join Kirschenbaum and Raley in trying to name what is truly new about the way language circulates today and how it is tied to the ever-growing entanglement of human and machine cognition (Hayles). Language is the territory staked out for corporate capture because it is a profitable conduit of the human mind, but, as Kirschenbaum and Raley suggest, the proliferation of automated inscription (synthetic text, image, voice, and so on) makes it difficult to maintain human-centered notions of language. When prompting a model supersedes the typing of keys as a scene of writing, then the words, pictures, and video that flow from the model belie an entire machinic architecture outside human perception.

The particular architecture of the LLM depends on high-performance chips and is in turn connected to a global infrastructure that includes undersea cables, landing stations, and data centers. Eighty percent of data center costs go to cooling (Gordon and Jones 3), and LLMs are energy and water intensive. Kate Crawford reports that residents in West Des Moines, Iowa, sued OpenAI in 2022 because the data center cluster serving *GPT-4* (the most advanced model at the time) used 6% of the district's water. Google and Microsoft revealed that their water usage went up 20% and 34%, respectively, in 2022 as they implemented language models for *Bard* and *Bing*. The insides of LLMs might be black boxes, but the outsides consist of fresh water, rare earth metals, and massive carbon footprints.

AI leaves its traces on the earth, like industrial revolutions before it, but unlike those revolutions, it also insinuates itself into the marrow of language. Frederic Kaplan has suggested that the merging of human-authored texts with machine-edited and automated texts creates a textual situation in which LLMs may coin their own words and expressions:

> As a consequence, these expressions may be suggested to us [humans] as statistically plausible forms by autocompletion algorithms. If we follow this hypothesis, natural languages could progressively evolve to seamlessly integrate the linguistic biases of algorithms and the economical constraints of the global linguistic economy. (62)

An LLM-driven evolution of language might sound nightmarish, but it would not be the first time that Internet argot changed natural languages. Recall 2015 when Oxford Dictionaries declared the tears of joy emoji (😂) to be the word of the year. The last two words of the year—*rizz* (2023) and *goblin mode* (2022)—also trace their explosive popularity to influencers and viral posts on social media platforms ("Oxford Word of the Year").

Evolutionary linguistics and sociolinguistics sometimes fall into the remit of language and literature departments, but they are not usually part of the training of literary scholars. I think the study of generative AI should bring linguistics and literary

https://doi.org/10.1632/S0030812924000579 Published online by Cambridge University Press

study closer today. This could be termed a computational return to philology in which theories and histories of computer-mediated language dovetail with qualitative and quantitative studies of literary and mass forms of writing. Whatever we call it, the coming ubiquity of language models poses a number of questions endogenous to literary studies even if the precipitating factor (the release of *ChatGPT*) is exogenous.

A brief list: How does a model's manipulation of symbols change what we mean by language if the machine's output is defined as the mapping of numbers into words rather than ideas into words? Will multimodal models demand a new "panaesthetic" theory that reactivates debates about the unity and diversity of the arts (Albright)? Various multimodal models like *Dall-E* (text-to-image), *LLark* (music-to-text), *MusicLM* (text-to-music), and *Sora* (text-to-video) use large datasets and vector representations to convert one medium into another. These models' operations seem like another iteration of the computer as a universal machine, but their numerical conversions might also be compared to the verbal conversions of nonverbal art forms like music and painting into criticism. As Daniel Albright writes, "Art is both a language and a not-language. If we ask an artwork to have a meaning, it will obey us by manifesting itself as speech or writing . . . every artistic medium is a language, but I can say this only because language understands everything as a language" (8). Albright's identification of verbal language as its own universalizing solvent renders words more like numbers when considered in the light of nonverbal media. How is the alliance of word and number shaping contemporary creative practices as well as those to come? Finally, how should language and literature departments prepare students for the world in which they live—the world to which the above questions speak?

Humanities scholars are uniquely positioned to talk about language models as writing technologies and potential teaching tools. Kirschenbaum and Raley mention the partnership between Arizona State University and OpenAI. We can of course imagine grim scenarios involving austerity politics and writing instruction by app, but university-based literature, writing, and foreign language instructors can still shape whether and how these technologies are used in the classroom. We should be at the forefront of those decisions, and we should welcome conversations at every level about the nature of language, the purpose of writing, and the uses of literature. This is not a bid for relevance, but an identification of a lacuna around the meaning of language in the public discourse on AI.

To think about what "language" means when we talk about LLMs is to realize that literature departments have already been there and done that. The most virulent debates over generative AI hinge on the terminology used to describe model capabilities. Leading figures in the current AI revolution (for example, Geoffrey Hinton and Yann LeCun) and their esteemed critics (for example, Emily M. Bender, Timnit Gebru, and Margaret Mitchell) disagree about the meaning of words like *intent*, *meaning*, and *understanding* when attributing agency to models. In the much cited paper "On the Dangers of Stochastic Parrots," Bender and her coauthors deny LLMs the capacity for linguistic understanding while granting models mastery over linguistic form. They insist on the difference between human-centered categories like "communicative intent" and machine-centered categories like "controlled generation," by which a larger system guides the output of a model toward certain styles and topics (616).

Kirschenbaum, in a *Critical Inquiry* forum titled "Again Theory," characterized "On the Dangers of Stochastic Parrots" as "red meat" for a graduate seminar on literary theory, the implication being that its account of linguistic meaning was too "linear" for scholars immersed in the play of the signifier. He was right, considering that the majority of participants in the forum could accept the political critique of dangerously large and biased LLMs while rejecting the authors' underlying philosophy of language. Adherents of poststructuralist theory (and ordinary language philosophy for that matter) refuse to ground meaning in intention or in the ideas of an experiencing subject, which speaks

https://doi.org/10.1632/S0030812924000579 Published online by Cambridge University Press

to the irony that LLMs have upended pedagogy in literature and language departments while at the same time confirming (without acknowledgment) diverse disciplinary orientations and insights.[1] Ted Underwood calls this development the "empirical triumph of theory" in his contribution to the forum, but empirical triumph in no way guarantees conceptual credit or other forms of triumph for beleaguered humanists contending with the downstream effects of models that write like people.

Whether or not we agree that models produce language or the simulation of language, generative AI is now an undeniable actor in the writing process. It occasions reflection on what writing is and why we still want students to do it. Amit Gupta, one of the founders of Sudowrite (a start-up established in 2020 and now pitched particularly to creative writers) described writing with models in the following way: "Your role starts to become deciding what's good and executing on your taste, not as much the low-level work of pumping out word by word by word. You're still editing lines and copy and making those words beautiful, but, as you move up in that chain, and you're executing your taste, you have the potential to do a lot more" (qtd. in Marche). In this world, "the craft of language becomes an afterthought" (Marche), a sentiment that has been described as "reductive if not insulting" (Elam 283).

Gupta's description of the writer's role as "deciding what's good" presupposes that the prompter knows what good writing looks like. I would not take this for granted. I would, however, welcome in any writing-based class a comparison of Gupta's statement with this one from Gottfried Wilhelm Leibniz: "For it is unworthy of excellent men to lose hours like slaves in the labor of calculation which would safely be relegated to anyone else if machines were used" (173). Leibniz wrote the original statement in Latin in 1685 to accompany his invention of a calculator that could add, subtract, multiply, and divide. I would like to know if students think that "writing" could substitute for "calculation" in Leibniz's statement and whether the labor reflected in their writing could "safely be relegated to anyone else if machines were used." I am ready to be surprised, affirmed, and disappointed by their various answers because the discussion would refine our collective understanding of writing, calculation, and these practices' differential relationships to thought.

Vauhini Vara has written that the purpose of writing for her, a professional novelist and journalist, is to "clarify the world from where I stand in it. That definition of writing couldn't be more different from the way AI produces language: by sucking up billions of words from the internet and spitting out an imitation." She made this statement in an essay reflecting on the virality of her partially AI-generated essay "Ghosts," and the implication was that she would recant her use of language models in the name of a literary project of particularizing human consciousness. Influenced by Zadie Smith's essay "Fail Better," she joins Smith in defining literary style as the amalgam of the written word and an individual consciousness where the refinement of style leads to the refinement of consciousness.

This is a strongly humanistic vision of literariness, which I respect but also want to challenge by looking at examples of contemporary literature that complicate the excision of data from the worthiness of style. Recent experimental novels including *The Nature Book*, by Tom Comitta, and *Alphabetical Diaries*, by Sheila Heti, also seem interested in the project of consciousness, but for what the ecologist David Abram calls a "more-than-human world." Abram has stated that, for some people, there is no getting past human exceptionalism, and he has wondered whether that limit is derived from the phenomenal experience of being human:

> our perception that humans are so qualitatively different—so diverse and nuanced in our creative and cognitive gifts—had something to do with the simple fact that we, ourselves, *are* human, and so know our own species *from the inside*. In other words, we view the other species necessarily only from the *outside*, so that only their most obvious and generalized traits are apparent, or *salient*, to us.

We will never know what it is like to be another species from the inside, but *The Nature Book* gives

https://doi.org/10.1632/S0030812924000579 Published online by Cambridge University Press

readers a portrait of the more-than-human world in a manner approximating the outside of human perspective. Comitta's novel, culled from 302 novels recycled into a new form, consists of found phrases and passages edited according to an Oulipo-like set of constraints in order to eliminate human characters from the narrative and delve into biospheres.

To achieve what they call a "literary supercut," Comitta treated text like data and books as material things ("Brief History"). They printed out fifteen hundred pages, physically cut them up, arranged them across twelve tables, and pasted words together to arrive at an 87,000-word novel. It is an artisan work, made partially by hand, but it is also born of a dataset and rules for pattern recognition. As Comitta says, the novel "contains no words of my own."[2] Its style is synthetic even if the text is not, and the techniques involved suggest that writers already have all the language they need in the storehouse of literature.

Heti performs an experiment similar to Comitta's with text as data. Instead of turning to the literary tradition, she turned to herself in *Alphabetical Diaries* by collecting 500,000 words from ten years of journaling and sorting them alphabetically through a spreadsheet. She then proceeded to edit the results but keep the alphabetical ordering. Here is one memorable sentence from the work: "A radical sympathy with all people based on their integrity as becomings, not beings; as people who experience the potential freedom of their own souls, so to radically know that people experience themselves from the inside, and not one person alive has experienced themselves from the outside" (7). The sentence can be imagined back into the original context of Heti's writing in her personal diary, but its alphabetical placement in the published work suggests that people alive today are also capable of experiencing themselves from the outside if we view our self-expression in tabular form, retrieve our search histories from *Google*, our voice histories from virtual assistants like *Alexa*, our reading histories from apps like *Kindle*, our viewing histories from *Netflix*, and so on. The externalization of self, now associated with surveillance capitalism, yields the data on which language models train. While novelists can try to resist these conditions through their writing, they can also clarify these conditions by allowing quantification into their creative processes in ways that change the form of fiction.

Comitta's and Heti's novels might be low-tech examples, but the point of their numerical literariness is not to use the latest language models (though Heti is also doing that in other projects). It is to think about the conjuncture of text and data and to contend with the irrefutable fact that datasets contain the foundational language of our time. As creative writers work more with data and with language models, they will regenerate our literary past through machinic processes that may rely on prediction but will lead to unpredictable kinds of literary characters and authorial voices.

I am referring here not just to Comitta and Heti but also to the computational poet Lillian-Yvonne Bertram, who developed *Warpland 2.0*, a language model based on *GPT-3* and fine-tuned on a corpus of texts by Gwendolyn Brooks and named after Brooks's first and second sermons on the Warpland (Bertram 5). In their poem "Tell me a Black story," Bertram repeats the titular prompt to *GPT-3* and *Warpland 2.0* and juxtaposes the results. The distance between the outputs is striking: *GPT-3* tells stories in the third person about danger and disenfranchisement, referring to Black characters as "she" and "they," while *Warpland 2.0* yields output in the first person in warm phrases that Brooks may never have uttered but that are inflected by her writerly voice. Not a synthetic style, but synthetic text mediated through a Bertram-assembled corpus and used for the refinement of consciousness.

Engineers like to call specialists in other fields domain experts. For deep learning (the name of the machine-learning method behind the success of LLMs) to breed more than deep disillusionment among humanists, we must recognize that our domain is language. Among all the disciplines in the university, the humanities have been theorizing, historicizing, and studying language in all its expressive forms the longest. Many forces inside

https://doi.org/10.1632/S0030812924000579 Published online by Cambridge University Press

and outside the university are diminishing the value of humanistic expertise, but generative AI is revealing it anew.

Notes

1. Ordinary language philosophy would ground meaning not in the play of the signifier but in the use of words. Its proponents would also demur from endorsing a general "theory" of language even though Wittgensteinian philosophy has also had a major impact on natural language processing. I would point here to the Wittgenstein-influenced linguist J. R. Firth's famous insight, "You shall know a word by the company it keeps!" (11), which is cited across the field of natural language processing and tied to the development of distributional semantics.

2. I am grateful to Deidre Lynch for introducing me to *The Nature Book*. For our discussion of the novel with Comitta, see "Narrative, Database, Archive," an episode of the podcast *Novel Dialogue*.

Works Cited

Abram, David. "On Being Human in a More-Than-Human World." *Center for Humans and Nature*, 22 July 2012, humansandnature.org/to-be-human-david-abram/.

Albright, Daniel. *Panaesthetics: On the Unity and the Diversity of the Arts*. Yale UP, 2014.

Bender, Emily M., et al. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜." *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Mar. 2021, pp. 610–23, https://doi.org/10.1145/3442188.3445922.

Bertram, Lillian-Yvonne. *A Black Story May Contain Sensitive Content*. New Michigan Press, 2024.

Comitta, Tom. "A Brief History of Citational Fiction and the Literary Supercut." *Literary Hub*, 5 Nov. 2020, lithub.com/a-brief-history-of-citational-fiction-and-the-literary-supercut/.

———. *The Nature Book*. Coffee House Press, 2023.

Crawford, Kate. "Generative AI's Environmental Costs Are Soaring—and Mostly Secret." *Nature*, 20 Feb. 2024, www.nature.com/articles/d41586-024-00478-x.

Elam, Michelle. "Poetry Will Not Optimize; or, What Is Literature to AI?" *American Literature*, vol. 95, no. 2, 2023, pp. 282–303.

Farahany, Nita. *The Battle for Your Brain: Defending the Right to Think Freely in the Age of Neurotechnology*. St. Martin's Press, 2023.

Firth, J. R. *Studies in Linguistic Analysis*. Basil Blackwell, 1957.

Gordon, Lori W., and Karen L. Jones. *Global Communications Infrastructure: Undersea and Beyond*. *Center for Space Policy and Strategy*, Feb. 2022, csps.aerospace.org/sites/default/files/2022-02/Gordon-Jones_UnderseaCables_20220201.pdf.

Hayles, N. Katherine. *Unthought: The Power of the Cognitive Nonconscious*. U of Chicago P, 2017.

Heti, Sheila. *Alphabetical Diaries*. Farrar, Straus and Giroux, 2024.

Kaplan, Frederic. "Linguistic Capitalism and Algorithmic Mediation." *Representations*, vol. 127, no. 1, 2014, pp. 57–63.

Kirschenbaum, Matthew. "Again Theory: A Forum on Language, Meaning, and Intent in the Time of Stochastic Parrots." *In the Moment*, 26 June 2023, critinq.wordpress.com/2023/06/26/again-theory-a-forum-on-language-meaning-and-intent-in-the-time-of-stochastic-parrots/.

Leibniz, Gottfried Wilhelm. "Leibniz on His Calculating Machine." Translated by Mark Kormes. *A Sourcebook in Mathematics*, edited by David Eugene Smith, McGraw-Hill, 1929, pp. 173–81.

Marche, Stephen. "The Computers Are Getting Better at Writing." *The New Yorker*, 30 Apr. 2021, www.newyorker.com/culture/cultural-comment/the-computers-are-getting-better-at-writing.

"Narrative, Database, Archive: A Discussion with Tom Comitta and Deidre Lynch (AV)." *Novel Dialogue*, 2 Nov. 2023, noveldialogue.org/2023/11/02/6-3-narrative-database-archive-a-discussion-with-tom-comitta-and-deidre-lynch-av/.

"Oxford Word of the Year." *Oxford Languages*, 2024, languages.oup.com/word-of-the-year/.

Underwood, Ted. "The Empirical Triumph of Theory." *In the Moment*, 29 June 2023, critinq.wordpress.com/2023/06/29/the-empirical-triumph-of-theory/.

Vara, Vauhini. "Confessions of a Viral AI Writer." *Wired*, 21 Sept. 2023, www.wired.com/story/confessions-viral-ai-writer-chatgpt/.

https://doi.org/10.1632/S0030812924000579 Published online by Cambridge University Press