

Article

What Is It Like to Sound Like a Bot?

Amanda Paxton

University of Toronto

Two years into my PhD, I became a teaching assistant in a first-year English course. One day, my course director, a very tenured mid-career professor, barreled into the office with a stack of printouts he'd made for the students, to be distributed along with the first batch of graded essays. On the printouts was a collection of common errors he'd compiled from the papers, with one specific extract that appeared to push him to the brink of collapse. He read aloud to me—"In the hustle and bustle of daily life..."—and then threw his hands up, cheeks inflamed, spluttering, "Who talks like that??" As a graduate student with little teaching experience, I blinked at him silently while discretely scanning for the closest exit. Looking back years later, I understand his frustration.

We all know that clichés abound, not only in student writing, but also in the discourse of business, bureaucracy, and—to be sure—academia. As William Zinsser so aptly put it, "writing is hard" (2006, p. 9). Indeed, our brains are inherently conservative organs that will relax into the phrase at hand even if it doesn't help crystallize a clear meaning—or make much meaning at all. Moreover, as one student of mine remarked this term, students are often incentivized to produce texts that are "almost robotic."

The question of clichés is, of course, fraught. What constitutes a cliché in itself has historically been determined by authority and native-speaker privilege. Orin Hargraves observes the inherent arbitrariness behind such determination: "if enough people think a form of words is overused, or if a person who is perceived as having some authority about language declares such a thing, then the word or phrase becomes a cliché" (2014, p. 4). Moreover, what counts as a cliché is culturally contingent. Su-Yueh Wu and Donald Rubin note that "using proverbs is considered a cliché according to English essay standards. On the contrary, it is a valued rhetorical device according to Chinese writing" (2000, p. 174). Ultimately, however, the central issue is conscious choice: we want to empower students to make thoughtful, deliberate rhetorical decisions rather than producing language based on predictability.

When I taught the first writing class of my own years later, I showed students a post from a veritable hive of the cliché—the Perez Hilton celebrity gossip site— and asked them to identify the clichés, hackneyed phrases, and filler words that added no meaningful information. I then asked them to consider the effect of such writing. The consensus was that the text sounded generated, as if it had been written not by a human but by a bot. This description isn't far from the one we see in George Orwell's "Politics and the English Language," when he describes worn out phrases associated with political speech:

When one watches some tired hack on the platform mechanically repeating the familiar phrases ... one often has a curious feeling that one is not watching a live human being but some kind of dummy: a feeling which suddenly becomes stronger at moments when the light catches the speaker's spectacles and turns them into blank discs which seem to have no eyes behind them. And this is not altogether fanciful. A speaker who uses that kind of phraseology has gone some distance toward turning himself into a machine. (1946)

We currently find ourselves at a moment in which the large-scale availability of *actual* generated text from Large Language Models like ChatGPT has variously elicited ebullience, hand-wringing, and public debates as to whether the college essay is in fact dead (or, perhaps, merely pining for the fjords). Instead, I suggest that this moment presents an opportunity for postsecondary instructors to have meaningful conversations with students about issues many of us may not ever have had the chance to broach with them: the connections between writing, language, and deep existential and political questions. We find ourselves in an unprecedented moment in that so much public discourse concerns the hard problem of consciousness; AI-generated text can help us capitalize on those discussions in the classroom. By taking these opportunities, we can guide students towards a heightened metalinguistic awareness that will allow them to make more informed rhetorical choices of their own, and it can help us grow as educators ourselves. Orwell tells us that we can "choose—not simply *accept*—the phrases that will best cover the meaning, and then switch round and decide what impression one's words are likely to make on another person" (1946; original emphasis). That double focus of choice and relationality is key to this discussion.

One of the problems my class in 2014 found in the Perez Hilton extract they analyzed was that it did little to convey meaning to the reader. The text's priority appeared to be *producing* language rather than delivering content through language: its main interest lay in the piling of words to reach a volume adequate for a celebrity gossip post. On a broader scale, the tired turn of phrase is symptomatic of a closed system approach to language that has little investment in the reader. Orwell

uses the metaphor of a machine because a text riddled with clichés leaves the reader feeling a lack of relationality, as if there's no person on the other side of the writing.

The question of relationality lies at the heart of theories of language and writing. The computational linguist Emily Bender has become famous for terming LLMs like ChatGPT “stochastic parrots”—stochastic in that they produce language on the basis of probability and chance, but also in the sense of the etymological root word meaning “guess”—and arguing that LLMs operate very differently from humans who produce language, not least because language is fundamentally a relational project (Bender, Gebru, et al., 2021). This focus on relationality equally resounds in rhetoric and writing studies, in which, as Keith Grant-Davie notes, “The roles of the rhetor and audience are dynamic and interdependent” (1997, p. 271). Even the contested notion of writerly “voice” is, in Zak Lancaster’s words, “best understood in dialogic terms, as negotiated through specific discursual interactions” (2019, p. 167).

ChatGPT and other LLMs use natural-language processing that requires significant human intervention in order to sound “natural” and to do the things we want them to do. Reinforcement learning from human feedback (RLHF) is a key component here, in which LLM outputs are ranked by human users. As Helen Toner of Georgetown’s Center for Security and Emerging Technology puts it, “AI is trying to maximize how much the humans will like [the] text it generates, based on what it learned about what humans like” (2023).

The irony, of course, is that this training in “natural language” often results in material that sounds *unnatural* to readers such as John Warner (2022), who describes generated text as “generic, voiceless.” By now, we’ve all been told in various workshops on academic integrity and GPTs to look for the central tell of generated text: its bland, robotic voicelessness. Leif Weatherby (2023) puts it well in *Jacobin*: “What GPT systems spit out is language, but averaged out around a selected center of words. It’s a mush with vague conceptual borders, English (or most any other language) but ironed out and set to the most middling version of itself.” My favourite example of this phenomenon comes from *The Atlantic* article “Welcome to the Golden Age of Clichés” (2023), in which the writer Kaitlyn Tiffany asks ChatGPT to write a “high-school-graduation speech without clichés,” only to be met with the response, “Today marks a significant milestone in our lives.”

The voicelessness, the prefabricated phrasing, the unremarkable prose: this is what it’s like to sound like a bot. And the reason behind this is that LLMs are closed systems, utterly non-situated, producing what Marchetti et al. (2023) call “purely ungrounded language.”

In 2022, I joined the University of Toronto Mississauga, where I teach Writing for University and Beyond, a first-year writing course with a standardized curriculum taught by many faculty members and offered in multiple sections across the disciplines. The first assignment asks students to situate themselves as writers: to identify the context from which they come, thereby setting themselves up for the rest of the term in which they will learn to make more conscious and deliberate choices about the ways that they relate to their readers through writing. This first essay invites them to consider their relationship to the various languages and versions of English they use; the various rhetorical situations in which they find themselves, depending on culture, environment, and upbringing; their linguistic relationships to others, to the academy, and to themselves. This is exactly the kind of interpersonal and relational context that chatbots lack insofar as they represent a closed system based on producing and reproducing language rather than practicing intersubjective exchange.

I propose that postsecondary instructors can use GPT technology as a model for students of what writing is *not*: produced by a mechanism that is utterly unsituated and non-relational, it has the flatness of text that operates from within a closed system, neither dynamic nor interdependent, to use Grant-Davie's terms. ChatGPT might be a classroom tool to demonstrate that the very situatedness of each student writer—the messy, complicated, multifaceted tensions, joys, and beauties that comprise each individual's relationship to language—is an asset rather than a liability. It makes relationality possible. The richness and vibrance of World Englishes, of nonstandard Englishes, of individualized experience and contexts are flattened in a world of GPT-generated text, and the result is a loss rather than a gain.

John Warner (2022) claims that contemporary writing instruction has—to paraphrase Orwell—gone some distance towards turning students into machines: “we have... incentivized them to behave like algorithms, creating simulations that pass surface-level muster.” Of course, the institutional pressures of an increasingly corporatized postsecondary context incentivize instructors to do so. GPT technology represents an opportunity to resist such tendencies by renewing an emphasis on relationality and dialogue in the classroom. Peter Elbow observes that, “When writers change their felt relationship to their readers..., they usually come up instinctively with better wording—and even more effective thinking” (2007, p. 178). LLMs can be a tool to model—and caution against—the way writing sounds when it *fails* to engage in a relationship to alterity, when it produces language for its own sake rather than as a means of relationship.

When using LLMs as a means of foregrounding exchange, relationality, and empathy in student writing, it's useful to remember Dawn Skorczewski's encouragement to recognize “our own clichés”

as instructors, “the marginal comments, and the habits of thought embedded in them that have become so familiar to us that we think of them as common sense” (2000, p. 236). If we want students to become active decision-makers in their own writing, we need to demand the same of ourselves in our response to that writing. The age of chatbots can remind us of the value and even urgency of a dialectic approach to teaching in the style of Martin Buber, assuming the pedagogical environment as one in which human beings encounter each other as human beings rather than as functionaries in a disciplinary system.

A student recently asked me why I can’t provide templates so that students can merely fill in the blanks; his reasoning was that the template would be able to say whatever he needed to say better than he could ever say it. Granted, templates like those in Gerald Graff and Cathy Birkenstein’s *They Say/I Say* (2006) continue to be valuable learning tools for a range of students, with native and non-native speakers benefiting from them. That said, our task as instructors to empower students as decision-makers and independent thinkers entails more than directing them to templates, lest they believe—as this student of mine appeared to—that their own voices and subject positions are a hindrance to communication rather than what makes communication organic and meaningful.

Perhaps counterintuitively, LLMs may be a way to encourage students to theorize writing not just as something to please an instructor, but as a significant reflection of students’ own humanity and relationship to others. Toward this end, I propose that we invite students to think about the things that separate human-produced language and thought from generated text. For instance, we can renew a classroom focus on examining language as it’s grounded in localized, culturally and historically dependent usage; thinking about arguments that are based in concrete specifics rather than bland abstractions; looking at real, shared experiences on a community level. Moving away from the valorization of Standard English as the only acceptable model of academic writing is a start; asking students to consider the dimensions of non-standard word choice and phrasing as it contributes and adds depth to one’s prose serves the double task of alerting students to the breadth of potential expression at their fingertips and highlighting a skill that escapes AI programs.

One exercise I’ve started using in my first-year writing courses is based on a shared, embodied, localized community site of experience: a notoriously dismal on-campus eatery. I show students a series of reviews of the establishment, some taken directly from Google Reviews, and some that I’ve generated myself through prompts that I’ve engineered to elicit creative negative restaurant reviews. Students are quick to identify the human-written reviews from the generated ones because of the inclusion of concrete details (such as mention of a supposedly vegan burger that was suspiciously

bloody) and the niche language of online videogame culture used in one review to warn against the establishment. Another human-written review uses arguably pro-social, inherently relational language in its advice that anyone looking to avoid writing an exam would do well to dine at this restaurant, since doing so will ensure that they will be hospitalized and exempt from any academic undertakings. By contrast, the generated reviews, although descriptive and lengthy, nonetheless rely on hackneyed phrasing (“left much to be desired”) that reads more like filler than descriptions of lived experience. As students identify substance versus semblance of substance, voice versus voicelessness, they become aware of the distinction between writing that is fluent yet banal and writing that is dense with meaning—a richness and vividness achievable in this case by their peers on Google Reviews, who have been unfortunate enough to visit this particular eatery.

On a larger scale, we can renew a focus on relationality in our classroom practices. In my first-year writing courses, capped at an enrollment of 25, I have the ability to give my students personalized feedback on each stage of their writing process. This weekly check-in allows me to learn their verbal idiosyncrasies, their interests, the fingerprint of their thinking, and to guide them in developing their thought and voice. I’ve also started asking students to choose topics for their papers not only on the basis of their own interest, but on the basis of something they think someone in their lives should know about. I want them to think about a reader on the other end of their writing, and I ask them to tell me who that intended reader is.

I find myself wanting to call this type of teaching a luxury—the luxury of knowing my students on an individual basis and giving them personalized attention in their development as writers. But it’s not a luxury: it’s a necessity. If the massive student disengagement in the age of COVID-19 lockdown teaching didn’t teach us that, then the age of chatbots is here to drive the lesson home. The need for teacher/student ratios that afford genuine interaction has grown increasingly urgent if we are to help students value their human, individual voices over generated text. Yet calls for smaller classes and more faculty stand in opposition to what Karen Gravett (2022) calls the “dominant and materializing discourse” that frames the modern university as “entrenched in academic capitalism, distorted by audit culture, governed by managerialism” (p. 22) and in which the student is positioned as a customer instead of a person.

As instructors, our job is to equip students to enter into academic conversations, and our first step—now more than ever—may be instilling in them the confidence to know that their own humanity is an asset in that endeavour rather than a liability. Ironically, by showing us what we are *not*, the bots might just help us rediscover what we are, and what we can be.

References

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *ACM Conference on Fairness, Accountability, and Transparency*. 610-23. <https://doi:10.1145/3442188>
- Elbow, P. (2007). Voice in writing again: Embracing contraries. *College English*, 70(2) 168-188. <https://www.jstor.org/stable/25472259>
- Grant-Davie, K. (1997). Rhetorical situations and their constituents. *Rhetoric Review*, 15(2) 264-279. <https://www.jstor.org/stable/465644>
- Gravett, K. (2022). *Relational pedagogies: Connections and mattering in higher education*. Bloomsbury Press.
- Hargraves, O. (2014). *It's been said before: A guide to the use and abuse of clichés*. Oxford University Press.
- Lancaster, Z. (2019). Tracking students' developing conceptions of voice and style in writing. In A. R. Gere (Ed.), *Developing writers in higher education: A longitudinal study* (pp. 163-184). University of Michigan Press.
- Marchetti, A., Di Dio, C., Cangelosi, A., Manzi, F., & Massaro, D. (2023). Developing ChatGPT's theory of mind. *Frontiers in Robotics and AI* 10 <https://doi.org/10.3389/frobt.2023.1189525>
- Orwell, G. *Politics and the English language*. The Orwell Foundation. <https://www.orwellfoundation.com/the-orwell-foundation/orwell/essays-and-other-works/politics-and-the-english-language/>
- Skorczewski, D. (2000). "Everybody has their own ideas": Responding to clichés in student writing. *College Composition and Communication*, 52(2) 220-239.
- Tiffany, K. (2023, February 21). Welcome to the golden age of clichés. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2023/02/ai-chatbots-cliche-writing/673143/>
- Toner, H. (2023, March 4). *First, some basics of how language models like ChatGPT work: Basically, the way you train a language model is by* [Tweet]. Twitter. <https://twitter.com/hlntnr/status/1632030583462285312>
- Warner, J. (2022, December 5). Freaking out about ChatGPT—Part 1. *Inside Higher Ed*. <https://www.insidehighered.com/blogs/just-visiting/freaking-out-about-chatgpt%E2%80%94part-i>

Weatherby, L. (2023, April 17). ChatGPT is an ideology machine. *Jacobin*.

https://jacobin.com/2023/04/chatgpt-ai-language-models-ideology-media-production?fbclid=IwAR2UD7Z0Ww8clYc4-Pa4LQKpWy7D5Fja_HpOe1_Y84ZE2l8d_gTtI1aXT6Y

Wu, S., & Rubin, D. (2000). Evaluating the impact of collectivism and individualism on argumentative writing by Chinese and North American college students. *Research in the Teaching of English*, 35(2), 148-178.

Zinsser, W. (2006). *On writing well: The classic guide to writing nonfiction*. Harpercollins.