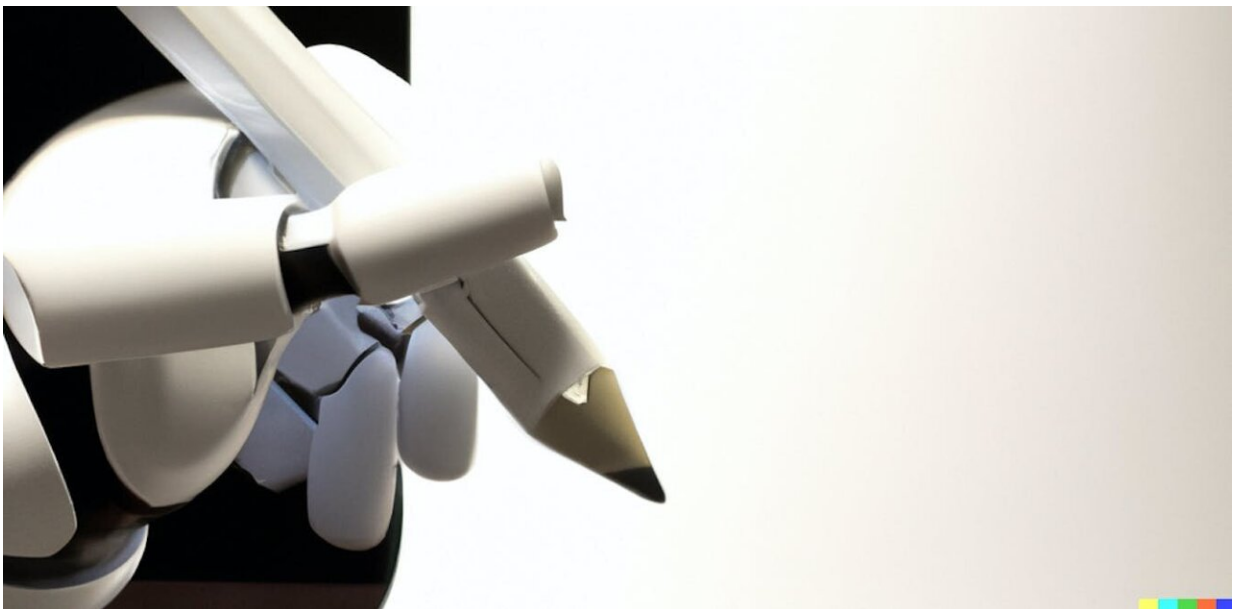# Commentary: ChatGPT threatens language diversity—more needs to be done to protect our differences in the age of AI

February 10 2023, by Collin Bjork

Credit: Collin Bjork, Author provided

The buzz around artificial intelligence (AI) technologies like ChatGPT is palpable. People are both optimistic and frightened by the possibilities of these tools. Clearly, these technologies will change *how* people write. But in terms of *what* people write, these technologies seem to be embracing the status quo.

In fact, the way these tools are currently built appears to homogenize writing—making everything sound the same. And writing that sounds the same is not just boring; it also perpetuates inequity.

When writing tools prioritize one way of writing over another, they reinforce existing hierarchies that unfairly position Standard American English (SAE) and the Queen's English over other languages and ways of writing.

How does ChatGPT work?

Technologies like ChatGPT are called large language models (LLMs). LLMs provide textual responses to human commands, by using machine learning to study patterns of words in a massive archive of texts.

Crucially, however, ChatGPT does not know the meaning of words. ChatGPT generates definitions by sorting through a mountain of definitions and then collating those into a single response that suits the context of a query.

In other words, without meaning as its guide, ChatGPT responds to queries by relying on context clues, stylistic structures, writing forms, linguistic patterns and word frequency.

This functionality means that, by default, ChatGPT perpetuates dominant modes of writing and language use while sidelining less common ones.

Erasing diversity

Dominant modes of writing don't become dominant by accident. They become dominant because one social group wants to assert power over

another social group.

There is not, for example, one kind of English. There are many Englishes.

The decision to prioritize Standard American English in many US classrooms, for example, means that speakers of Black English—a language with its own grammar, lexicon and remarkable history of resistance—are penalized and shamed for writing as they speak.

Similarly, in Aotearoa New Zealand, the Queen's English became dominant not because it's intrinsically better than te reo Māori. Rather, European colonizers wanted to stamp out Māori culture, and writing in the Queen's English became a key tool for furthering that objective. In the 20th century, students were regularly beaten for speaking Māori in schools.

Going against the default

Supporters of ChatGPT will be quick to note that ChatGPT can read, analyze and generate content in many languages, including in Black English and te reo Māori.

But the concern is not about what ChatGPT can do.

It's about what its default settings are. It's about how ChatGPT is configured to treat some forms of writing as normal, typical and expected. And it's about how ChatGPT requires a special request to generate non-normative forms of writing.

This problematic default behavior also occurs in ChatGPT's sister program, Dall-E 2. This image-generating AI was asked to create an image for this article based on this prompt: "close up photo of hands

typing on a laptop." The program created four images. All had white masculine hands.

The program needed a more specific prompt to generate an image that included a person of color because even the ways that AI visualizes writing is dominated by [white men](link).

Ultimately, this kind of algorithmic bias continues to make white English-speaking men the standard of writing culture, while ushering everyone else to the margins.

How did it get like this?

It's no surprise that ChatGPT's default functionality seems to prioritize forms of English writing developed by [white people](link). White English-speaking men have long dominated many writing-intensive sectors, including journalism, law, politics, medicine, computer science and academia.

These white English-speaking men have collectively written billions of words, many times more than their colleagues of color. The sheer volume of words these authors have written means that they likely constitute the majority of ChatGPT's learning models, even though ChatGPT's parent company, [OpenAI](link), doesn't publicly reveal its source material.

So when users ask ChatGPT to generate content in any of these disciplines, the default output is written in the voice, style and language of those same white English-speaking men.

Challenging the norm

Some people will say that we need defaults and standards in writing. They argue that we need to teach people to write in the Queen's English or SAE so that people don't miss out on jobs and promotions because they write in a different way.

But that line of thinking just means capitulating to workplace prejudice and reinforcing an unjust system through our participation in it. Instead, other scholars say we need to challenge those unfair writing standards and encourage writers to embrace the rich rhetorical possibilities in their linguistic diversity.

Educators who want to embrace linguistic diversity might be tempted to ban text-generating AI from their schools and universities.

But it's worth remembering that writing itself is a technology that has been, and still is, used to further inequality. Literary scholar Alice Te Punga Somerville calls this "the inextricability of writing from historical and ongoing violence.."

In response to this threat, however, Professor Somerville does not advocate abandoning writing altogether. Rather, she insists on using the tool of writing critically and creatively to resist oppression.

Taking her lead, educators might instead encourage students to develop new ways of deploying these tools to compose a more equitable future. Doing so means, as Professor Vershawn Young says in "Black English", "that good writin gone look and sound a bit different than some may now expect. And another real, real good result is we gone help reduce prejudice."

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Citation: Commentary: ChatGPT threatens language diversity—more needs to be done to protect our differences in the age of AI (2023, February 10) retrieved 22 June 2024 from https://phys.org/news/2023-02-commentary-chatgpt-threatens-language-diversitymore.html