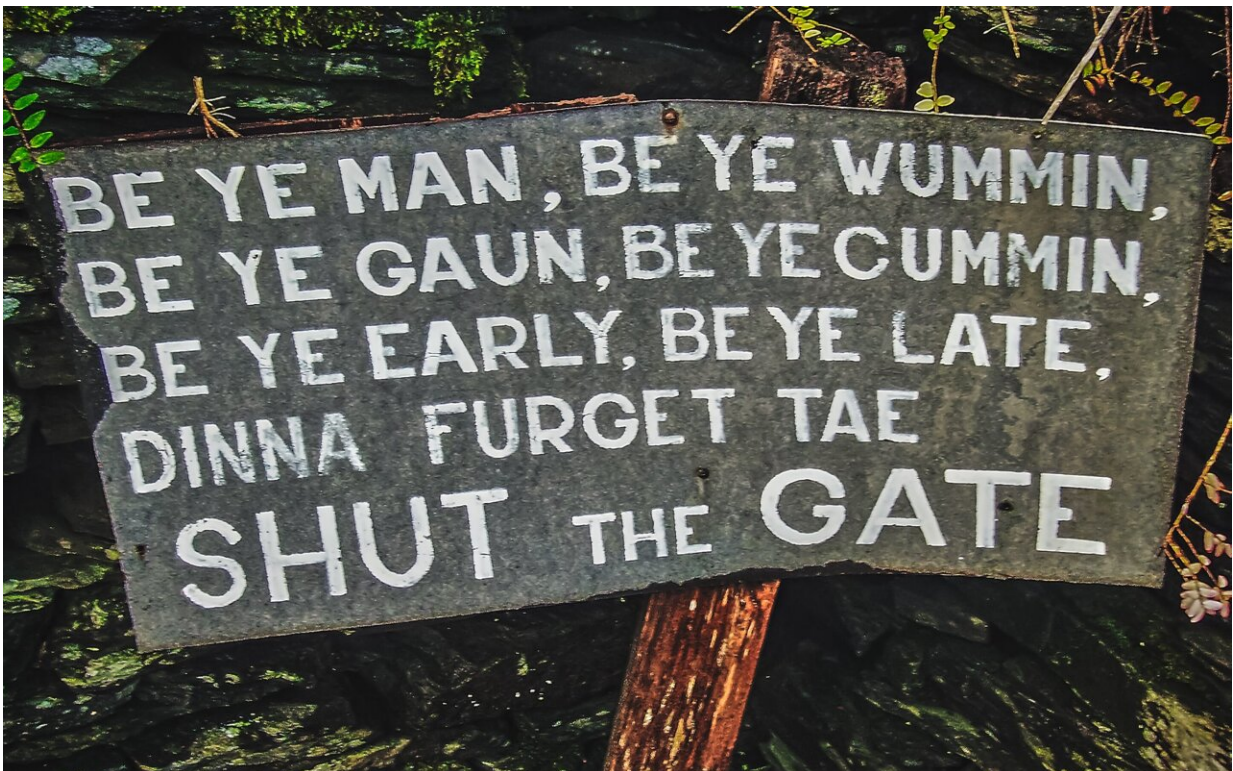


Why AI software 'softening' accents is problematic

January 11 2024, by Grégory Miras



Credit: Unsplash/CC0 Public Domain

["Why isn't it a beautiful thing?"](#) a puzzled Sharath Keshava Narayana asked of his AI device masking accents.

Produced by his company, Sanas, this recent technology seeks to

"soften" the accents of call center workers in [real-time](#) to allegedly shield them from bias and discrimination. It has sparked widespread interest both in the [English-speaking](#) and [French-speaking world](#) since it was launched in September 2022.

Far from everyone is convinced of the software's anti-racist credentials, however. Rather, critics contend it plunges us into a [contemporary dystopia](#) where technology is used to erase individuals' differences, identity markers, and cultures.

To understand them, we could do worse than reviewing what constitutes an [accent](#) in the first place. How can they be suppressed? And in what ways does ironing them out bend far more than [sound waves](#)?

How artificial intelligence can silence an accent

"Accents" can be defined as a set of oral clues (vowels, consonants, intonation, etc.) that contribute to the more or less conscious elaboration of hypotheses on the identity of individuals (e.g. geographically or socially). An accent can be described as regional or foreign according to different narratives.

With start-up technologies typically akin to [black boxes](#), we have little information about the tools deployed by Sanas to standardize our way of speaking. However, we know most methods aim to at least partially transform the structure of the sound wave in order to bring certain acoustic cues closer [to a perceptive criteria](#). The technology tweaks the timbre of certain vowels and consonants and parameters such as rhythm, intonation or accentuation. At the same time, the technology will be looking to safeguard as many vocal cues as possible to allow for the recognition of the original speaker's voice, such as with [voice cloning](#), a process that can result in [deepfake vocal](#) scams. These technologies make it possible to dissociate what is speech-related from what is voice-

related.

The automatic and real-time processing of speech poses technological difficulties, the main one being the quality of the sound signal to be processed. Software developers have succeeded in overcoming them by basing themselves on [deep learning](#), [neural networks](#), as well as [large data bases of speech audio files](#), which make it possible to better manage the uncertainties in the signal.

In the case of foreign languages, Sylvain Detey, Lionel Fontan and Thomas Pellegrini identify [some of the issues inherent in the development of these technologies](#), including that of which standard to use for comparison, or the role that speech audio files can have in determining them.

The myth of the neutral accent

But accent identification is not limited to acoustics alone. Donald L. Rubin has shown that listeners can [recreate the impression of a perceived accent](#) simply by associating faces of supposedly different origins with speech. In fact, absent these other cues, speakers are [not so good at recognizing accents](#) that they do not regularly hear or that they might stereotypically picture, such as German, which many associate with ["aggressive" consonants](#).

The wishful desire to iron out accents to combat prejudice raises the question of what a "neutral" accent is. Rosina Lippi-Green points out that [the ideology of the standard language](#)—the idea that there is a way of expressing oneself that is not marked—holds sway over much of society but has no basis in fact. [Vijay Ramjattan](#) further links recent colossal efforts to develop accent "reduction" and "suppression" tools with the neoliberal model, under which people are assigned skills and attributes on which they depend. Recent capitalism perceives language as

a skill, and therefore the "wrong accent" is said to lead to reduced opportunities.

Intelligibility thus becomes a pretext for blaming individuals for their lack of skills in tasks requiring oral communication according to [Janin Roessel](#). Rather than forcing individuals with "an accent to reduce it", researchers such as [Munro and Derwing](#) have shown that it is possible to train individuals to adapt their oral abilities to phonological variation. What's more, it's not up to individuals to change, but for public policies to better protect those who are discriminated against on the basis of their accent—[accentism](#).

Delete or keep, the chicken or the egg?

In the field of sociology, Wayne Brekhus calls on us to pay specific attention to the invisible, weighing up what isn't marked as much as what is, the "lack of accent" as well as its reverse. This leads us to reconsider the power relations that exist between individuals and the way in which we homogenize the marked: the one who has (according to others) an accent.

So we are led to Catherine Pascal's question of [how emerging technologies can hone our roles as "citizens" rather than "machines"](#). To "remove an accent" is to value a dominant type of "accent" while neglecting the fact that other co-factors will participate in the perception of this accent as well as the emergence of discrimination. "Removing the accent" does not remove discrimination. On the contrary, the accent gives voice to identity, thus participating in the phenomena of humanisation, group membership and even empathy: the accent is a channel for otherness.

If technologies such AI and [deep learning](#) offers us untapped possibilities, they can also lead to a dystopia where dehumanization

overshadows priorities such as the common good or diversity, as spelt out in the [UNESCO Universal Declaration on Cultural Diversity](#). Rather than hiding them, it seems necessary to make recruiters aware of how accents can contribute to customer satisfaction and for politicians to take up this issue.

Research projects such as [PROSOPHON at the University of Lorraine \(France\)](#), which bring together researchers in applied linguistics and work psychology, are aimed at making recruiters more aware of their responsibilities in terms of [biais awareness](#), but also at empowering job applicants "with an accent". By asking the question, "Why isn't this a beautiful thing?" companies like SANAS remind us why technologies based on internalized oppression don't make people happy at work.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Why AI software 'softening' accents is problematic (2024, January 11) retrieved 29 April 2024 from <https://phys.org/news/2024-01-ai-software-softening-accents-problematic.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.