

Automated speech recognition less accurate for blacks: study

March 24 2020, by Edmund L. Andrews



Credit: CC0 Public Domain

The technology that powers the nation's leading automated speech recognition systems makes twice as many errors when interpreting words spoken by African Americans as when interpreting the same words spoken by whites, according to a new study by researchers at Stanford Engineering.

While the study focused exclusively on disparities between black and white Americans, similar problems could affect people who speak with regional and non-native-English accents, the researchers concluded.

If not addressed, this translational imbalance could have serious consequences for people's careers and even lives. Many companies now screen job applicants with automated online interviews that employ [speech recognition](#). Courts use the technology to help transcribe hearings. For people who can't use their hands, moreover, speech recognition is crucial for accessing computers.

The findings, published on March 23 in the journal *Proceedings of the National Academy of Sciences*, were based on tests of systems developed by Amazon, IBM, Google, Microsoft and Apple. The first four companies provide online speech recognition services for a fee, and the researchers ran their tests using those services. For the fifth, the researchers built a custom iOS application that ran tests using Apple's free speech recognition technology. The tests were conducted last spring, and the speech technologies may have been updated since then.

The researchers were unable to determine whether the companies' speech recognition technologies were also used by their virtual assistants, such as Siri in the case of Apple and Alexa in the case of Amazon, because the companies do not disclose whether they use different versions of their technologies in different product offerings.

"But one should expect that U.S.-based companies would build products that serve all Americans," said study lead author Allison Koenecke, a doctoral candidate in computational and mathematical engineering who teamed up with linguists and computer scientists on the work. "Right now, it seems that they're not doing that for a whole segment of the population."

Unequal error rates

Koenecke and her colleagues tested the [speech recognition systems](#) from each company with more than 2,000 speech samples from recorded interviews with African Americans and whites. The black speech samples came from the Corpus of Regional African American Language, and the white samples came from interviews conducted by Voices of California, which features recorded interviews of residents of different California communities.

All five speech recognition technologies had error rates that were almost twice as high for blacks as for whites—even when the speakers were matched by gender and age and when they spoke the same words. On average, the systems misunderstood 35 percent of the words spoken by blacks but only 19 percent of those spoken by whites.

Error rates were highest for African American men, and the disparity was higher among speakers who made heavier use of African American Vernacular English.

The researchers also ran additional tests to ascertain how often the five speech recognition technologies misinterpreted words so drastically that the transcriptions were practically useless. They tested thousands of speech samples, averaging 15 seconds in length, to count how often the technologies passed a threshold of botching at least half the words in each sample. This unacceptably high error rate occurred in over 20 percent of samples spoken by blacks, versus fewer than 2 percent of samples spoken by whites.

Hidden bias

The researchers speculate that the disparities common to all five

technologies stem from a common flaw—the machine learning systems used to train speech recognition systems likely rely heavily on databases of English as spoken by white Americans. A more equitable approach would be to include databases that reflect a greater diversity of the accents and dialects of other English speakers.

Unlike other manufacturers, which are often required by law or custom to explain what goes into their products and how they are supposed to work, the companies offering speech recognition systems are under no such obligations.

Sharad Goel, a professor of computational engineering at Stanford who oversaw the work, said the study highlights the need to audit new technologies such as [speech](#) recognition for hidden biases that may exclude people who are already marginalized. Such audits would need to be done by independent external experts, and would require a lot of time and work, but they are important to make sure that this technology is inclusive.

"We can't count on companies to regulate themselves," Goel said. "That's not what they're set up to do. I can imagine that some might voluntarily commit to independent audits if there's enough public pressure. But it may also be necessary for government agencies to impose more oversight. People have a right to know how well the technology that affects their lives really works."

More information: Allison Koenecke et al. Racial disparities in automated speech recognition, *Proceedings of the National Academy of Sciences* (2020). [DOI: 10.1073/pnas.1915768117](https://doi.org/10.1073/pnas.1915768117)

Provided by Stanford University

Citation: Automated speech recognition less accurate for blacks: study (2020, March 24)
retrieved 2 May 2024 from

<https://techxplore.com/news/2020-03-automated-speech-recognition-accurate-blacks.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.