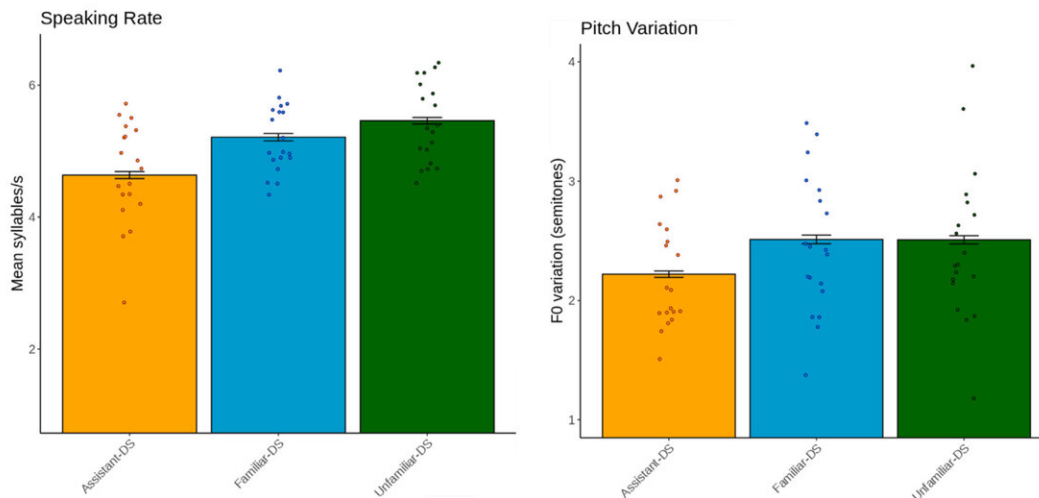# Study explores how African American English speakers adapt their speech to be understood by voice technology

April 30 2024

African American English speakers adjust rate and pitch based on audience. Credit: Michelle Cohn, Zion Mengesha, Michal Lahav, and Courtney Heldreth

Interactions with voice technology, such as Amazon's Alexa, Apple's Siri, and Google Assistant, can make life easier by increasing efficiency and productivity. However, errors in generating and understanding

speech during interactions are common. When using these devices, speakers often style-shift their speech from their normal patterns into a louder and slower register, called technology-directed speech.

Research on technology-directed speech typically focuses on mainstream varieties of U.S. English without considering speaker groups that are more consistently misunderstood by technology. In *JASA Express Letters*, researchers from Google Research, the University of California, Davis, and Stanford University wanted to address this gap.

One group commonly misunderstood by voice technology are individuals who speak African American English, or AAE. Since the rate of automatic speech recognition errors can be higher for AAE speakers, downstream effects of linguistic discrimination in technology may result.

"Across all automatic speech recognition systems, four out of every ten words spoken by Black men were being transcribed incorrectly," said co-author Zion Mengesha. "This affects fairness for African American English speakers in every institution using voice technology, including health care and employment."

"We saw an opportunity to better understand this problem by talking to Black users and understanding their emotional, behavioral, and linguistic responses when engaging with voice technology," said co-author Courtney Heldreth.

The team designed an experiment to test how AAE speakers adapt their speech when imagining talking to a voice assistant, compared to talking to a friend, family member, or stranger.

The study tested familiar human, unfamiliar human, and voice assistant-directed speech conditions by comparing speech rate and pitch variation. Study participants included 19 adults identifying as Black or African

American who had experienced issues with voice technology.

Each participant asked a series of questions to a voice assistant. The same questions were repeated as if speaking to a familiar person and, again, to a stranger. Each question was recorded for a total of 153 recordings.

Analysis of the recordings showed that the speakers exhibited two consistent adjustments when they were talking to voice technology compared to talking to another person: a slower rate of speech with less pitch variation (more monotone speech).

"These findings suggest that people have mental models of how to talk to technology," said co-author Michelle Cohn. "A set 'mode' that they engage to be better understood, in light of disparities in speech recognition systems."

There are other groups misunderstood by voice technology, such as second-language speakers. The researchers hope to expand the language varieties explored in human-computer interaction experiments and address barriers in technology so that it can support everyone who wants to use it.

Citation: Study explores how African American English speakers adapt their speech to be understood by voice technology (2024, April 30) retrieved 21 May 2024 from