TCSJ
TECHNICAL COMMUNICATION
& SOCIAL JUSTICE

Technical and Professional Communicators as Advocates of Linguistic Justice in the Design of Speech Technologies

Halcyon M. Lawrence, Towson University

Technical Communication & Social Justice, Vol 2, No. 1 (2024), pp. 1-22

Keywords: speech recognition, speech technologies, bias, linguistic representation, linguistic justice

Abstract: Despite claims of being a revolutionary technology, research demonstrates that speech technologies, broadly, and automatic speech recognition systems, specifically continue to demonstrate listening (recognition) bias against languages and dialects spoken by people of color, foreign speakers of English, and marginalized communities. Yet despite evidence of bias, listening devices are increasingly being used in US schools, prisons, courts of law, and workplaces such as call centers in India and the Philippines: spaces all disproportionately represented by people of color and foreign English speakers. The paper reframes the on-going conversation around linguistic representation in speech technologies as an urgent linguistic justice issue by highlighting the sociopolitical contexts in which these devices are used. I argue that technical communicators—both researchers and practitioners—are in an ideal position to advocate for a more socially just design of speech devices and to assess and mitigate potential harms to marginalized communities with varying language backgrounds.

Author Bio: Dr. Halcyon Lawrence transformed technical communication research, teaching, and practice by highlighting the importance of linguistic justice in the design of speech technologies. An award-winning leader, teacher, and practitioner from Trinidad and Tobago, Dr. Lawrence will continue to shape how technical communicators take up their responsibility to design accessible and equitable technologies. Her work can be accessed on her website: http://www.halcyonlawrence.com/.

© Halcyon Lawrence, *Technical Communication & Social Justice* Vol. 2, No. 1 (2024), pp. 1-22.

Problem Statement

For years, speech technologies have been described as revolutionary in the media and by companies (Miller, 2020; Webb, 2018) often without real discussion about their potential harm or embedded biases. That trend is changing as more journalists of technology for whom English is not a first or only language are writing about these uneven, biased interactions with speech devices (Rangarajan, 2021; Félix, 2021; Lloreda, 2020). Additionally, research is beginning to support these anecdotal accounts. For example, research findings show that personal assistants like Google Home perform best with midwestern and western accents, and Amazon's Echo perform best with southern and eastern US accents. Both devices perform poorly for speakers with Indian, Chinese, or Spanish accents (Harwell, 2018).

The uneven experiences of people of color and speakers of "non-standard[1]" English with accents might not be immediately apparent. On the one hand, the lack of guidance and standards for how speech is designed and used technology might just be experienced as an inconvenience for some (e.g., a misheard name over an airport's PA system); on the other hand, the growing use of speech technologies had exposed inequity, bias, and limited accessibility for speakers of African American English[2] (Martin, 2022; Koenecke, 2021).

Despite the growing use of speech technologies and the increasing documentation of potential harm and bias, the design of speech (and sound) as a medium of communication remains under-examined in the field of technical communication and information design (Hocutt, 2021). For example, a search for the terms of "speech technology" or "voice user interface" in our discipline's major journals and conference proceedings[3] (between 2000 to 2022) yielded approximately 17 articles (four journals returned no results for the search terms). The need to address the lack of representation of language and accent varieties in speech technologies is a growing social justice issue as voice—and the data it produces—is becoming a medium of surveillance and policing in prisons, schools, and banking (Gillum and Kao, 2019). In fact, Safiya Noble (2017) predicted, "artificial intelligence (AI) will become a major human rights issue in the 21st century.

This paper provides technical communication practitioners and teachers with an orientation to the linguistic justice issues that the design of speech devices raises along with a discussion of the

[1] I wish to acknowledge how problematic and harmful terms like non-standard and non-native are in discussions about language and accents. These terms are wholly colonial ways of talking about language, which center white, Western belief systems. I encourage you to read work by Cheng et al. (2021) who suggest alternate ways of characterizing language experience and use specifically for research purposes (e.g., primary vs. secondary language). See also April Baker Bell's (2020) intentional use of "White Mainstream English " instead of "standard English". However, you will see use of the term standard English and similar terms in this paper as I refer to studies that have used these terms.
[2] African American English (AAE) is defined as a "linguistic variety predominantly used by African Americans and in places where African Americans live." (Hudley et al., 2022)
[3] Proceedings for SIGDOC, IEEE ProComm, STC Summit; *Technical Communication Quarterly* (TCQ), *Technical Communication; Journal of Business and Technical Communication* (JBTC); *Business and Professional Communication Quarterly*; *Journal of Business Communication*, *Computers and Composition*; and *KAIROS: Rhetoric, Technology, Pedagogy*. There have been articles that have been written for special collections (e.g., Martin and Schultz, 2012; Lawrence, 2019; Lawrence, 2021) and researchers in our field have published outside of our traditional TPC journals and proceedings (e.g., Kim et al., 2007).

sociopolitical contexts in which they are used. I ask why and for whom this work matters and why technical communicators are well suited for this work, despite our relatively low research engagement with this type of communication technology. I end with an urgent call to action to the field of technical communication to advocate and provide guidance in the design of speech technologies.

Positionality Statement

Growing up in Trinidad and Tobago, I did not think about accents in the same way that I do now having lived in the United States for 15 years. Speaking the "Queen's English" was and is still an expectation of an educated citizen—even after 60 years of independence. Like many in Trinidad and Tobago, I grew up bi-dialectical, switching between the not-so-mother tongue of England and our distinctive creole influenced by English, French, Spanish, and Hindi.

In 2008, I moved to the US to pursue my graduate studies. I lived on the south-side of Chicago for nearly 6 years where my black skin allowed me to move undetected as a foreigner until I opened my mouth to speak. Once I realized that my British-inflected English was welcomed, I used my accent to negotiate the culture of anti-Blackness I was beginning to witness daily in America. I can recount being followed around in a convenience store, being surveilled from afar, but once I opened my mouth to speak, my perceived threat would dissipate. This was the first time in my life that I understood that accent was not only a marker of identity, but a tool of translation I could use—on my terms—to navigate unwelcoming anti-Black spaces.

In a respectful communication exchange, there is a beautiful linguistic negotiation that goes on between speakers of different languages or with different accents. It's like a dance. There are linguistic and facial cues that tell you that your conversational partner didn't quite get "it" and you begin to think about other ways, and other words, that you could use so they might understand you better. Yet, despite many of these "tricky" exchanges, I had never felt the expectation put upon me to change my accent to be understood. The translation process happened on my terms. That was until I started interacting with speech devices.

My daily human interactions in the last 15 years in the US did not mirror the same linguistic experiences I was having in digital spaces. Calling a bank and getting an automated teller would be a frustrating experience, as my letter [a] (as in apple] was often heard as [e] (as in echo). There was simply no room to dance with the device when it could not "understand" me. Even if an American listener didn't understand me, there was always room for negotiation. Yet in the digital sphere, none of my clear speech strategies worked with these devices: slower, louder, and hyper-articulated speech was often mistaken by digital listeners as angry or distorted speech. In many digital spaces, my accent was not understood. Engaging with speech technologies often robbed me of my ability to engage with translation practices on my terms and instead disciplined my tongue, forcing "[my mouth to] move in ways that felt foreign or strange" (Cooley and Gonzales, 2023, p. 2).

In 2010, while I was working on my doctorate in Chicago, I received a panicked call from my mother who lives in Trinidad. She had just received a call from the credit card company indicating unusual activity on my account—my credit card had eight transactions! Since I was not actively using the card, and concerned that my account had been compromised, I made an international call (not inexpensive for a student) to my local bank's credit card center. The first thing that struck me was that I was greeted with an automated speech recognition (ASR)[4] system that used an American English accented speaker. When I did eventually talk with a live agent, they indicated that there was in fact, "a" single transaction on my account, not the "eight" that my mom misheard. Having taken a series of linguistics courses as part of my PhD program, it was immediately clear to me how my mother could have acoustically and phonetically confused "eight transactions" (eɪt trænˈzækʃənz) with "a transaction" (ə trænˈzækʃən) especially on a system that did not speak our variant of English. It occurred to me then that this was a design problem, much like the design problems we identify and address in technical communication for written and visual media. For example, I understood this to be an intelligibility problem that might have been avoided if better design decisions had been made about word choice (for example, the voice interface could have been designed to say "one transaction" instead of "a transaction") and about which English accent would have produced higher rates of speech intelligibility[5] for local listeners.

This encounter led to my dissertation work; in it, I examined the perception and intelligibility of non-standard and non-native accented speakers of English. I wanted to answer if, beyond the documented negative perceptions of non-standard and non-native English speech,[6] there were speech-mediated contexts in which these accents were both intelligible and positively perceived (Lawrence, 2013). To me, these questions were analogous to the questions technical communicators asked about the legibility and perception of fonts in written and visual texts (Redish & Selzer, 1985; Brumberger, 2004).

Despite my interest in this topic, as a field, we weren't talking as widely about speech and sound interfaces in the same way that textual and visual design was part of our disciplinary concern. As Cynthia Selfe (2009) had earlier argued within rhetorical studies, "our contemporary adherence to alphabetic-only composition constrains the semiotic efforts of individuals and groups who value multiple modalities of expression (p. 616)." The same was (and is still) true for the field Technical and Professional Communication (TPC); however, as a young academic, I lacked the skill (and perhaps confidence) to articulate how and why this work should similarly matter to our field.

Given the focus of this special issue, and the exigencies articulated in the problem statement, this paper is an effort to reframe my ongoing work on bias in speech-mediated communication within the context of linguistic justice. Moreover, it is an opportunity for us as a field to examine the ways in which speech technologies can mediate, support, hinder, or obscure translation practices of speakers. As Natasha Jones challenges us, "at this point in history, scholars concerned with the social, economic, and political implications of their work must now consider ways to critique, intervene in, and create communicative practices and texts that positively impact the mediated experiences of individuals" (2016, p. 345).

---

[4] I define this and other terms in appendix A at the end of the paper.
[5] Speech intelligibility refers to the clearness of a speech signal. It is not to be confused with comprehensibility (i.e., whether what is said is understood by the listener), but instead, what is said, is clearly perceived by the listener.
[6] For example, non-native speakers are perceived negatively; perceived to be less intelligent and less loyal, and less competent. A summary of these perceptions is provided in Table 2.3 in Lawrence (2013).

To this end, I begin with a literature review that offers some insight in the sociopolitical contexts in which speech technologies are used locally and internationally, and the potential harm that marginal groups of English speakers (AAE, foreign-accented, non-standard accented, for example) face because of the uneven representation of their languages and dialects in speech devices.

Literature Review

Speech technologies can be defined as any device using computing technology that recognizes, analyzes, or produces speech. The technology is used in a range of devices including automatic tellers, tele-help systems, personal assistants, global positioning systems (GPS) to name a few. This range of use and number of users of speech technology devices continues to grow globally; for example, by 2026, the global digital assistant (e.g., Apple Siri or Google Home) market is expected to reach 33.771 billion USD from 4.793 billion USD in 2019. Such technologies are used in business and industry (e.g., automated tellers); tertiary education (e.g., automated transcription services); home and personal use (personal assistants, GPS, smart speakers, and IoTs) health industry (e.g., used to generate medical documentation, patient check in, and to communicate with patients with dysarthric speech).

The benefits and challenges of speech as a medium of communication

Researchers have demonstrated that speech as a mode of communication provides several benefits: first, speech is distinct as it is omni-directional (signals can be received from multiple directions); attention- grabbing, and is able to communicate information, even when listeners are not paying attention (Noyes et al., 2006; Rivenez et al., 2004). As an ambient medium, other activities, such as seeing, or writing can be conducted while using speech. Additionally, speech as an input medium can facilitate a range of tasks. Speech is also scalable, both for input/output of data compared to the screen and keyboard as input/output devices. Speech provides a form of natural expression for a wide range of users, and for those with specific disabilities, speech may allow users to engage with technology in ways that textual and graphical interactions do not permit. But speech as a medium also presents a number of challenges as well.

Perhaps one of the greatest challenges of speech is that it is ephemeral, so speech interactions must be designed with this transience in mind. Additionally, given the very probabilistic nature of speech (i.e., a single idea can be expressed in multiple ways), designing interactions that capture the breadth of expressions becomes important to ensuring accuracy of input and output.

Possibilities and limitations of automatic speech recognition discussed in TPC

In the field, perhaps one of the earliest articles written about Automatic Speech Recognition (ASR) which discusses its possibilities and limitations appears in the online journal, *Kairos: Rhetoric, Technology, Pedagogy*. In it, Harrison (2000) argues that ASR deserves both our careful attention and debate. While Harrison sees the possibility of ASR transforming the compositionist and the writing that they do, he cautions against any narrative that suggests a deterministic view of ASR. This deterministic view about technology is still heavily critiqued to this day. For example, Broussard (2018) suggests that "the way we talk about technology is out of sync with what digital technology can actually do," (p. 6) and Hutter and Lawrence (2021) argue that even the discourse around the innovation of technologies promotes "an uncritical acceptance that, for the sake of innovation, new technologies could be developed independent of well-articulated problems and well-defined communities of users" (p. 151).

Harrison also asks researchers to resist approaching ASR as a "neutral tool for doing one's bidding." Although Harrison doesn't offer a definition of the term "*neutral*," the term has come to be synonymous with "bias-free" in discussions about artificial intelligence (AI). Since Harrison's writing, there is even more evidence of the bias embedded and reproduced by AI. As recent examples, work by Noble (2021) critiques racial and gender bias in search engines; Buolamwini (2017) focuses on racial bias in facial recognition technologies; Benjamin (2019) discusses how technologies reinforce white supremacy and deepen social inequality; Eubanks (2018) demonstrates how technologies police and punish the poor; and Lawrence (2021) discusses how speech technologies discipline non-native and non-standard speakers of English.[7] The following short cases demonstrate three contemporary contexts in which speech technology is being used. The purpose here is to illustrate the potential and often unexamined harms for people of color or already marginalized communities both locally, in the US and internationally, in countries like India and the Philippines, who live and work in spaces where speech technologies are being used.

Three Cases of Speech Technologies

The use of speech recognition in US schools

In the US, given the resistance of Congress to enact gun control laws, some schools and hospitals administrations have started installing speech surveillance devices as a means to identify public threats to safety.[8] Schools purchase and inconspicuously install microphones using aggression detection software, although testing has found them less than reliable[9] (Gillam & Kao, 2019). Using vocalizations to detect or predict violence is problematic. First, many of these surveillance systems aren't designed to analyze speech but instead analyze the non-verbal markers (such as pitch, volume, and rate) of the vocalization which then is matched to an emotion (angry, happy, sad, etc.). However, it has been repeatedly established that objective markers of emotion and vocalizations are not easily identifiable because some vocal expressions of emotions are both individually and culturally specific (Sauter et al., 2009).

Additionally, African American speech (particularly performed by women) is routinely marked as loud and often misunderstood as aggressive or as "having an attitude" (Koonce, 2012), making this group particularly vulnerable to being targeted for negative emotional speech. As Neville (2022) notes "[w]e can anticipate that the speech and voices of racialized children and youth will be disproportionately misinterpreted as aggressive sounding. This troubling prediction should come as no surprise as it follows the deeply entrenched colonial and white supremacist histories that consistently police what Stoever (2016) refers to as the "sonic color line."

---

[7] See also work by Cathy O'Neil (2016) and Mullaney et al. (2021)

[8] A bipartisan bill of $300 million to increase school security was unveiled in the US Congress in June 2022 to make schools more hardened to threats of violence.

[9] You can listen to actual recordings: https://projects.propublica.org/graphics/aggression-detector-data-analysis

Second, in view of known concerns, researchers are actively working to improve violence detection systems using vocal signals of emotion. However, the research methods used in these types of studies can also be problematic. Very often, for these studies, emotions are simulated (acted or induced) by participants, sometimes even actors. This is done as naturally occurring speech is difficult to capture and is often of poor quality to analyze. Since many of these systems are designed based on emotional speech interpreted by an actor or simulated violence experiments (see Han et al., 2018 as an example), it is quite possible that real emotional speech can be misunderstood by these devices. Finally, speech characteristics (e.g., pitch, volume, etc.) of children is subject to more variability because of the physiological changes children go through as they get older, making it virtually impossible for an ASR detection system not to misunderstand emotions produced by school-aged children. Taken together, the design of many of these systems don't challenge, but in fact, reinscribe bias that is already experienced by communities of color.

The use of automatic speech recognition in US courts

Not only can speech technology reinscribe existing bias, but it can then render that bias invisible. Take for example the use of automatic speech recognition systems in courtrooms (also called digital court reporting). To save millions of dollars in wages, at least 14 states now use these devices for the transcriptions of court proceedings (Jaafari & Lewis, 2019). One stenographer, Christopher Day (personal interview, January 11, 2022), with whom I spoke as part of my on-going research is part of a small, but growing number of stenographers who are not only concerned that the use of ASR in the court system is sidelining their profession, but he also raised the concern that ASR will further marginalize people of color based on the courts system's inexperience with African American English. For example, recent research demonstrated that accuracy for stenographers in the transcription of African American English can be as low as 60% (Jones et al., 2019). Day notes that while he is a part of the New York State Court Reporters Association which is actively trying to redress this through training, he suggests that no such transparency exists for digital transcription as the software is both proprietary and closed source, so it is difficult to assess accuracy rates of digital transcriptions.

While we have as yet no data to support Day's specific concerns, we do know that ASR systems exhibit bias against African American English (Martin, 2022; Koenecke, 2021)[10]. To illustrate a simple transcription error, I asked Amazon's Alexa to make a 6:30 pm appointment to "chat with Pat." As Alexa regularly has challenges understanding my accent, the device saved the appointment as "Sharks with pots" (see figure 1). Rather than marking the text as unintelligible or unclear (which makes the error visible) the recognition feature provides a nonsensical, but readable entry into my calendar based on what it has heard. Again, while the example of this incorrect transcription is mildly inconvenient (and maybe even a little comical), in a court of law, the cost of these errors for defendants of color can be a matter of life or death, as "incoherent and incomplete transcripts make it harder to craft appeals…" (Jaafari & Lewis, 2019).

[10] Also called African American Language.

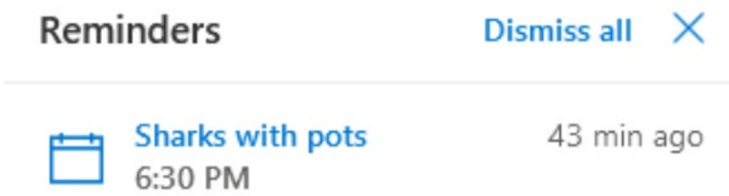

Figure 1: Alexa appointment posted as "Sharks with pots" instead of "Chat with Pat"

Use of speech technology in the workplace (call centers)

In 2022, Sanas, a Silicon Valley startup launched a real time accent translation tool marketed initially to offshore call centers. The founders claimed that the tool would, "revolutionize communication by giving multilingual speakers a choice when it comes to how they communicate." (Sanas, 2022).

Essentially, overseas workers on a help call could—by the press of a button—have their non-American accents replaced by an American accent. The company raised 32M USD in venture capital, the largest first round of financing ever raised by a speech technology company. Despite the financial success, the backlash was swift (abc7news.com,) as many saw the speech technology as engaging in accent erasure, doing nothing to—as the company claimed— "change the world." But complicating the critique of yet another form of digital whitewashing, was the company's own origin story. Started by four Stanford graduates, who were part of the renowned Stanford Artificial Intelligence Lab (SAIL), the company boasts that 80% of their team are made up of immigrants and draws from the personal experience of one founder, Sharath Narayana who was born in India, and worked at a call center, but refused to go through accent reduction training so as to maintain his identity. As you scroll through the company's website, you are presented with a number of compelling, and perhaps personal taglines which include, "Change the World, not Your Self," "Sound Local, Globally," Your Voice. Your Choice."

Sanas is a problematic technology because as Huatong Sun (2012) suggests, cross-cultural communication often values efficiency over cultural sensitivity. Fundamentally, the designers of this technology have taken what Laura Gonzales (2022) sees as a reductive approach to identifying and addressing the problem experienced by offshore call centers: the accent. If the accent is removed then problem of caller abuse is fixed "magically,"[11] all without needing to address the systemic problems brought about by colonial legacies that should lead us to ask why and by what means have these people come to be speaking English in the first place.

Sanas is a problematic technology because it erases the real work of translation which it claims to engage in. It eliminates glossodiversity (diversity of languages) which Pennycook (2008) argues is necessary for us to take global language diversity seriously. In fact, the erasure of the Filipino or Indian accent does nothing to bring dignity or empower individuals, to advance equality or deepen empathy, as the company's website suggests. It is simply an efficient way to ensure that American customers are no longer frustrated by hearing foreign accents on the other end of their help desk calls.

[11] A term used on the Sanas' website and in interviews with company leadership.

Significant evidence of linguistic bias in Automatic Speech Recognition (ASR) technologies

In 2020, Koenecke and her collaborators at Stanford University published important evidence of racial and linguistic bias in listening technologies. In this study, the researchers were interested to see if black speakers and white speakers were understood accurately by personal assistants (Alexa, Google Home, Watson, Siri, and Cortana) by comparing error rates in transcriptions of interviews of the different speakers. In the study, speakers of African American English (AAE) experienced, on average, a word error rate of 35% compared to 19% for white speakers across all five devices. The study confirmed exactly what many speakers with a non-standard accent (or dialect) already knew: that speech technologies generally, and more specifically personal assistants, exhibit bias.

Similar findings were earlier documented for non-white talkers using Bing Speech and YouTube Automatic Captions (Tatman and Kasten, 2017). More recently, Martin (2022) tested contemporary speech corpora, Deep Speech and Google Cloud Speech and found that the habitual “be”—a syntactic and morphological feature of AAE—was more error prone than the non-habitual “be” used in “standardized” forms of English. Martin and Wright (2023) offer other compelling cases around the use of ASR and the impact of hiring and advancement of African American Language speakers in the workforce and the use of ASR for the transcription of medical documents in healthcare settings. They argue that the already disparate treatment of speakers of African American Language risks exacerbation when ASR systems are employed. Falbo and LaCroix (2021) theorize that one of the reasons for such poor performance of these devices within communities of color is that the design of the technology doesn’t account for cultural code-switching[12] practice. They term this exclusionary design practice, “cultural smothering.” This smothering can have powerful psychological effects. For example, research by Mengesha et al. (2021) demonstrate that African American users of ASR technologies have difficulty having their intent recognized, ultimately leading them to feel like the technology is not made for them, triggering emotions such as frustration, disappointment, and self-consciousness. As a result, most participants in their study reported making linguistic accommodation (speech modification) to be understood. Such linguistic accommodation, Lawrence (2021) argues, is a form of linguistic discipline, not unlike the colonial linguistic practices used to subjugate conquered peoples and erase cultural identity.

Response to the critiques of bias in ASR

One of the responses to these critiques of the absence of Black language representation in personal assistants was to introduce Black celebrity voices[13] such as Samuel L. Jackson and Shaquille O’Neal (for the low, low price of $4.99), which only reinforced long-held stereotypes (and the appropriation and commodification of African American Language) about the usefulness of Blacks to entertain, but nothing more. In 2021, a little over a decade since Siri was launched, the platform offered what many listeners perceived to be “black-sounding” speakers

[12] Code switching is defined as the ability or practice of shifting between linguistic varieties of languages depending on the context, setting, or topic, among other factors (Hudley et al., 2022)
[13] Google Home also had limited time cameos by Issa Rae and John Legend.

(Holliday, 2022)[14]. But even then, the representation of Black voices reinforced long held stereotypes against African Americans. Using the newly released voices, Holliday conducted a study and found that listeners engage in racialized judgments of digital voice assistants (in other words, even when voices are racially ambiguous, listeners will still assign race to a voice). Additionally, Holliday found that these racialized judgments also interact with perceptions about the personality of the devices, not unlike the kind of judgments we make about people in the physical (i.e., non-digital) world. For example, in the study, listeners rated the Black-sounding voice as less competent and less professional than the other voices—the same stereotype experienced by African American males in America.

What the bias in ASR tells us about industry

The bias embedded in speech technologies demonstrates industry's preference for standard English over other variants of English. While it is estimated that there are over 160 different English dialects spoken in the world, by 2021, Amazon Alexa, Google Home and Apple Siri supported a combined seven English dialects (e.g., British, American, Australian, New Zealand, Irish, Canadian) (Lawrence, 2021). Lawrence (2021) argues that decisions to support so few English dialects are driven by which Englishes are deemed profitable (in the case of Hinglish or Singaporean English) or are considered prestige Englishes, or what Mufwene (2001) describes as the legitimate offspring of English. Given that English is the Lingua Franca, this limited representation of world Englishes in speech technologies is unfortunate, as currently, there are more non-native speakers of English than there are native speakers of English.

Part of the challenge of identifying and addressing language bias is that it's very much still internalized as an acceptable form of bias. When people are discriminated against on the basis of language, the underlying bias is harder to identify. Discrimination based on race, gender, and disability, all present obvious means of detection. Language bias is veiled, harder to identify, and often not understood by the listener themselves. For example, it is perfectly acceptable to express a preference for one's own accent or native language, but it is seen as problematic to express a preference for one's own race or gender. We make judgments about people all the time based on their use of a language (intelligence, professionalism, competence, etc.), yet it would not be acceptable to make similar judgments based on one's race or gender. As a result, linguistic bias is insidious and easily replicated in speech technologies, when design and interface decisions are made to (1) have devices produce speech with a standard language or dialect or (2) recognize only standard accented speakers and discipline/ignore/misunderstand speakers whose language patterns don't conform to that standard.

What can be done to address representation and support translanguaging in ASR?

To increase the representation of languages using speech technology requires vast amounts of voice data for machine learning (ML) to happen. On the one hand, many of the early speech technologies were based on a corpus of midwestern accented speakers, so the collection of voices to support other languages and their regional dialects is an expensive endeavor. On the

[14] Apple has since confirmed that these were indeed Black voice actors. To hear these four voices, you can go to: https://www.consumerreports.org/digital-assistants/apples-new-siri-voices-resonate-with-some-black-iphone-users-a5978242346/

other hand, there is evidence that speech technologies can be designed to support translanguaging practices like Hinglish, a hybridized language in which speakers move between Hindi and English. Martin and Shultz (2012) point to the usefulness of speech technologies, more generally, in supporting the preservation of "dying languages" or languages with few speakers, despite the relative unprofitability that this language support would create.
In 2019, I suggested that we might look to open-source projects such as Mozilla's Common Voice project as a source of voice data for the development of non-proprietary software (Lawrence, 2019); however, this process has proven to be a lengthy one. For example, started in June 2017, the stated goal for each language is to have 10,000 hours collected and validated to train a speech-to-text system (Mozilla Common Voice, 2017). To date, while there are an impressive ninety-six languages represented in the corpora, none of these languages have reached the collection goal, with English having collected a little over three thousand hours. Additionally, for the collected sentences, validation (having 2-3 speakers confirm the sentence) is slow.

When I first threw my support behind Mozilla's Common Voice project in 2019[15], I was hopeful about what the open-source community could do with access to a large and diverse corpora of voice data. Reading about people's excited responses to hearing "their voice" on these devices seemed to be adequate reasons to support such initiatives. Today, I am less optimistic about the outcomes and increasingly aware of the potential harms. Decisions to support the collection of voice data for the sake of representation is increasingly at odds with the need to protect already vulnerable communities. Increased representation means that listening devices such as emotion detections now have a more diverse database of Black and Brown voices to draw from.

Holliday's (2022) work has demonstrated that the same biases and perceptions get replicated in the digital realm. Additionally, asking communities of color—often uncompensated—to participate in projects that reduce their voices into data points presents opportunities for further stigmatization and exploitation. Language has always been used strategically by oppressed communities to navigate that oppression and in turn redefine identity. Making their translation practices explicit by lending their voices to these initiatives, might also reduce their ability to use their language to navigate oppression. As Cooley and Gonzales (2023) point out, "[t]ranslation is also understanding how the same language can undergo metamorphoses in the mouths of the marginalized, giving words new meaning and life as a means of survival amidst the dominance of whiteness" (p. 2). This tension is summed up by Nee and colleagues (2021):

> "Because language invites us to make assumptions about others' identities, it can serve as a tool for propagating harmful bias and discrimination by proxy. At the same time, because language and social reality are mutually reinforcing (meaning our language reflects the world around us and influences how we think and what we do), language can also serve as a mechanism for advancing social justice" (p. 1).

[15] In an attempt to provide translation for more languages on the web, in March 2022, Facebook's MetaAI launched a crowdsourcing NLP project, "No Language Left Behind." https://ai.facebook.com/research/no-language-left-behind/

More recently, I have had to delay and even set aside research projects to ask more urgent questions: when we support the "capture" [16] of the voices of minority communities, how do they benefit, and what possible harms can we account for? In the pursuit of linguistic justice, questions about what we advocate for and what projects we support must be decided with care. We cannot simply reduce the problem of linguistic justice to an issue of language representation, as we run the risk as Agboka fears that we may in fact, "otherize" or recolonize users, if we don't consider the "ideology, power, economics, knowledge, politics, law, and ethics all as dimensions of a locale" (2013, p. 29). Therefore, the work that we do in TPC requires a special care as "technical communicators are not only designers of information but are also transmitters, translators, or articulators of cultural values" (Agboka, 2013, p. 31). Since we too run the risk of bringing harm to the very communities we purport to help, centering the community of users means an explicit and intentional focus on mitigating marginalization (Jones et al., 2016). Given these concerns, what then is a way forward for TPC to pursue linguistic justice in speech technology research?

Discussion

Given the exigencies raised in the research findings discussed above, technical communicators are in an ideal position to advocate in the design of speech technologies, especially considering the increasing role of technical communicators as agents of accessibility, social justice, and change (Jones, Moore, & Walton, 2016), and their skill as rhetoricians with expertise in information and user-centered design.

Harris (2001) suggests that technical communicators understand the rhetorical work that "scripting" a voice user interface involves, which includes the work of
"charting paths around a site, anticipating user-agent scenarios, preparing for contingencies, developing personalities for the agents, adhering to the context-sensitive principles of natural language, repairing misconstruals—addressing, in short, the massive array of linguistic demands made by (1) taking information meant to be discerned visually, (2) making it truly accessible acoustically, and (3) rendering it vocally interactive" (p. 221).

Additionally, Lawrence (2019) suggests that given our field's work in textual and visual communication, refocusing on the oral/aural is possible. This requires both a conceptual shift (e.g., what does it mean that speech is clear?) and technical shift (e.g., how is that clarity measured?) since many of the concepts in speech design have analogous concepts in textual and visual design.

Despite the obvious valuable role we can play, Harris (2001) in their review of the book, *Designing Effective Speech Interfaces* by Weinschenk and Barker (2000) laments about the often secondary (and too late) contributions that technical communicators make to these speech interaction design initiatives:

> "[c]ue the technical communicators, right? Well, nope, not yet. Nobody seems to have given much thought to who should script the interfaces…Rest assured, you'll be getting

[16] I use capture here both in the technical sense of recording and in the philosophical sense to refer to the ways in which colonialism was built on the physical possession of Black and Brown bodies.

> the calls. Users will revolt, companies will panic, and someone will yell, "Bring in the writers!" But wouldn't it be nice if they realized before they hit the marketplace that the people to put the chat in Cathy [the software] are the professional wordsmiths?" (p. 221).

While I agree that Harris is correct that we can make significant contributions as wordsmiths, the work goes far beyond our contribution as skilled users of words, and I address these contributions fully in the next section.

I also see a clear role for technical communicators, and have pointed out that to do this work, technical communicators might need to acquire new skills including familiarity with the International Phonetic Alphabet (IPA) system, phonological and syntactic rules of a language, technical knowledge with speech analysis software, and range of research methodologies, including scientific research methods (Lawrence, 2021). Finally, Hocutt (2021) also suggests that if we are to engage with this work, technical communicators need to understand how search results are generated and communicated through speech devices.

Charting a Way Forward for TPC: A Research Framework for Pursuing Linguistic Justice in Speech Technology Research

Drawing from the ongoing work on social justice in TPC, Johanna Phelps (2021) applies a transformative paradigm from Mertens (2007) to illustrate how a framework built on the tenets of axiology, ontology, epistemology, and methodology might allow us to pursue socially just research as we

> "… grapple with demands to illustrate validity and reliability, address calls for replicability, and incorporate or address empirical practices to further establish our discipline while simultaneously honoring Indigenous, decolonial, and participatory methodologies that push back against these demands" (Phelps, 2021 p. 204).

Mertens' transformative paradigm situates research as both etiological (addressing the origin or cause) and teleological (addressing the purpose). Following Phelps, I engage with Merten's four tenets to raise a number of broad questions for speech technology research in TPC. In the tables below, I draw from both Phelps and Mertens to define each tenet and indicate what concerns each tenet addresses. Given these concerns, I raise what I think are relevant questions for speech technology research vis-à-vis each tenet:

Axiology	
How is it defined?	• A theory of ethics or principles of value for research, which categorically influences WHAT we investigate (Phelps p. 207)
What concerns does it address?	• Addresses standards of rigor: credibility, transferability, and dependability (Phelps p. 208). • Addresses the portability and durability of research (Phelps p. 208).

What questions are raised for speech technology research?	• How will communities of color and marginalized (by race, language, socio-economic status, etc.) communities benefit from speech technology research? • How will communities of color be further disadvantaged by this research? • Is the research engaging with social justice and ethical theories? • Where and how will this research be disseminated? • Are there ways in which findings might be used that are not initially intended by the researcher? • What does the community want and need from this research? • What should happen when the community no longer needs or wants this work?
• **Ontology**	
How is it defined?	• A theory of our perception and understanding of what "is". (Phelps p. 208)
What concerns does it address?	• Recognizes and addresses competing and multiple social realities. • Acknowledges which realities (and experiences) are valued over others.
What questions are raised for speech technology research?	• In speech technology design, whose realities are privileged and whose are decentered? • Who is making decisions on behalf of communities? • Who is being consulted in the research process? • How are power dynamics being replicated in the design of speech technologies? • What is being defined (e.g., nativeness, fluency, and intelligibility) and by whom?
Epistemology	
How is it defined?	• A theory of how we come to know what we know and how we make knowledge (Phelps p. 209)
What concerns does it address?	• Allows shareholders and partners to articulate issues of power and privilege.
What questions are raised for speech technology research?	• In what way is this research historically and socially situated (e.g., does the research adequately consider the

	broader issues of how language is used to discipline speakers)? • Who is designing the research? How can this research be designed to be emancipatory? How can this research be designed to be participatory? • What do all parties gain or benefit from this research (e.g., can this research help preserve linguistic identity within communities)? • What is the researcher's positionality to the research? What tensions does this positionality create? (e.g., does the researcher speak the language? What are the researcher's biases about this language and its speakers?)
Methodology	
How is it defined?	• The systematic and concordant selection, design, and use of methods to suit an investigation of a particular set of problems (Phelps 2021, p. 210).
What concerns does it address?	• Methodological integrity: Are these methodologies aligned to the axiological, ontological, and epistemological goals of the project?
What questions are raised for speech technology research?	• Which communities are represented by the chosen methodologies? • What research methods are historically used in this type of research? Do they in fact reinforce colonial ideologies? • Are communities being asked to contribute to research without articulated benefit or risks? • Is the language of informed consent accessible to participants? • What languages are being used to collect, analyze, and report data and findings?

What can we advocate for and provide guidance on?

Given the questions raised by applying the tenets of the transformative paradigm, there are a number of ways that technical communicators can advocate for linguistic justice in the design of speech technologies. Here, I draw on the work of Julia Nee and her colleagues (2021), highlighting some recommendations that align with technical and professional communication competencies:

Advocating about language use in speech interface design MUST include:

1. Flagging of and use of better word choices for more inclusive language that avoids linguistic stereotypes (e.g., the Blacks).

2. Avoiding the use of terms in coding that are harmful language in favor of neutral or positive alternatives (e.g., master and slave vs. primary and secondary).
3. Being aware of the contextual and social nature of language (derogatory language used in one context, can be empowering in another context because the terms might be reclaimed).
4. Having a zero-tolerance policy on language that propagates hate.
5. Using active language that assigns agents and their responsibilities to actions for accountability rather than erasure. Using present tense to describe on-going activities and groups.

Advocating about representation MUST include:

1. Supporting the design of systems to allow people to self-identify and are built to respect differences in self-identification (e.g., Latino/Latina v. Latinx or person with a disability v. disabled person). This is especially crucial in the academic community as we can have strong theoretical reasons for using certain language which does not reflect how communities self-identify.
2. Advocating for representation of languages and dialects when beneficial to that community of speakers.
3. Getting involved in dataset curation projects and advocating for the labeling of data that doesn't center whiteness (e.g., descriptors like non-white or non-standard English).

Advocating about socially just methods MUST include:

1. Ensuring that data collected respects and protects the privacy of marginalized communities.
2. Thinking about potential for harmful use of data beyond the initial intention.
3. Engaging with methodologies that are community-based and participatory. Collaborating with members of marginalized groups at all stages of the design and development process is crucial.

Conclusion: A Call To Action

As I sat to write this conclusion, I was alerted to a newly released CFP for a special issue on social justice and digital interfaces. I read excitedly as I thought that it might be another good venue to present my research, but when I got to the list of examples of the kinds of digital interfaces we might consider writing about for the special issue, they were all textual and visual in nature.[17]

Journal editors: maybe this is a place to begin for our field. How can we make explicit and intentional calls that underscore that speech/sound mediated communication is as important as the visual and the textual? How can our CFPs underscore that focusing on speech can help us better account for a broader set of languages and translation practices in our research and praxis? The work is urgent.

[17] The editor quickly responded to my inquiry indicating that work on speech/sound interfaces were certainly welcomed!

Researchers: how do we expand our research agenda to account for a broader set of languages and by extension a broader set of linguistic experiences of users, including that of Indigenous languages? The work is urgent.

Program directors: what kinds of courses are part of your curriculum that can get students thinking about linguistic justice and speech-mediated technologies? Does your curriculum offer courses in linguistics, translation, or intercultural communication, for example? Are there courses in the curriculum that center sonic literacies? The work is urgent.

Instructors: how can we broaden our syllabi and readings to elevate the elements of speech and sound interfaces alongside textual and visual ones? What projects can we engage our students with that help them to think about the contexts in which speech devices silence and discipline users, and by extension languages? How can user experience instruction expand to think about sonic interfaces and experiences? The work is urgent.

Practitioners: who are you insisting participate and contribute to this developmental work? What skills do you need to develop to provide guidance and recommendations on the design of speech technologies? The work is urgent.

Technical communicators: speech technologies and speech mediated interfaces will continue to develop without us; so why not with us? To it we bring our design expertise, our concern for social justice, our rhetorical sensitivities, and our field's commitment to empower vulnerable communities of users. The work is urgent.

Appendix A: Glossary of terms

Term	**Definition**
Accent	A manner of pronunciation with other linguistic levels of analysis (grammatical, syntactical, morphological and lexical) more or less comparable with the standard language (Gluszek and Dovidio, 2010 p. 215).
Artificial Intelligence	Refers to the simulation or mimicry of human-intelligence by machines.
Automated speech recognition (ASR)	Refers to the capability of a program to process human speech and transliterate it into a written format.
Discrimination	Specific behaviors or actions directed at a group, or its individual members based solely on the group membership (Ng 2007)
Machine Learning (ML)	Refers to the process by which computers learn and make predictions and draw inferences.
Natural Language Processing	A subfield of AI which helps machines understand and process human language
Native language	A speaker's first language/s
Non-standard accent	Speech produced by native speakers of a language yet considered to be non-standard because of pronunciations associated with speakers' region, socio-economic status, ethnicity, caste or social class; for example, the British Cockney accent.
Standard English/Language	A form of English/Language "accepted" as a correct, widely used form.
Speech intelligibility	Speech intelligibility refers to the clearness of a speech signal. It is not to be confused with comprehensibility (i.e., whether what is said is understood by the listener), but instead, what is said, is clearly perceived by the listener.
Speech technology	Any device using computing technology that recognizes, analyzes, or produces speech

References

abc7news.com (2022). *Co-founder of Silicon Valley startup says accent-removing technology is not racist*. ABC7 San Francisco. (2022, August 25). Retrieved February 13, 2023, from https://abc7news.com/sanas-voice-technology-silicon-valley-startup-accent-remover-translator/12162646/

Agboka, G. (2013). Participatory localization: A social justice approach to navigating unenfranchised/disenfranchised cultural sites. *Technical Communication Quarterly*, *22*(1), 28-49.

Agboka, G. (2014). Decolonial methodologies: Social justice perspectives in intercultural technical communication research. *Journal of Technical Writing and Communication, 44*(3), 297–327. doi:10.2190/TW.44.3.e

Baker-Bell, A. (2020). *Linguistic justice: Black language, literacy, identity, and pedagogy*. Routledge.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new jim code*. Polity Books.

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world.* MIT Press.

Brumberger, E. (2004). The rhetoric of typography: Effects on reading time, reading comprehension, and perceptions of ethos. *Technical Communication*, *51*(1), 13-24.

Buolamwini, J. (2017). Gender shades: Intersectional phenotypic and demographic evaluation of face datasets and gender classifiers [Doctoral dissertation, Massachusetts Institute of Technology]. MIT. https://www.media.mit.edu/publications/full-gender-shades-thesis-17/.

Cheng, L., Burgess, D., Vernooij, N., Solís-Barroso, C., McDermott, A., & S. Namboodiripad. (2021). The problematic concept of native speaker in psycholinguistics: Replacing vague and harmful terminology with inclusive and accurate measures. *Frontiers in Psychology*, *12*, 1-22. doi: 10.3389/fpsyg.2021.715843

Cooley, S., & L. Gonzales. (2023). Introduction to special issue on social justice and translation in technical communication. *Technical Communication and Social Justice*, *1*(1), 1-7.

Crystal, D. *English as a global language*. 2003. Cambridge University Press, Cambridge.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Falbo, A., & T. LaCroix. (2021). Est-ce que vous compute? Code-switching, cultural identity, and AI. *arXiv preprint arXiv:2112.08256*.

Félix, N. (2021, November 15). Hey, Siri: Accent bias in speech recognition technology. Distilled Periodical. Retrieved August 2, 2022, from https://www.yaledistilled.com/post/hey-siri-how-many-times-do-i-have-to-repeat-myself-before-i-simply-decide-to-do-it-myself

Gillum, J. & J. Kao. (2019). Aggression detectors: Unproven, invasive surveillance technologies schools are using to monitor students. Pro-Publica. Accessed October 15, 2019. https://features.propublica.org/aggression-detector/the-unproven-invasive-surveillance-technology-schools-are-using-to-monitor-students/

Gluszek, A., & J. Dovidio. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, *14*(2), 214-37.

Gonzales, L. (2022). *Designing multilingual experiences in technical communication*. University

Press of Colorado.
Han, T., Zhang, J., Zhang, Z., Sun, G., Ye, L., Ferdinando, H., ... & S. Yang. (2018). Emotion recognition and school violence detection from children's speech. *EURASIP Journal on Wireless Communications and Networking*, *2018*(1), 1-10.
Harris, R. (2001). Book reviews. *Technical Communication*, 48(2), 218-22.
Harrison S. (2000). Cyborgs and digital soundwriting. *Kairos: A Journal of Rhetoric, Technology, and Pedagogy*, *5*(1). Accessed July 21, 2022.
Harwell, D. (2018). “The Accent Gap.” *The Washington Post*, 19 July 2018, https://www.washington post.com/graphics/2018/business/alexa-does-not -understand-your-accent/.
Hocutt, D. (2021, October). Interrogating Alexa: Holding voice assistants accountable for their answers. In *Proceedings of the 39th ACM International Conference on Design of Communication* (pp. 142-50).
Holliday, N. (2022, January 6-9). “Siri, you’ve changed: Acoustic properties and racialized judgments of voice assistants” [Poster presentation]. LSA Annual Meeting. Washington DC.
Hudley, A., Mallinson, C., & M. Bucholtz. (2022). *Talking college: Making space for Black language practices in higher education*. Teachers College Press.
Hutter, L., & H, Lawrence. (2021, October). The discourse of technological innovation: A new domain for accountability. In *The 39th ACM International Conference on Design of Communication* (pp. 151-56).
Jaafari, J., & N. Lewis. (2019, February 14). In court, where are Siri and Alexa? The Marshall Project. Retrieved August 2, 2022, from https://www.themarshallproject.org/2019/02/14/in-court-where-are-siri-and-alexa
Jones, N. (2016). The technical communicator as advocate: Integrating a social justice approach in technical communication. *Journal of Technical Writing and Communication*, *46*(3), 342-61.
Jones, T., Kalbfeld, J., Hancock, R., & R. Clark. (2019). Testifying while Black: An experimental study of court reporter accuracy in transcription of African American English. *Language*, *95*(2), e216-e252.
Jones, N., Moore, K., & R. Walton. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, *25*(4), 211-29.
Kim, L., McCauley, T., Polkosky, M., D'Mello, S., Craig, S., & B. Nikiforova. (2007, March). Finding information and finding locations in a multimodal interface: A case study of an intelligent kiosk. In *Proceedings of the Second IASTED International Conference on Human Computer Interaction* (pp. 111e117).
Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., ... & S. Goel. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, *117*(14), 7684-89. https://doi.org/10.1073/pnas.1915768117
Koonce, J. (2012). Oh, those loud Black girls!”: A phenomenological study of Black girls talking with an attitude. *Journal of Language and Literacy Education*, *8*(2), 26-46.
Lawrence, H. (2013). *Speech intelligibility and accents in speech-mediated interfaces: Results and recommendations*. [Unpublished doctoral dissertation]. Illinois Institute of Technology.
Lawrence, H. (2019). BEYOND THE GRAPHIC USER INTERFACE: Speculations on the

Future of Speech Technology and the Role of the Technical Communicator. In S, Sundvall (Ed.), *Rhetorical Speculations* (pp. 226-48). Utah State UP.

Lawrence, H. (2021). Siri disciplines. In Mullaney, T., Peters, B., Hicks, M., & K. Phillip (Eds.), *Your Computer Is On Fire* (pp, 179-98). MIT Press.

Lloreda, C. (2020, July 5). Speech recognition tech is yet another example of bias. *Scientific American*. Retrieved August 2, 2022, from https://www.scientificamerican.com/article/speech-recognition-tech-is-yet-another-example-of-bias/

Martin, J. (2022). Automatic speech recognition systems, spoken corpora, and African American language: An examination of linguistic bias and morphosyntactic features [Doctoral dissertation, University of Florida]. Proquest. https://www.proquest.com/docview/2676147443

Martin, J., & K. Wright. (2022). Bias in automatic speech recognition: The case of African American language. *Applied Linguistics 44*(4), 613-30. doi.org/10.1093/applin/amac066

Martin, J., & T. Schultz. (2012). Multimodal and speech technology. In A. Mehler & L. Romary (Eds.), *Handbook of technical communication* (pp. 189-254). De Gruyter Mouton. https://doi.org/10.1515/9783110224948.189

Mengesha, Z., Heldreth, C., Lahav, M., Sublewski, J., & E. Tuennerman. (2021). "I don't think these devices are very culturally sensitive." Impact of automated speech recognition errors on African Americans. *Frontiers in Artificial Intelligence*, *4*, 169.

Mertens, D. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research*, *1*(3), 212–25. doi:10.1177/1558689807302811

Miller, E. (2022, February 7). How voice technology will drive business communications in 2022, *Forbes Technology Council post*. Retrieved February 15, 2023, from https://www.forbes.com/sites/forbestechcouncil/2022/02/04/how-voice-technology-will-drive-business-communications-in-2022/?sh=301baf1a1562

Mozilla Common Voice. Common Voice. (n.d.). Retrieved August 5, 2022, from https://commonvoice.mozilla.org/

Mufwene, S. (2001). *The Ecology of Language Evolution*. Cambridge UP. doi:10.1017/CBO9780511612862.005

Mullaney, T., Peters, B., Hicks, M., & K. Phillip (Eds.). (2021). *Your computer is on fire*. MIT Press.

Nee, J., Macfarlane Smith, G., Sheares, A., & I. Rustagi. (2021). Advancing social justice through linguistic justice: Strategies for building equity fluent NLP technology. Equity and Access in Algorithms, Mechanisms, and Optimization Conference, New York (pp. 1-9). https://dl.acm.org/doi/pdf/10.1145/3465416.3483301

Neville, S. (2022, July 14). 'Hey Siri': Virtual assistants are listening to children and then using the data. *EconoTimes*. Retrieved August 2, 2022, from https://www.econotimes.com/Hey-Siri-Virtual-assistants-are-listening-to-children-and-then-using-the-data-1637396

Ng, S. (2007). Language-based discrimination: Blatant and Subtle Forms. *Journal of Language And Social Psychology, 26*: 106-22.

Noble. S. (2021) Your robot isn't neutral. In Mullaney, T., Peters, B., Hicks, M., & K. Philip. (Eds.). *Your computer is on fire*. MIT Press. pp. 199-212.

Noble, S. (2017, 1 June). "The role of artificial intelligence in society: The next human rights issue of the 21st Century." Safiyanoble.com. https://safiyaunoble.com/role-artificial-intelligence-society-next-human-rights-issue-21st-century/. Accessed 28 Mar. 2022.

Noyes, J., Hellier, E., & J. Edworthy. (2006). Speech warnings: a review. *Theoretical Issues in Ergonomics Science*, *7*(6), 551-71.

O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Pennycook, A. (2008). English as a language always in translation. *European Journal of English Studies*, *12*:1, 33-47. doi: 10.1080/13825570801900521

Phelps, J. (2021). The transformative paradigm: Equipping technical communication researchers for socially just work. *Technical Communication Quarterly*, *30*(2), 204-15.

Rangarajan, S. (2021, February 23). Hey Siri-why don't you understand more people like me? *Mother Jones*. Retrieved August 2, 2022, from https://www.motherjones.com/media/2021/02/digital-assistants-accents-english-race-google-siri-alexa/

Redish, J., & J. Selzer. (1985). The place of readability formulas in technical communication. *Technical Communication, 32*(4), 46-52.

Rivenez, M., Darwin, C., & A. Guillaume. (2004, July). Perception of unattended speech. Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia.

Rubin, D., Healy, P., Gardiner, T., Zath, R., & Moore, C. (1997). Nonnative physicians as message sources: Effects of accent and ethnicity on patients' responses to AIDS prevention counseling. *Health Communication*, *9*(4), 351-68.

Sanas. (2023). Breaking barriers - one conversation at a time. Retrieved February 13, 2023, from https://www.sanas.ai/

Sauter, D., Eisner, F., Ekman, P., & S. Scott. (2009). Universal vocal signals of emotion. In *31st Annual Meeting of the Cognitive Science Society (CogSci 2009)* (pp. 2251-55). Cognitive Science Society.

Selfe, C. (2009). The movement of air, the breath of meaning: Aurality and multimodal composing. *College Composition and Communication 60*(4), 616-63.

Sun, H. (2012). *Cross-cultural technology design: Creating culture-sensitive technology for local users*. Oxford UP.

Stoever, J. *The Sonic Color Line: Race and the Cultural Politics of Listening*. NYU Press, 2016.

Tatman, R., & C. Kasten. (2017, August). Effects of talker dialect, gender & race on accuracy of Bing speech and YouTube automatic captions. *Proc. Interspeech 2017*, 934-38, doi: 10.21437/Interspeech.2017-1746

Webb, D., (2018, January 11). On the brink of a speech revolution. Talk Technologies. Retrieved February 15, 2023, from https://talktech.com/on-the-brink-of-a-speech-revolution/

Weinschenk, S., & D. Barker. (2000). *Designing effective speech interfaces*. John Wiley & Sons.