Considerations of Voice Design for Cross-Cultural Robotic Identity

Jaye Nias
Spelman College
Department of Computer Science
Atlanta, Georgia, USA
jnias@spelman.edu

ABSTRACT

In recent years, the computing community has been focusing their efforts in understanding how algorithmic biases impact or amplify current societal biases. And in a world that is increasingly becoming mediated through computer and algorithmic engagements, it is important that at all levels of design and consideration – algorithmic fairness is appropriately enacted. When specifically considering voice-mediated interactions, whether robotic or agent-based, there are still many factors that amplify challenges that subsets of English language speakers face in their daily lives. This position paper presents an examination that challenges current axiomatic dives into the nuances of human-robotic interaction to consider the epistemic realities where current resources create challenges to already under served users from with diversified vernacular considerations.

CCS CONCEPTS

• **Human-centered computing → User centered design**.

KEYWORDS

human-robotic interaction, cultural linguistics, algorithmic bias, cultural design

1 INTRODUCTION

A 2018 study of African English speakers who used voice based assistive devices found that while many found their assistants useful for interaction, many more found that there were thresholds in the amount of understanding that the majority be gathered from the device [1]. More than half believed that the device did not understand them at all. And when asked what could make their devices better, the majority pointed to the need for more accent recognition and also the inclusion and recognition of specific non-English phrases. These limitations are enacting a form of exclusion that is

problematic in AI based systems. Digital discrimination is the form of discrimination in which automation enacted by AI agents or intelligent AI based systems, treat users unfairly, unethically or just differently based on their personal data [2]. The facilitators of this workshop pose following questions: *Philosophy* - **What is artificial identity?** *Affect* - **How are artificial identities delivered via emotional expression of agents?** *Design* - **How do an agent's voice and speech convey emotions in constructing its artificial identity?** For much of this paper, we will reference the terms culture and race interchangeably. For continuity, we will utilize the definition of culture as presented by Borgman (1992) - "Culture includes race and ethnicity as well as other variables and is manifested in customary behaviors, assumptions and values, patterns of thinking and communication style" [3]. We will reference cultural and racial constraints that impact voice- based agent interactions based on the authors experiences in robotic interaction design and personal lived experiences. It is from an anti-discriminatory design praxis that we approach these questions with seemingly more questions than answers. And in consideration of a constructive, intersectional lens, we hope to offer a culturally and socially inclusive discourse on the future of robo-identity.

2 IDENTITY

What is artificial identity? The facilitators of this workshop propose the idea that identity is rooted in a sense of *belonging* and *expression of emotions* and *information.* While the intended focus of this position would be from that of the human perception of robotic identity. The authors seek to challenge both the human-robotic perceptions and the robotic-human constructions that facilitate appropriate engagement. We will address factors that impact the robot's ability to understand and perceive human identity in a way that is sensitive to the cultural and social nuances of the human agent. This paper will address several cultural considerations and limitations to voice based agents though the parameters of belonging, emotional expression and information.

2.1 Belonging: Human Identity in Culturally Responsive HRI

Race and culture scholars have found that names are deeply rooted in person-hood and power [12]. It is not uncommon for initial human-human engagement to begin by identifying their personal nomenclature. This can be seen mirrored in many languages and cultures. "My name is", is often one of the first phrases taught to learners of a new language. Our names are often one of the most public and personal forms of identity. Further, in circumstances

where the humans first introduction to a name is not audible, communicants will commonly ask the correct pronunciation as a sign of respect in discourse.

Research has shown that robots that regard humans using their name can have improve engagements[5]. But what becomes of voice-based technologies who improperly infer name pronunciation based on linguistic patterns of a dominant group within the society. When voice-based AI are unable to construct or learn proper naming conventions it further silences the representation and identity of people from a variety of ethnic or cultural backgrounds [4]. Names are culturally and socially significant parts of one's identity. In some cultures, a person's name can represent:

- their family lineage as passed on or in remembrance of those beloved who have passed on,
- the inferred characteristics of the child (one who brings joy to the house, wealth is her crown, he comes to end the tears),
- spiritual artifacts from one's culture, community, or spiritual system – (named after the home country of an immigrant)
- cultural naming conventions based on the circumstances of birth (day of the week, order of children, time of the year, cauled, breech, during a festival, after the death of a parent).

As the use voice-based systems become more pervasive, this issue only has the option to be exacerbated. We will more often need to reference those in our lives. For example, consider when a user prompts an agent to "Compose a text message to Ominike". In this example the name, pronounced oh-me-knee-kay, is going to be translated as ohm-in-ike. As a result, the user needs to determine whether to change their pronunciation of the name to create recognition by the agent, or to edit the name in their system to something different. Either response echoes the systematic erasure that is prevalent for many who have diverse cultural names in our society [13]. Further, in some tonal languages the emphasis or tonal quality of the name generated by the agent could change the meaning entirely.

Users who wish to construct a mental model of person-hood identity of a voice based agent would need to model that sense of identity on cultural and personal values exhibited by others in the society. The agent's expression of misnaming someone only seeks to increase the gulf of evaluation that creates emotional connections to artificial agents. Additionally, lack of cultural considerations in areas like nomenclature, could further exacerbate cultural stigma that enact historically and socially discriminatory experiences for users.

3 EMOTIONAL EXPRESSION

How are artificial identities delivered via emotional expression of agents?

Many prominent psychologists have theorized about the similarities of emotional features across cultures [6]. And while cultural emotional expression can have many similarities, emotional regulation across cultures differs[9]. Emotional expression in language can vary across many ethnic and cultural,gender and age groups. When talking to a friend within a West African Drum and Dance community it was stated, "Have you ever noticed how the sounds of their instruments mimic the tonal quality of their language?" Whether the high pitched statacco rhythms of the sabar drum that

are said to sound like "rain on a tin roof" or the deep rounded notes of the Atumpan Drum. It can't be concealed that the languages, Wolof and Akan can be felt in the tonal qualities of the language. A study on the affective language of African-Americans vs Euro-Americans showed that there are significant differences in the level of affect between the two groups [11]. As early as 1982 researchers studied and found that there were negative social effects of 'sounding Black' [8]. Black women, highlighted here as an intersectional example, struggle with notions of emotional responses being perceived as angry. Tone variations is something that is deeply cultural [10]. It should be considered that emotion through voice alone can be potentially misconstrued, due to systemic bias, by AI systems and also mirrored by voice-based agents.

4 INFORMED DESIGN CONSIDERATIONS

How do an agent's voice and speech convey emotions in constructing its artificial identity?

In considering the how of the prompt above, we look to how we can reconstruct more inclusive design methods for voice based agent interactions for robotic identity. We present the following questions and considerations as theorized by Aran, et. al.[2]:

- Consideration Discrimination and Human Biases: Cultural decentering allows designers to embrace cultural diversity by asking the principle question: How can we create a robust and enriching agent based experience for all members of a community or society?"
- Consideration for Bias in Language: Human Centered Design methods allow for expression and exploration of the nuances of cultural linguistic and social considerations when designing new applications.
- Understanding and Attesting AI Biases: In formal research methods,the limitations of a study should be made available to readers who may wish to utilize or further test their methodologies. In the same way, many in the AI community are calling for transparency in methods through discrimination aware AI [7] by developing AI based systems that denote potential discriminatory features and outcomes.

In addition to anti-bias design considerations. Voice based systems need to be inherently agile about multilingual capabilities. It is not enough to have systems that can switch from wholly one language to another. The identities of humans are intersectional and should provide agency in engagement style and practice.

REFERENCES
[1] Lameck Mbangula Amugongo. 2018. Understanding What Africans Say. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3180301
[2] Xavier Ferrer Aran, Jose M Such, and Natalia Criado. 2019. Attesting biases and discrimination using language semantics. *arXiv preprint arXiv:1909.04386* (2019).
[3] Christine L Borgman. 1992. Cultural diversity in interface design. *ACM SIGCHI Bulletin* 24, 4 (1992), 31.
[4] Mary Bucholtz. 2016. On being called out of one's name. *Raciolinguistics: How language shapes our ideas about race* (2016), 273–289.
[5] Xiangyu Chen and Andrew Williams. 2020. Improving Engagement by Letting Social Robots Learn and Call Your Name. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 160–162.
[6] Hillary Anger Elfenbein and Nalini Ambady. 2002. On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin* 128, 2 (2002), 203.

[7] Xavier Ferrer, Tom van Nuenen, Jose M Such, Mark Coté, and Natalia Criado. 2021. Bias and Discrimination in AI: a cross-disciplinary perspective. *IEEE Technology and Society Magazine* 40, 2 (2021), 72–80.

[8] Fern L Johnson and Richard Buttny. 1982. White listeners' responses to "sounding black" and "sounding white": The effects of message content on judgments about language. *Communications Monographs* 49, 1 (1982), 33–49.

[9] Batja Mesquita, Dustin Albert, et al. 2007. The cultural regulation of emotions. *Handbook of emotion regulation* 2 (2007), 284–301.

[10] Robert J Podesva. 2016. Stance as a window into the language-race connection. *Raciolinguistics* (2016), 203–220.

[11] Randall G Rogan and Mitchell R Hammer. 1998. An exploratory study of message affect behavior: A comparison between African Americans and Euro-Americans. *Journal of Language and Social Psychology* 17, 4 (1998), 449–464.

[12] Betsy Rymes. 1996. Naming as social practice: The case of Little Creeper from Diamond Street. *Language in society* 25, 2 (1996), 237–260.

[13] Xian Zhao and Monica Biernat. 2017. "Welcome to the US" but "change your name"? Adopting Anglo names and discrimination. *Journal of Experimental Social Psychology* 70 (2017), 59–68.