
Eschewing Gender Stereotypes in Voice Assistants to Promote Inclusion

Andreea Danielescu

Accenture Labs

San Francisco, CA 94105, USA

andreea.danielescu@accenture.com

Abstract

The wide adoption of conversational voice assistants has shaped how we interact with this technology while simultaneously highlighting and reinforcing negative stereotypes. For example, conversational systems often use female voices in subservient roles. They also exclude marginalized groups, such as non-binary individuals, altogether. Speech recognition systems also have significant gender, race and dialectal biases [1,14,15]. Instead, there is an opportunity for these systems to help change gender norms and promote inclusion and diversity as we continue to struggle with gender equality [21], and progress towards LGBTQ+ rights across the globe [11]. However, prior research claims that users strongly dislike voices without clear gender markers or misalignments between a voice and personality [10]. This calls for additional research to understand how voice assistants may be designed to not perpetuate gender bias while promoting user adoption.

Author Keywords

Conversational interfaces; inclusion and diversity; LGBTQ+; personality design

CSS Concepts

• **Human-centered computing**~**Human computer interaction (HCI)**; *Interaction paradigms*; Natural language interfaces;

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

© 2020 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-6819-3/20/04.

DOI: <https://doi.org/10.1145/3334480.XXXXXXX>

**update the above block & DOI per your rightsreview confirmation (provided after acceptance)*

Introduction

UNESCO's 2019 report, entitled "I'd Blush If I Could," claims that voice assistants propagate harmful gender biases, such as reinforcing that women should be in subservient roles [19], while media coverage and research continue to argue that tech companies need to do better. To spark a discussion around gender stereotypes and to promote inclusion of LGBTQ+ communities in the design of conversational systems, Copenhagen Pride, Virtue, Equal AI, Koalition Interactive & Thirty Sounds Good teamed up to create Q, a genderless voice aimed at ending gender bias in AI assistants [20]. In Q's words: "I'm created for a future where we are no longer defined by gender, but rather how we define ourselves." However, prior research has found that "an ambiguous voice is classified as strange, dislikable, dishonest and unintelligent." Similar penalties are applied to systems where the user perceives a misalignment between the gender they assign a voice and the gender they assign a personality [10], raising questions about whether this research is outdated, or whether a non-binary voice would be rejected by users of conversational systems. To complicate things further, cultural norms can vary significantly by country and language and many gendered languages do not support non-binary references and speech.

Conversational systems do not just propagate biases through the gender and personality of the assistant, but biases are also engrained in the automatic speech recognition (ASR) models that are used. While many believe that ASR is a solved problem in conversational systems and we can achieve 95% accuracy or better for male speakers with standard dialects, research shows that recognition accuracy for women is reduced

by about 13% [1,14]. Accented speech suffers further penalties, with accuracy rates in the 50-80%.

Reducing gender bias in voice assistants is worth pursuing, but it raises a number of practical, cultural and ethical questions that we need to address first. From a practical point of view, how does a gender-neutral voice affect adoption and how can we design a gender-neutral personality to match this voice? Culturally, how do we design for languages that are heavily gendered? And ethically, can a gender-neutral voice assistant truly change gendered norms, or will it ultimately disquiet users instead? These are just a few of the questions that we need to ask while pursuing this goal.

User Preference and Gender

Research has found that people have an overall preference towards female voices across cultures and genders. Clifford Nass, author of *Wired for Speech* and *The Media Equation*, notes that people can discern the gender of the voice within seconds [10] and that women's voices tend to be higher pitched while men's tend to be deeper. In an interview with CNN, Nass said: "It's much easier to find a female voice that everyone likes than a male voice that everyone likes. It's a well-established phenomenon that the human brain is developed to like female voices" [5]. More recently, Mitchell and his colleagues conducted studies in 2008 with university-aged students in the Midwest and found that female voices were perceived as "warmer." They also found that although both genders say that they prefer women's voices, only women actually held a subconscious preference for them [9].

Echoing gender stereotypes, Nass also found that people tend to perceive female voices as helpers or assistants who are helping us solve our problems, while male voices are viewed as authority figures who tell us the answers to our problems. In addition, he also found that women's speech includes more personal pronouns (I, you, she), while men's uses more quantifiers (one, two, some more).

Research with robots found that users ascribed male or female gender to a robot based on the function it was meant to perform. Robots programmed to perform security work were viewed as male, while the same robot programmed for guidance (i.e. giving directions to passersby) was viewed as female [17]. In another study, Trovato and colleagues found that the form of the robot also influences perception. A robot with a straight torso or large shoulders was viewed as male, while those with more curves were viewed as female [18]. This was consistent across cultures.

While some of this research highlights biological differences between sexes, much of this research instead mirrors gender stereotypes and cultural norms that have been shifting over the past couple of decades. Nass's work, although extensive and well-known, was conducted in the 1990s and early 2000s. Social norms change over time, and this research was conducted several decades ago, meaning it may be outdated. However, more recent research has not shown significant progress when it comes to which gender users ascribe to assistants. This opens up the following questions:

- Why is progress on gender equality not mirrored in voice assistants and robots?

- Would prolonged exposure to voice assistants that do not echo gender norms change user perception over time?

Personality in Popular Voice Assistants

Regardless of whether or not a designer explicitly designs for a personality, people will attribute one to it, strongly guided by societal norms and expectations. While anthropomorphism varies by individuals and cultures, we all have the inclination to anthropomorphize non-human objects and creatures. This is especially true for those things that we can talk to. When we anthropomorphize, we also tend to ascribe associated social expectations. Meaning that we view:

- a system that interrupts you as rude
- a system that mispronounces a common word as stupid
- a system that doesn't remember what you just said as forgetful

Additionally, integrating anthropomorphic qualities in the design of a conversational system will encourage users to fall back on existing social norms and expectations. Increasingly, people view their interactions with intelligent assistants as social interactions, with 41% of people who own a voice assistant saying it feels like talking to a friend or another person [8]. This may also explain why we hear many cases of users insulting voice assistants, especially when it misunderstands or fails to do what the user asked.

Before Siri, Cortana, Alexa or the Google Assistant came to the market, each of their creators designed a personality, some of it inspired by research on user

preferences [9, 14]. When they launched, every single one of the big four tech companies had personal assistants that defaulted to a female voice in US English [2]. All but one has a female name.

In our own work on Radar Pace, a conversational coaching system for running and cycling, we conducted extensive research in all five of the countries in which we launched the product (US, Spain, Italy, France, and Germany). We chose female voices for two of our personalities: English and Spanish, and male voices for the other three. The top priority for our personality was to engender trust from the user. After all, we were developing a product with the goal of changing user's behaviors to make them better runners and cyclists. So when picking a male or female voice, we looked at prevalence of male vs. female coaches in each country (male coaches dominated in every country), well-known coaching personalities, user preference, and social norms and expectations [3]. For example, we had to take into account social hierarchy and collectivist vs. individualist views of a culture, because that would influence adoption, trust and rapport with the virtual coach.

Q: The Genderless Voice

This discussion of gendering in voice assistants leads us to Q, the genderless voice developed through a collaboration between Copenhagen Pride, Virtue, Equal AI, Koalition Interactive & Thirty Sounds Good [20]. The team developed the voice to sit at the intersection of male and female vocal ranges. Male voices are usually pitched between 85 to 180 hertz (Hz), while female ones are between 140 to 255 Hz. In an interview with Reuters, Nis Norgaard, a sound designer at Thirty Sounds Good studio, also mentions that "men

tend to have a 'flatter' speech style that varies in pitch less and they also pronounce the letters 's' and 't' more abruptly" [4] This is consistent with research by Anna Jørgensen, in which she found that while other vocal features are involved in gendering a voice, pitch is the most important feature used to change the perception of a trans individual's gender [6].

The voice was tested by over 4,000 non-binary individuals from Denmark, the UK and Venezuela, half of which said they couldn't tell the gender, while the other half were fairly evenly split between male and female. While this is a great first step in moving towards gender-neutral voices, personality is dictated by more than just the sound of the voice.

Globalization, Cultural Norms and Gendered Languages

While the English language has been adapted to support those that identify as non-binary and gender-fluid by overloading the use of the plural pronouns they/them/theirs, not all languages have. For example, what do you do with that voice when you are creating an experience in a gendered language, such as French? In French, every object or person is either masculine or feminine. And when referring to a group, one male in a group of females automatically makes that group masculine. In addition, there are still professions in French that only have a masculine form. In France, there has been a push to make language more gender neutral, but it's spawned some significant controversy [16]. Much more work needs to be done in these languages to support gender-neutrality. LGBTQ+ communities in countries with gendered languages are adopting their own changes for self-expression, but these changes are not recognized widely yet.

Putting It All Together

When thinking about designing an inclusive assistant, there are so many research questions that are yet unanswered.

- If you stray too far from gender norms will your voice assistant be disliked and suffer from lack of adoption?
- How does one design a gender-neutral personality to match a gender-neutral voice?
- Will context of use and word choice gender a gender-neutral voice?
- What will it take to construct gender-neutral assistant in a gendered language?
- Does prolonged exposure to a gendered assistant that does not propagate gendered stereotypes change user perception over time?

These are just some of the questions we will be exploring over the next 6 months as we create a non-binary TTS voice and look at approaches to reducing bias in existing ASR models. This workshop will be a great venue for fostering future collaborations, sharing our learnings so far, and hearing about additional challenges others in the field are tackling.

Initial adoption of voice assistants was a critical first step to thinking about how we might change social norms with this technology. As Londa Schiebinger, the John L. Hinds Professor of History of Science at Stanford University and Director of the EU/US Gendered Innovations in Science, Health & Medicine, Engineering, and Environment Project, puts it, "We don't know if gendering a robot to meet human expectations will encourage compliance on the part of the humans" [13]. However, she also says "What if we

surprised people? What if we made robots that did not meet human expectations? That would loosen up gender roles in human society...this will influence the user to think about gender roles and norms. This eventually comes back to change gender roles in society."

References

1. Joan Palmiter Bajorek. 2019. Voice Recognition Still Has Significant Race and Gender Bias. *Harvard Business Review*. Retrieved from <https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases>.
2. Tim Capes, Paul Coles, Alistair Conkie, et al. 2017. Siri on-device deep learning-guided unit selection text-to-speech system. *Interspeech*, 4011–4015.
3. Andreea Danielescu and Gwen Christian. 2018. A Bot is Not a Polyglot: Designing Personalities for Multi-Lingual Conversational Agents. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, ACM Press, 1–9.
4. Sonia Elks. 2019. Alexa, are you male or female? "Sexist" virtual assistants go gender-neutral. *Reuters*. Retrieved from <https://www.reuters.com/article/us-global-tech-women/alex-a-are-you-male-or-female-sexist-virtual-assistants-go-gender-neutral-idUSKBN1QS25B>.
5. Brandon Griggs. 2011. Why computer voices are mostly female. *CNN Business*. Retrieved from <https://www.cnn.com/2011/10/21/tech/innovation/female-computer-voices/index.html>.
6. Anna K. Jørgensen. 2016. Speaking (Of) Gender: A sociolinguistic exploration of voice and transgender identity. .
7. Dan Kedmey. 2015. Here's What Really Makes Microsoft's Cortana so Amazing. *TIME*. Retrieved from <https://time.com/3960670/windows-10-cortana/>.

8. Sara Kleinberg. 2018. 5 ways voice assistance is shaping consumer behavior. *Think with Google*. Retrieved from <https://www.thinkwithgoogle.com/consumer-insights/voice-assistance-consumer-experience/>.
9. Wade J. Mitchell, Chin Chang Ho, Himalaya Patel, and Karl F. MacDorman. 2011. Does social desirability bias favor humans? Explicit-implicit evaluations of synthesized speech support a new HCI model of impression management. *Computers in Human Behavior* 27, 1: 402–412.
10. Clifford Nass and Scott Brave. 2005. *Wired for Speech*. MIT Press.
11. Kenneth Roth. 2015. *LGBT: Moving Towards Equality*. .
12. Ben Fox Rubin. 2017. Alexa, be more human. Inside Amazon’s effort to make its voice assistant smarter, chattier, and more like you. *CNET*. Retrieved from <https://www.cnet.com/html/feature/amazon-alexa-echo-inside-look/>.
13. Stanford Engineering Staff. 2019. Londa Schiebinger: Why does gender matter. *Artificial Intelligence, Technology & Society, Transportation & Robotics: The Future of Everything*. Retrieved from <https://engineering.stanford.edu/magazine/article/londa-schiebinger-why-does-gender-matter>.
14. Rachael Tatman. 2017. Gender and Dialect Bias in YouTube’s Automatic Captions. *Ethics in Natural Language Processing*, Association for Computational Linguistics, 53–59.
15. Rachael Tatman and Conner Kasten. 2017. Effects of Talker Dialect, Gender & Race on Accuracy of Bing Speech and YouTube Automatic Captions. *INTERSPEECH 2017*, ISCA, 934–938.
16. Annabelle Timsit. 2017. The Push to Make French Gender-Neutral. *The Atlantic*. Retrieved from <https://www.theatlantic.com/international/archive/2017/11/inclusive-writing-france-feminism/545048/>.
17. Gabriele Trovato, Alexander Lopez, Renato Paredes, and Francisco Cuellar. 2017. Security and guidance: Two roles for a humanoid robot in an interaction experiment. *RO-MAN 2017 - 26th IEEE International Symposium on Robot and Human Interactive Communication*, 230–235.
18. Gabriele Trovato, Cesar Lucho, and Renato Paredes. 2018. She’s Electric—The Influence of Body Proportions on Perceived Gender of Robots across Cultures. *Robotics* 7, 3: 50.
19. Mark WEst, Rebecca Kraut, and Han Ei Chew. 2019. *I’d blush if I could: Closing Gender Divides in Digital Skills Through Education*. .
20. Meet Q The First Genderless Voice. Retrieved from <https://www.genderlessvoice.com/>.
21. 2019. Mind the 100 Year Gap. *World Economic Forum*. Retrieved from <https://www.weforum.org/reports/gender-gap-2020-report-100-years-pay-equality>.