

Why Project Q Is More Than the World's First Nonbinary Voice for Technology

Insights

- AI is a communication medium between developers and users, not just users-to-users.
- As with other communication mediums, symbolic representation is important to people using the medium. Furthermore, symbolic annihilation of representation negatively affects everyone.
- Humanlike AI interfaces must be crafted in a way that offers fair representation of potential users.

There is a prevalence of gendered technologies emerging in our world. By *gendered technologies*, I specifically mean tech-based products and interfaces designed in a way to purposefully send social cues that it has humanlike gender. Up until recently, this social gender implication was typically presented to users in one of two binary choices, male or female. Siri, Alexa, and Google Home voice-based home assistants all default to feminized voices, although Google Home and Siri can be changed by users to a male voice. As a cultural scholar who works in the emerging technology space, I have been paying attention to this phenomenon because

gender is an enormously powerful tool for socialization. Gender is one way we categorize people, so its presence or absence in social actors—like AI in some mediums and contexts—signifies beliefs and roles we have been acculturated to expect or assign to groups of people. Additionally, voice—although bounded by every person's physical capabilities—is not simply a manifestation of language or verbal communication. Voice also has a learned component and conveys socially constructed cues by speakers and listeners. There is no “normal” or standardized default type of voice. Although societies and cultures may

have expectations of what a normal voice sounds like, those interpretations are built on societal power structures where *normal* is often cultural code for “mainstream acceptance.”

People who design technology understand that a gendered voice is a cue for how people are to interact with a thing, and how to regard it in relationship to themselves. That is why designers, developers, and other decision makers make conscious choices to establish gender cues via a name or voice assigned to consumer goods and services that integrate various levels of persona-driven AI, whether it’s a home assistant, vehicle GPS, telephony voice-guided system, or other medium. These designers and developers of emerging robust technologies that use voice to communicate have the responsibility to apply thoughtful design considerations. That means not reinforcing mainstream ideas of *standard*, *normal*, or similar ways of considering voice that provide limited options for users in a world where we understand voice has strong meaning for identity and that there is a spectrum of gender identities. We also know that voice is a cue for our relationship with an Other—even a technological one.

As with other communication mediums, AI-based technologies that omit representation of groups of people can have negative social ramifications. For example, the omission of groups of people in media representation can be viewed as an attempt to erase their very existence and contributions to the world by not acknowledging and including them as represented social models of being. Moreover, people who see themselves as outside the non-represented group(s) may regard the omitted people as invisible, hidden, unwanted, or of lesser significance around them in society because of that lack of representation in everyday models of social interaction.

If one goal of a voice-assisted product or service design is to create an intuitive, humanlike communication model and gender is designed into the product to align with and leverage our existing expectations for that model, then why force only one or two options onto users when we know people identify across a spectrum of gender? In response, companies will often cite research claiming that people expect and even prefer female or male voices in different situations, such as a male voice in authoritative roles and a female voice in assistive roles. I’ve done studies with humans interacting with robots that have demonstrated similar user expectations and beliefs that align with historically “traditional” American gendered design social cues. However, as seen in my own work in this area [1] and the work of others [2,3], often the same research that descriptively reports those views was not intended to be *applied as a restriction* on product development, feeding into existing cultural biases and stereotypes. Rather, the research findings demonstrate that people are indeed projecting gender-related socially constructed biases onto technology when given even the shallowest lifelike design cues. This type of research is generally intended to be the foundation of understanding how and why people interact with lifelike technologies, not a weapon to ingrain these biases by re-creating them in technological artifacts. I argue that instead of feeding into current rhetoric of design for the low-hanging fruit of user bias reinforcement, we can do better and responsibly design technologies in ways that appease expectations and include a wider range of identities with options for affordances like voice.

In October of 2018, I was approached by Ryan Sherman from global media network Vice’s Copenhagen-based office. He explained

a nascent project idea, Q, a nonbinary voice for use in technology. Project Q was being led by Vice and Virtue, their design arm, and later joined in the collaboration by Copenhagen Pride. They asked for my involvement as a research consultant based on my body of work about behavior, culture, and human-technology interactions. After meeting with people from these groups—including linguist Anna Jörgenson [4], whose work focuses on gender and identity—I was excited to volunteer to help develop a new way for people to interact with voice technology. Jörgenson’s research, writing, and perspective on voice explains well the extraordinary personal challenges that people who identify beyond a gender binary face when they are not represented or poorly represented in voice media.

Emil Asmussen, a Project Q creative lead for Virtue, explains how the project got started: “The initial spark was a talk held at the agency by a representative from The Webbs that discussed hidden bias in AI. The example was a little kid yelling daily at Alexa; how is that going to shape his view on women in the future?” Asmussen and his team were inspired to do a lot of reading and thinking about gender, robots, and AI, which led to a revelation: “The world is increasingly acknowledging [many] gender options; why are there still only two options in AI? AI is born genderless so it seems stagnant that there’s not a genderless option,” says Asmussen. Virtue then began interviewing nonbinary individuals, which made it clear that other gender options in AI would make them feel much more visible and acknowledged.

Thus, created by a group of linguists, social science researchers, technologists, and sound designers, Q began as a proof of concept project whose goals I personally identified as 1) contributing positively to the ongoing global discussion about gender, identity, and technology, 2) demonstrating that developing a human-based voice as a gender-neutral voice option is possible, and 3) showing evidence of the significance of omission, inclusivity, and representation in AI research and design. To develop Q, the workgroup first recorded the voices of six people who identify as male, female, transgender, or nonbinary to authentically blend a voice that did not typically fit within the

People who design technology understand that a gendered voice is a cue for how people are to interact with a thing, and how to regard it in relationship to themselves.

male-female binary. To find this voice, engineers worked on the pitch, tone, and the format filter to blend them to a genderless-sounding single voice. Via Copenhagen Pride, the team surveyed more than 4,600 people who were asked to rate versions of these human voices that had been woven into a single voice. However, at the end of several sound-engineered iterations and accompanying user feedback, survey respondents reported preferring a modulated single voice that fit within a frequency range that was perceived as most gender-neutral. From this experiment, audio engineers were able to technically define and achieve Q.

We all respond to gender cues; in fact, we seek them out in others to give us hints about their social role in relation to our own. My argument for some time—along with some others in the field—is that AI, when integrated into a system for human interaction, can be an emerging social category in some situated uses. Whether something is designed with deliberate gender cues or not, if we project a lifelike narrative onto an intelligent object, one way we interact with it is assigning the gender values we hold and project onto that object’s narrative. This narrative is fundamentally social, something that people construct in order to communicate seamlessly with an object with some sort of humanlike AI output, such as voice and natural language.

Some argue that AI does not need to be gendered or to even imply gender, which is sometimes identified as a kind of deception wherein a technology mimics aspects of humanlikeness and therefore encourages people to regard it as more of a social actor than it is capable of being. However, whether something artificial is designed for social interaction or not, sometimes people treat it that way depending on a combination of factors such as its aesthetics, affordances, movement/embodiment, role in relation to people, task capabilities, and limitations. When all of these factors align in someone’s mind as indicating a thing is lifelike, they may then begin to project a persona onto the object or thing. None of us knows for sure if this human inclination to create social narratives for some smart things will continue

Q'S NOMINATIONS AND AWARDS

**Winner: Bronze Glass Lion
for Innovation (2019)**

**Winner: Radio and Audio Cannes
Bronze Lions (2019)**

- Not-for-profit / Charity / Government
- Sound Design
- Use of Audio Technology/Voice – Activation

Shortlist:

Beazley Designs of the Year award (2019)

on a similar path or evolve differently as we increasingly normalize artificial social presence into our societies. But for now, as AI continues to be a medium for expression and communication and is incorporated into our everyday decision-making processes, it is imperative that sociocultural biases in representation—and omission of representation—are addressed.

As we work toward the goal of Q 1.0 with the responsibility of its further development as a voice option ready to integrate with technology, there are many important nuanced decisions to make. Every day, we continually revisit ideas about how we ourselves are culturally embedding Q through our design and storytelling choices. For example, we have reassessed and then discarded the problematic wording of *genderless* or *gender neutral* for Q and replaced it with *nonbinary* for accuracy and inclusion of the widest variety of people. We intentionally seek external feedback to check our ideas, have a development team with a wide variety of backgrounds and experiences, and employ active self-reflexivity of our own ideas as parts of our ongoing work philosophy. In addition to Project Q, as an organization, Vice is making a concerted internal effort to persistently develop and enforce inclusive policies and actions such as new systems of pay equity, human resources policies, and internal affinity groups.

The hope for our collective workgroup is that when Q is completely digitized, we can make it available for

everyone to access. Furthermore, Q can potentially be used in existing human-voice technology-interaction scenarios that are not necessarily AI-based. As Sherman said to me recently, “Q isn’t just about voice in AI. When you hear a voice at a train station announcing arrivals and departures or safety information on the address system, that could be Q.” In other words, Q 1.0 will be usable in ways that are not necessarily limited to socioeconomically elite groups of people using expensive or emerging technologies.

Virtue’s Asmussen told me via email that the goal for Q is still more or less the same as when the project launched: “We’re working hard to get Q implemented as a third gender option in AI products and believe it will be available before 2020,” adding “The project sparked a global conversation far beyond what we had imagined and only strengthened our belief that this is an important step for AI.”

ENDNOTES

1. Carpenter, J., Davis, J., Erwin-Stewart, N., Lee, T., Bransford, J., and Vye, N. Gender representation in humanoid robots for domestic use. *International Journal of Social Robotics (special issue) 1*, 3 (2009), 261–265; https://www.researchgate.net/publication/220397347_Gender_Representation_and_Humanoid_Robots_Designed_for_Domestic_Use
2. Nass, C.I. and Brave, S. *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. MIT Press, Cambridge, MA, 2005.
3. Davies, S., Papp, V.G., and Antoni, C. Voice and communication change for gender nonconforming individuals: Giving voice to the person inside. *International Journal of Transgenderism 16*, 3 (2015), 117–159.
4. Jörgensen, A. *Speaking of Gender: A Sociolinguistic Exploration of Voice and Transgender Identity*. Master’s Thesis. University of Copenhagen, 2016; https://www.slideshare.net/slideshow/embed_code/key/LSfMJNj3Uxe6aG

Julie Carpenter’s principal research discovers and describes patterns of human behavior with emerging technologies and situates them within larger cultural contexts and social systems to offer a framework for understanding what phenomena are occurring, explain why these interactions are playing out the way they are, and predict future emerging patterns of human behaviors based on these findings.

→ julie@emergingethics.com

jgcarpenter.com