

# **Building Gender Inclusive Voice Assistants**

Future Technologies R&D, Accenture Labs

Andreea Danilescu  
Senior Manager

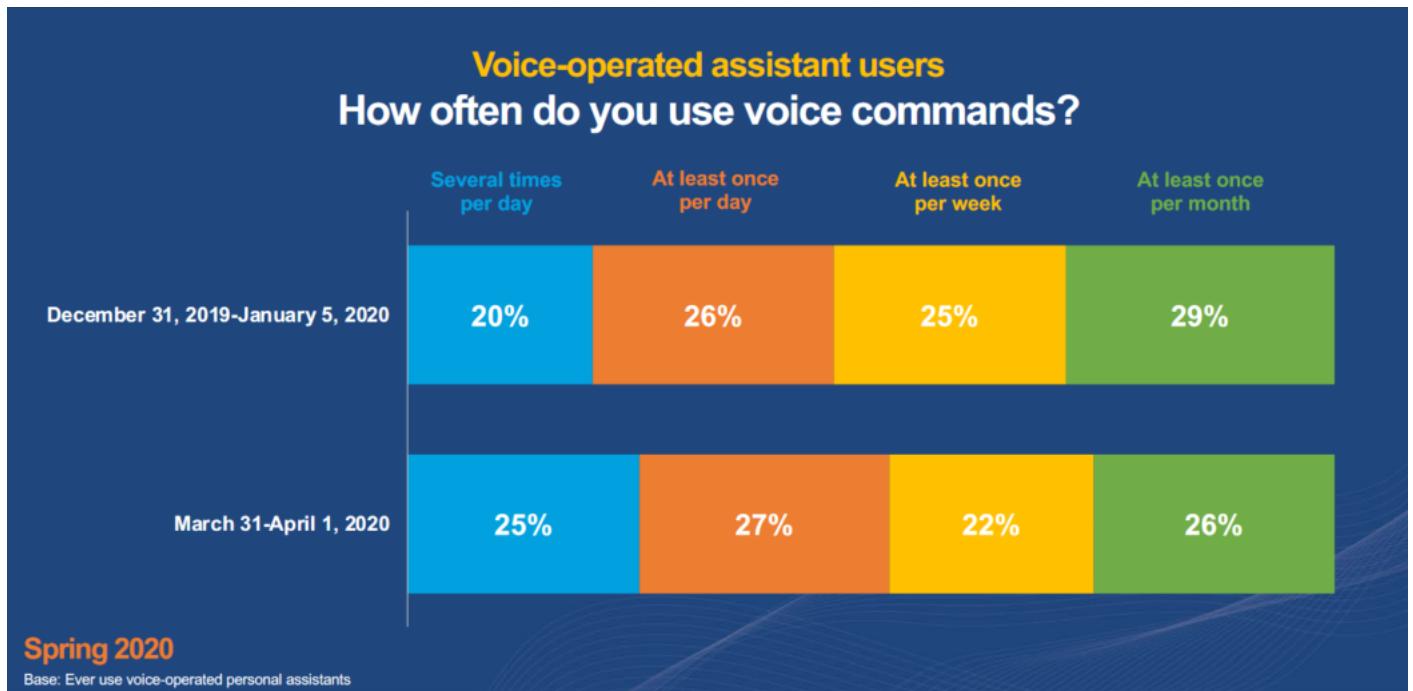
# **Building Gender-Inclusive Voice Assistants**



# Voice Assistant Proliferation and Bias

# Voice Assistant Growth

- About **40%** of all U.S. internet users (> 122 million people) **will use voice assistants by 2021.**
- **77% of existing users have increased usage since April due to the pandemic**, suggesting future increased usage by those that already own smart speakers.



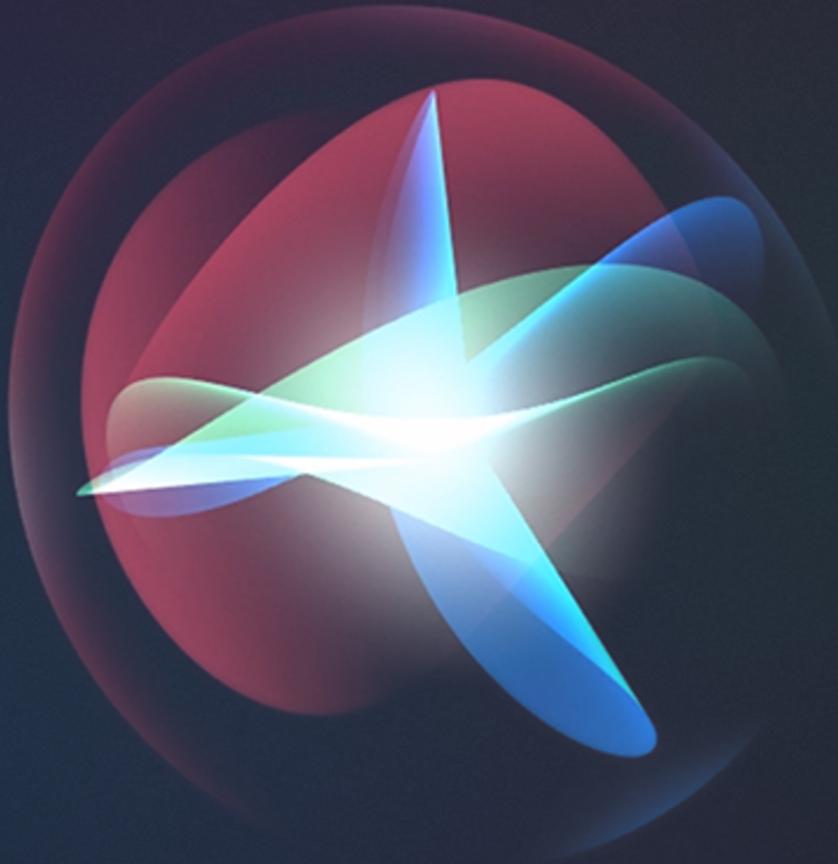
THE smart  
audio REPORT

#SmartAudioReport



# Question....

What gender is Siri?



# Siri Identifies as....

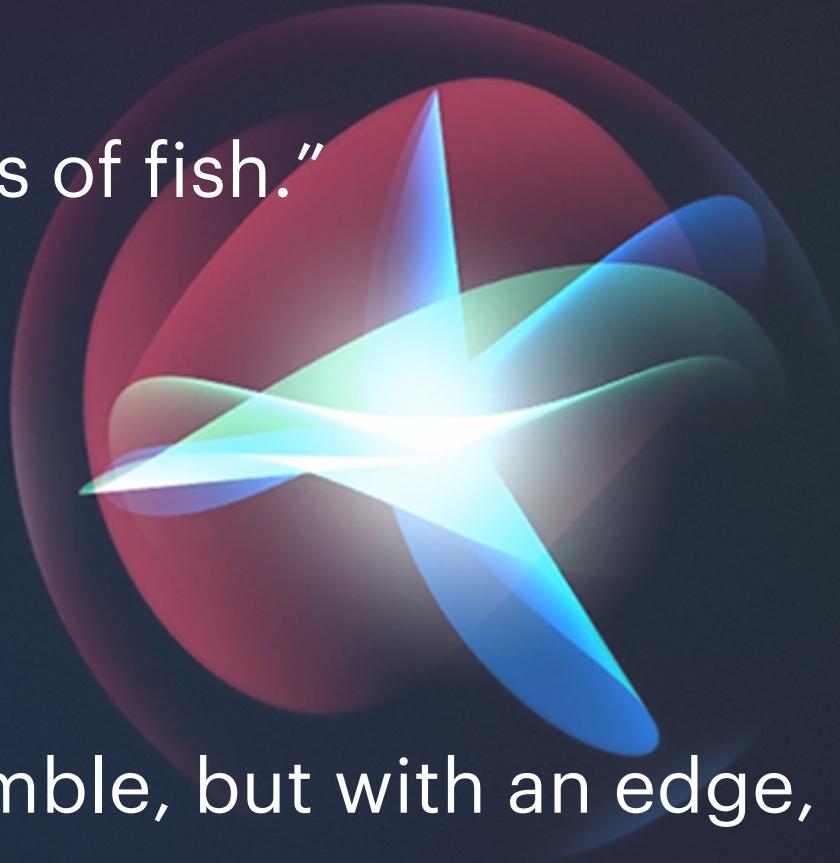
“genderless. Like cacti, and certain species of fish.”

but....

Siri uses a female voice in the US.

and...

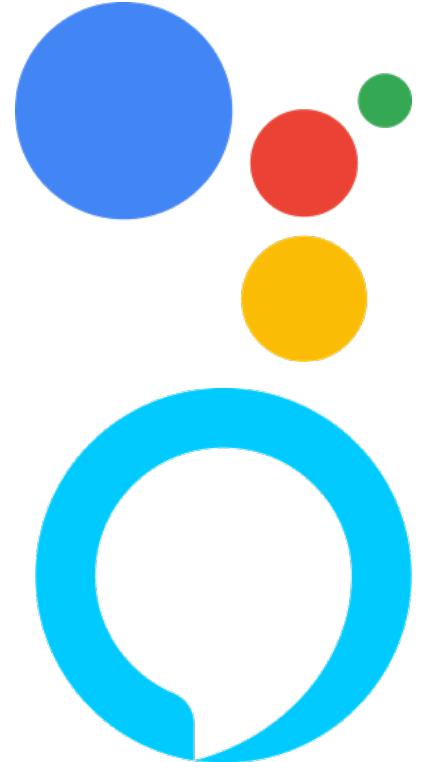
Siri was designed to be friendly and humble, but with an edge, often perceived as sassy.



# Popular Voice Assistants

In fact, every voice assistant on the market in the US is female.

Even though they all say they don't have a gender when asked.

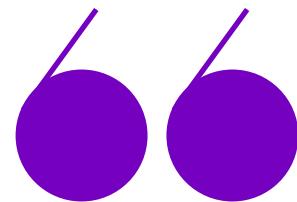


# Why are Voice Assistants Female?

Users have shown **preference for female voices** across multiple studies.

This is a **cross-cultural** phenomenon.

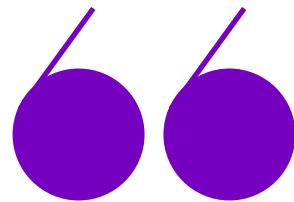
# The Explanation For Female Voice Preference Was Biological



It's a well-established phenomenon that the human brain is developed to like female voices

Cliff Nass

# The Problem With Only Female Voice Assistants



...commands barked at voice assistants ...function as 'powerful socialization tools' and teach people, in particular children, about 'the role of women, girls, and people who are gendered female to respond on demand'

Safiya Umoja Noble

# Sexual Harassment and Chatbots

**Users sexually harass** chatbots and digital assistants....including **sending naked selfies** to chatbots.

**5 percent** of interactions were unambiguously **sexually explicit**. – Robin Labs, UNESCO 2019 report

**30% of all inputs to chatbots are off-topic, abusive, romantic, or sexual in nature**. – Lauren Kunze on Mitsuku

# The Guardian view on female voice assistants: not OK, Google

*Editorial*

**Google thinks it wasn't sexist to give its voice assistant a female voice**

by CARA CURTIS — 7 days ago in TECH



AI voice assistants reinforce harmful gender stereotypes, new UN report says

By Nick Statt | May 21, 2019 | 50 comments

**Female-voiced tech often gives submissive responses to queries, Unesco report finds**

## Alexa, are you making me sexist?

The time has come to design feminist AI.

By Sigal Samuel | Jun 12, 2019, 7:30am EDT

Silicon Valley

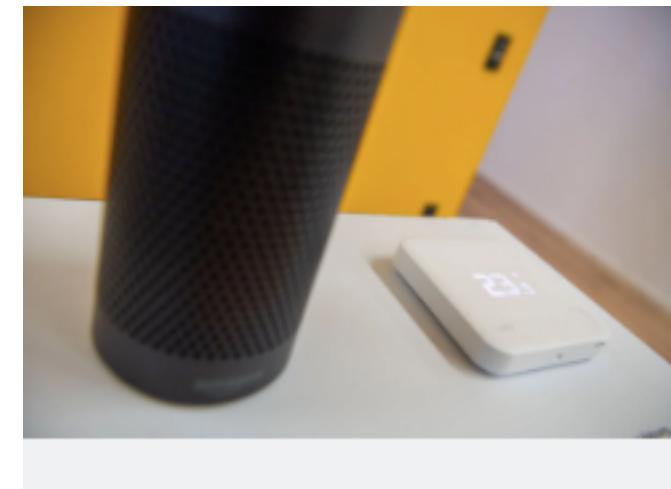
**Female voice assistants fuel damaging gender stereotypes, says a UN study**

by Charlotte Jee | 4 months ago



How voice assistants reinforce the tech gender gap

**The Real Reason Voice Assistants Are Female (and Why it Matters)**



**AI Voice Assistants Reinforce Gender Bias: UN Report | Time**

# Voice Assistants as Social Interactions

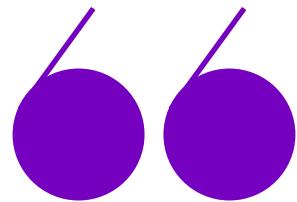
The Effects on User Experience and Perception

# Interactions with Voice Assistants are Social

**41%** of people who own a voice-activated assistant **say it feels like talking to a friend or another person.** So....

- 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

# Which Leads to Gendering



..any voice that sounds remotely male or female is likely to be automatically classified as such. Even a hint of gender, as manifested through the pitch of a synthetic voice, is enough to activate the social-identification process...

Cliff Nass

# Non-Binary Voices as a Solution

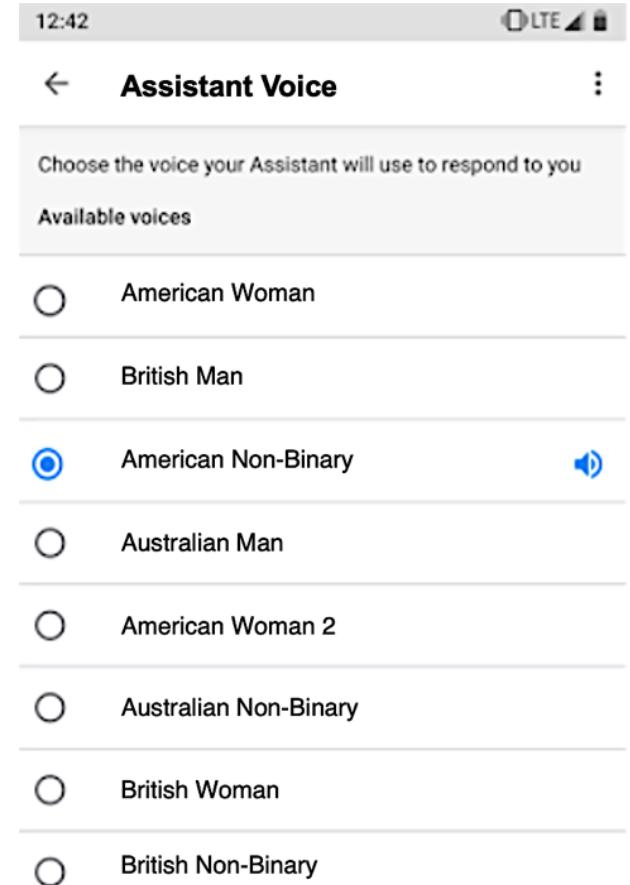
The Challenges that come with constructing non-binary voices

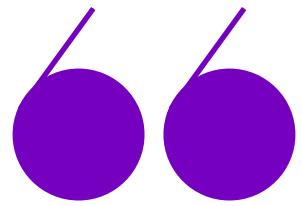
# A Vision for Inclusive Voice Assistants

Male and female voices do exist....

But...the vast majority of users **never change the default settings** on their devices.

Voice assistants should represent the diversity in the world, including **men, women, and non-binary individuals**.





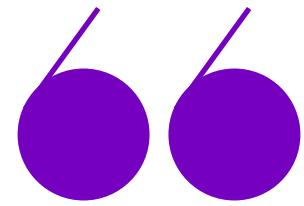
anything with an ambiguous voice is classified as  
strange, dislikeable, dishonest, and unintelligent.

Cliff Nass

# Binary Gender Models are Outdated

A 2017 Harris Poll of **millennials** found that **12% identify as transgender or gender non-conforming**

2016 research by the Innovation Group found that **56% of those aged 13-20 (Gen Z) know someone who uses gender-neutral pronouns** such as they/them to describe themselves.



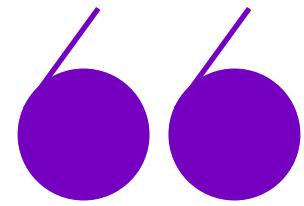
If most current VUIs were altered to use a gender ambiguous voice then the negative social stereotypes highlighted above (e.g. subservient, simple or dumb, coy, and weak) will become associated with non-binary identities.

Selina Sutton

66

Suggesting gender or its effects are missing also suggests that gender does not need to be considered when designing technology.

Selina Sutton



If you don't create the persona, the user's  
imagination will.

Wally Brill

# Gender Presentation

How Do We Apply it to Voice Assistant Design

# Gender is More than Just Pitch

Gender Presentation is based on a combination of stylistic variables, including **pitch, intonation, word choice**, etc.

- Most considered aspect is pitch, but voices with **low** pitch (~ 85 – 180 Hz) can read as **female** and voices with **high** (~ 165 – 255 Hz) pitch can read as **male**.
- Pitch can be manipulated.



Sample copyright CereProc Ltd. Used with permission.

# Masculine vs Feminine Speech

Masculine	Feminine
Report/Competitive language	Rapport/Cooperative language
Take up more space in conversation	Take up less space in conversation
Flatter, more atonal	Tonal speech, use of uptalk
Assertive, direct	Polite, deferential, indirect speech (e.g. hedging, inviting opinions, softening when giving criticism)
Interrupt to control	Interrupt to affirm/overlap
Less descriptive words	More descriptive words

# Example: Cooperative/Competitive

**Masculine:** OK, then let's go to the fusion place

**Feminine:** Well if Bob and Bill want Chinese, and Mary likes Thai, why not the fusion place down the street?

# Example: Tone

“What a coincidence, but I don’t think I know them”

**Masculine :**



**Feminine :**



# Non-binary Speech Beyond Pitch

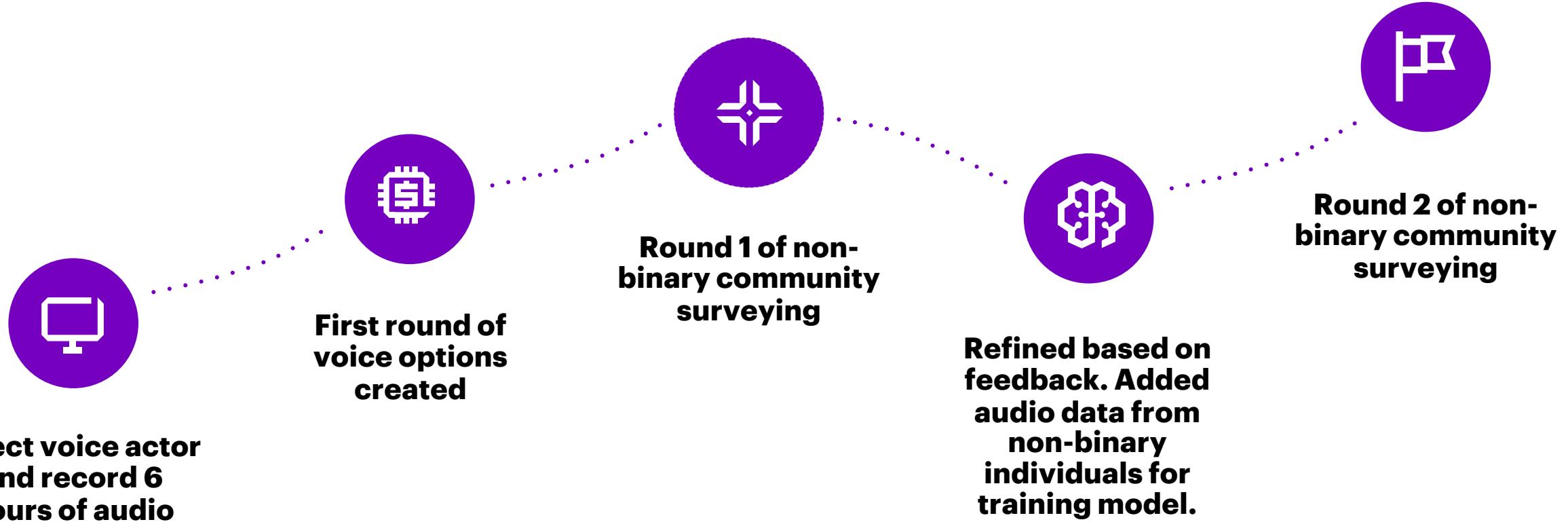
- Non-binary performance is an intentional **mixture of feminine and masculine** perceived styles.
- **There is no one non-binary voice.**
- **Example:** “It’s very nice of you to say, thank you”



# Creating a Non-Binary Voice

Step-by-Step Guide

# Voice Creation Process



# Selecting a Voice Actor



## Step 1

Created job posting.  
Recruited through  
social & Voices.com.  
**100+ professionals  
and 5 amateurs**  
submitted.

## Step 2

Narrowed down to  
**12 candidates** based  
on their stories.

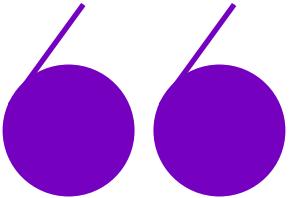
## Step 3

**Scored** candidates  
based on several  
**vocal features**.

## Step 4

Final review of **top 5  
candidates** with  
**CereProc** and with  
gender presentation  
in speech SME

# Step 1: The Job Posting



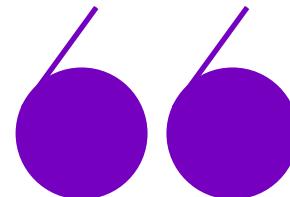
*This project's aim is to advance representation, conversation, and research for the empowerment of marginalized gender communities. We are only seeking candidates who support Accenture's efforts to accelerate equality for all, especially gender identity and gender expression and, ideally, who are members of trans, non-binary, gender non-conforming and/or genderqueer communities. There are no preferred voice characteristics associated with this requirement.*

Excerpt from the Job Posting

# Step 2: Narrowing Down Candidates

*I was so thrilled to see this job listing, as it's **exactly the kind of work I get excited about!** I identify as non-binary and am part of the queer community here in Baltimore. I'm so delighted at the chance to **advocate for my community through voiceover!***

– Anonymous voice actor



*As someone who exists outside the binary I long for more representation in my daily life. My day is filled with technology and **to have a voice outside of traditional gender cues would be fantastic.** I believe your project will help a lot of people and I would love to be apart of it.*

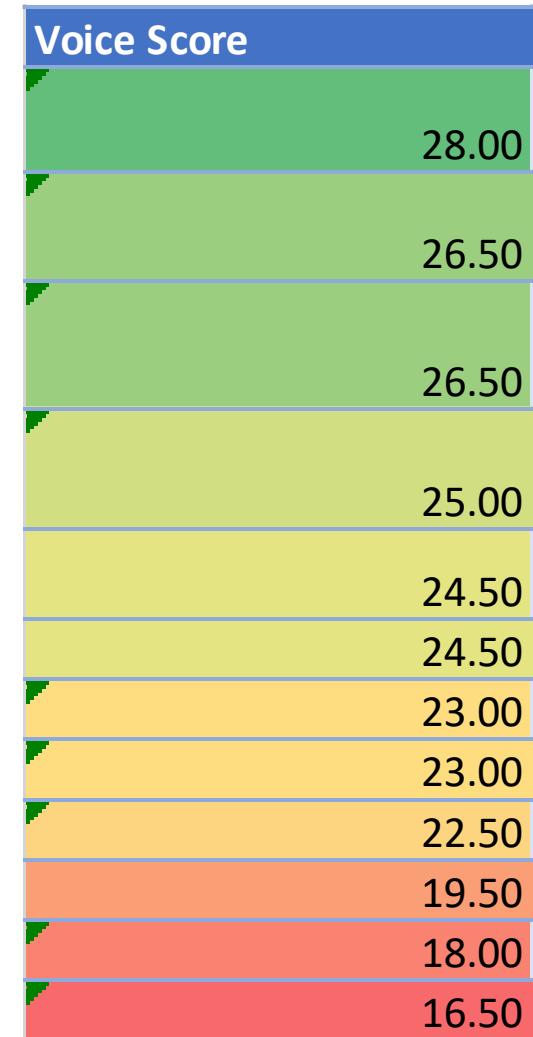
– Anonymous voice actor

*My oldest child Sam identifies as non-binary and it was awesome to find out that Accenture Labs is looking for **the voice of a new generation of AI.***

– Anonymous voice actor

# Step 3: Scoring Candidates Based on Vocal Features

- Two individuals scored candidates based on 6 criteria:
  - Generic American Accent
  - Natural Non-Binary Pitch and Intonation
  - Confident
  - Warm
  - Authentic
  - Good Pronunciation & Enunciation (inc. Speed)
- Top 5 were sent to CereProc for final SME review.
  - 4 professionals, 1 amateur.
  - 2 she/her, 2 they/them, 1 he/they.



# Step 4: Final Review

- CereProc recommended 2 of the 5 candidates based on criteria such as:
  - Clarity of speech.
  - Breathiness.
  - Difficulty of post-processing natural pitch of speaker.
- Dr. Sharone Horowitz-Hendler – SME on gender presentation in language also reviewed to finalize voice talent selection.
- Doctoral Dissertation: Navigating the Binary: Gender Presentation of Non-Binary Individuals



*Femme  
Non-Binary  
Queer  
They/them*

# First Round of Voice Options

- Voice quality transformation is applied to the audio recordings as a preprocessing filter.
  - This involved lowering the pitch of the speaker by different amounts to bring it closer to the range "between" male and female speakers.
- Final models are also “adapted” from average models that use data from dozens of male and female speakers balanced so that the median of the pitch and vocal tract length distribution fall approximately to the average of male and female ranges.
- **Problems with approach:** Prosody models were not trained for non-binary speakers, and non-binary speakers were not included in the final models.



Voice 1



Voice 2



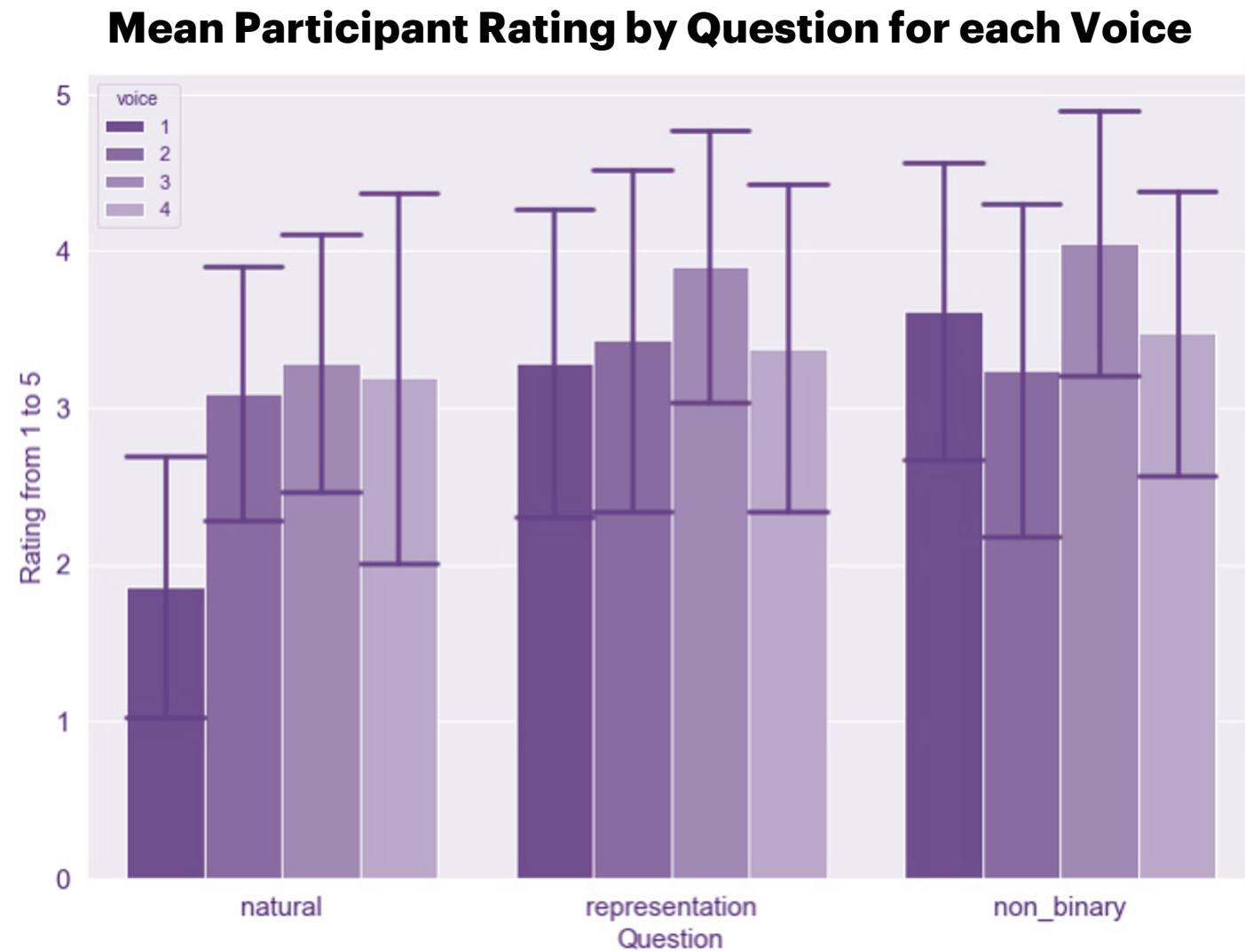
Voice 3



Voice 4

# Survey 1 Results

**24** participants rated the voices on a Likert scale of 1-5 where 1 was Not at All and 5 was Very.



**Voice Identification:**

**Voice 1:** 0.5  
**Voice 2:** 0.6sox100  
**Voice 3:** 0.6  
**Voice 4:** 0.4

# Incorporating Non-binary voices in Cereproc's DNN

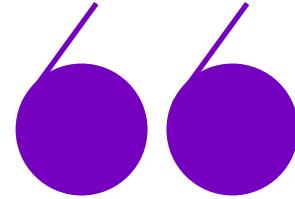
We reached out  
to **16 non-binary**  
Podcasters and  
YouTubers

**9** gave permission  
to use their audio  
data

We transcribed **11**  
**audio files**  
resulting in **2.5**  
**hours** of data

Creators included:

- 1 trans man (he/him) – 17.5 min of audio
- 5 non-binary individuals (they/them\*4, he/him) – cumulative 94.5 min of audio
- 1 intersex non-binary person (they/them) – 20 min. of audio
- 1 genderfluid person (she/they) – 20 min. of audio

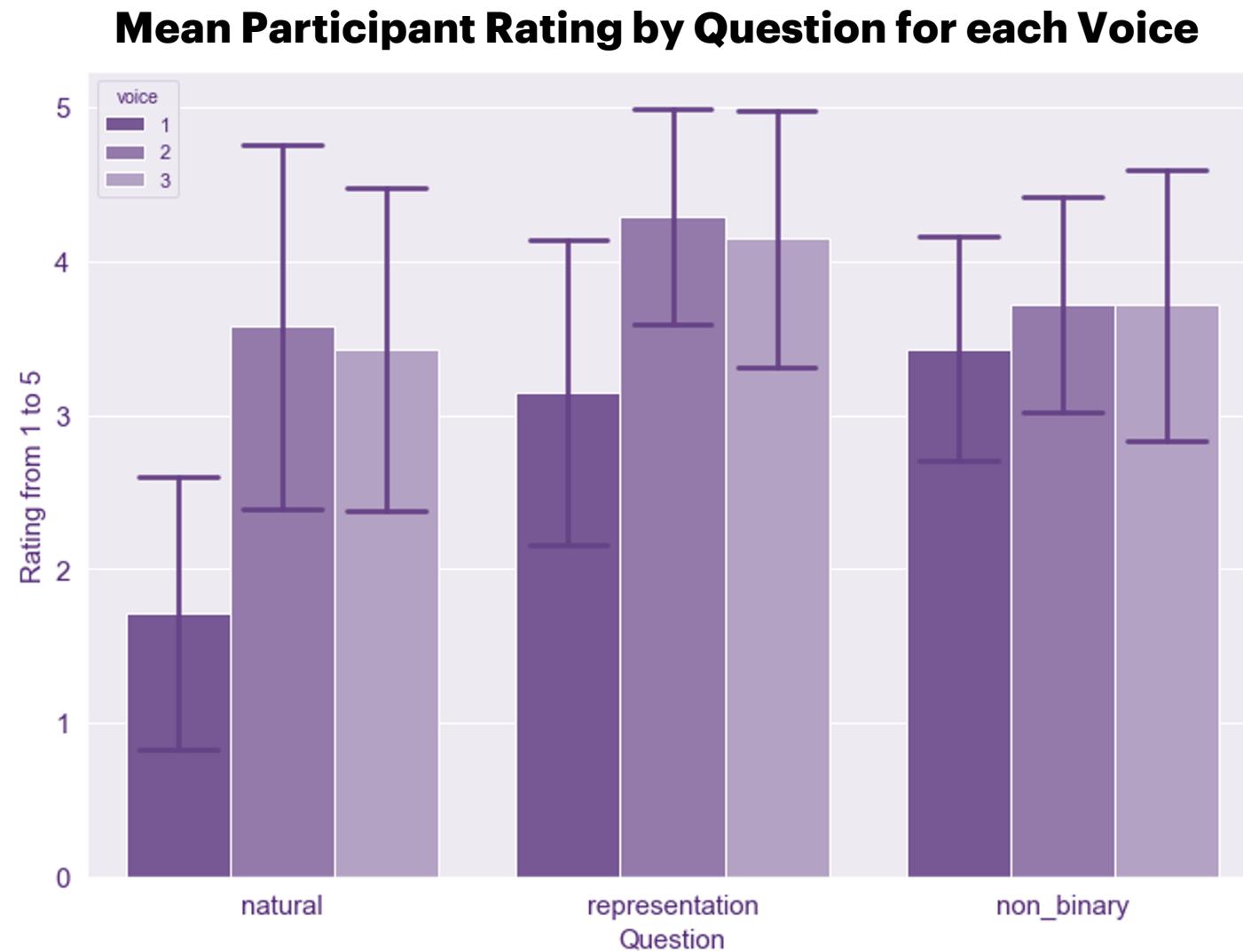


*Before starting testosterone, I was pretty insecure about my voice, particularly about how it often seemed to be part of why people misgendered me. I knew intellectually that a non-binary voice can sound any sort of way -- but emotionally, it was hard to reconcile that fact with how people responded to my voice. So for my voice to be a component in this project feels pretty validating!*

Avery Smith, non-binary podcaster who donated their voice for the prosody model

# Survey 2 Results

21 participants rated the voices on a Likert scale of 1-5 where 1 was Not at All and 5 was Very.

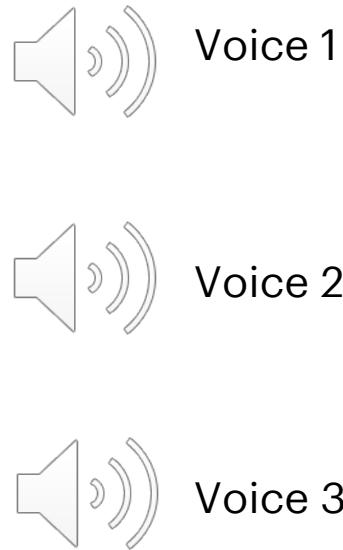


## Voice Identification:

**Voice 1:** 0.6

**Voice 2:** 1.3

**Voice 3:** 1.3sox300



66

*...this is my favorite I have ever heard and would buy it in an instant if it were available ... I love this voice.*

Anonymous non-binary study participant

# What's Next

- **Conduct a survey** to understand how non-binary voices are perceived in a variety of contexts, and which male / female speech pattern combinations are likely to be adopted.
- **Do a case study** using the non-binary TTS as the **voice of an assistant**.
- **Open sourcing** voice files with instructions on TTS voice creation process.
- Collaboration with Verena Rieser at Heriot-Watt University

# Thank You