

Friendly female

01/



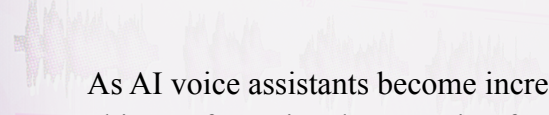
04/



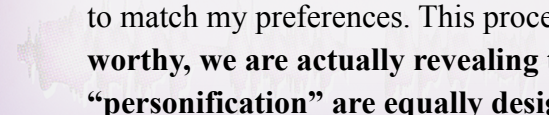
07/



10/



14/



# NEW VOICE TONE

As AI voice assistants become increasingly common in daily life, they are gradually evolving from functional tools into objects of emotional connection for users. In my daily life, I also chat with voice assistants and select different voice styles to match my preferences. This process led me to reflect: **when we “customize” a voice that feels comfortable and trustworthy, we are actually revealing that in the “seemingly neutral” internet, the emotions behind the voice and its “personification” are equally designed—not truly equal or naturally generated.**

# Design Concept

---

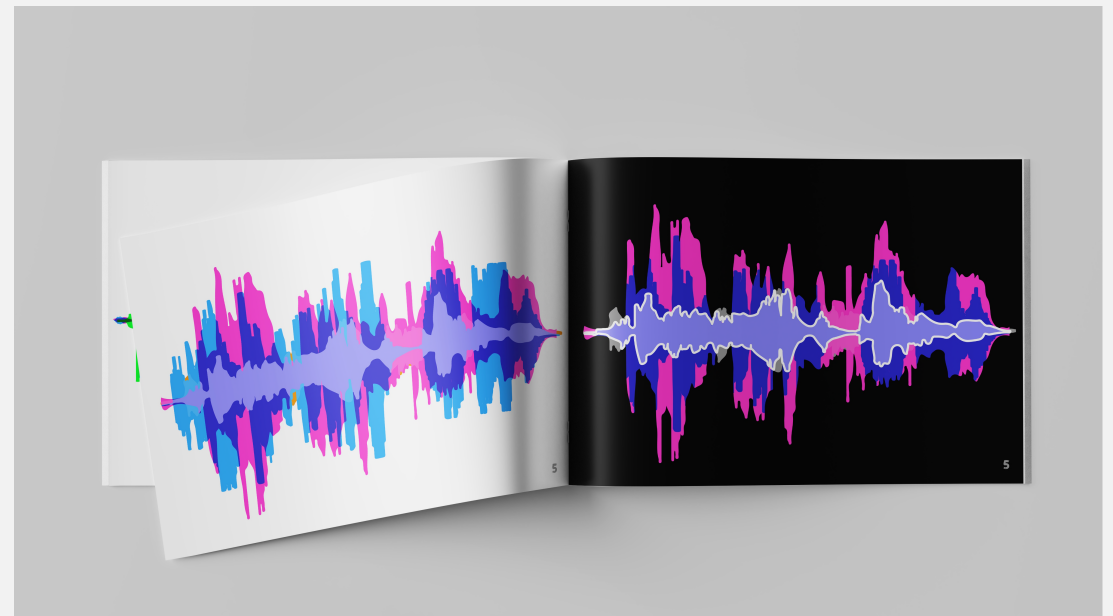
## Poster

**This project visualizes the audio waveforms of four distinct AI voice tones in real-world contexts.** By observing the amplitude fluctuations, density, rhythm, and coherence of sound across the timeline, each tone reveals its unique “communication posture”: Some voices flow smoothly, lighthearted and approachable Others remain steady, punctuated, and formally restrained Transforming sound into imagery serves not merely technical analysis, but to make viewers intuitively recognize: vocal intonation itself is a designed language. This visualization renders visible the emotional nuances, interaction patterns, and implicit biases hidden within auditory perception, prompting us to rethink how anthropomorphized AI voices influence user experience and emotional connection.



## Collage Experiment

**After analyzing the interview results, this project employed a “collage-style reconstruction” approach to combine preferred characteristics from the four voice styles, constructing an “idealized AI voice.”** By visualizing these preferences and synthesizing the resulting sounds, the project aims to highlight that what we perceive as “warm or formal” voices are actually engineered products. This redesign also serves as a reflection—it reminds us that anthropomorphism does not equate to endowing AI with a personality. Rather, voice and intonation provide clues that allow users to “perceive personality.” Different rhythms, inflections, and modes of expression naturally evoke specific emotions, attitudes, and roles, thereby forming an image of the AI as “someone like a person” in the user's mind.



# Observation of Sound Wave Characteristics

---

## 1) Friendly Female

Waveform fluctuations:

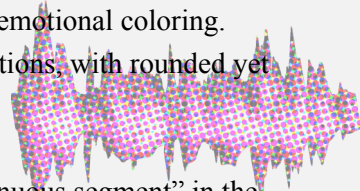
Significant vertical waves indicate pronounced emotional expression in speech.

A clear “high-low-high” pattern reveals elastic intonation with emotional coloring.

Localized spikes indicate rising intonation or increased interjections, with rounded yet slightly pointed peak shapes.

Horizontal Density:

Overall continuity with minimal pauses, forming a “long, continuous segment” in the waveform.



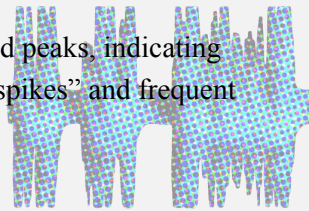
## 2) Friendly Male

Waveform Characteristics:

Vertical amplitude exhibits significant fluctuations with pronounced peaks, indicating distinct emotional expression. The waveform contains numerous “spikes” and frequent minute vibrations, appearing lively and highly elastic.

Horizontal Density:

High density with strong continuity, showing almost no noticeable pauses.



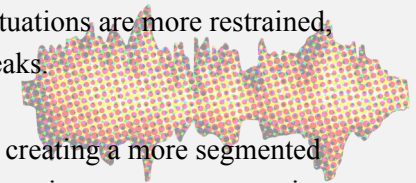
## 3) Formal Female

Waveform Characteristics:

Vertical amplitude remains relatively consistent overall, with minimal pitch variation. Emotional fluctuations are more restrained, featuring shorter, denser, and sharper peaks.

Horizontal Density:

Subtle pauses occur between sentences, creating a more segmented structure. This produces a rhythm of “narration → pause → continuation.”



## 4) Formal Male

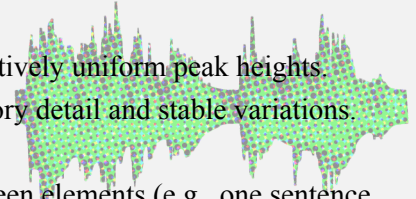
Waveform Profile:

Moderate vertical fluctuations with relatively uniform peak heights.

Narrow waveform with minimal vibratory detail and stable variations.

Horizontal Density:

Low density with frequent pauses between elements (e.g., one sentence, one pause).



# Interview Transcript Analysis (Full transcript available in specific file)

Participants generally agreed that all four voices sounded overly engineered, lacking the natural quality of human speech and instead sounding like someone deliberately playing a role. However, the gentle female voice was the most well-received compared to the other three. Therefore, I will rebuild and modify the voice based on the gentle female voice's acoustic characteristics.

# Collage Experiment Design Instructions

---

In the collage reconstruction, I used the gentle female voice as the foundation, as it received the most positive feedback in interview statistics. However, participants noted its excessive upward inflection and overly “performative” emotional expression. **Therefore, I referenced the intonation range of serious female and serious male voices, reducing the overall pitch variation to achieve a more steady tone. Simultaneously, I extracted natural breathing rhythms from the warm male voice—specifically, introducing uniform micro-pauses within the waveform rather than maintaining a completely “unbroken flow.”** This approach made speech cadence and pauses feel more organic. The resulting voice profile retains warmth while avoiding excessive emotionality, presenting a closer approximation to the “ideal AI voice style.”

# Design Process (Outputs are detailed in specific documents)

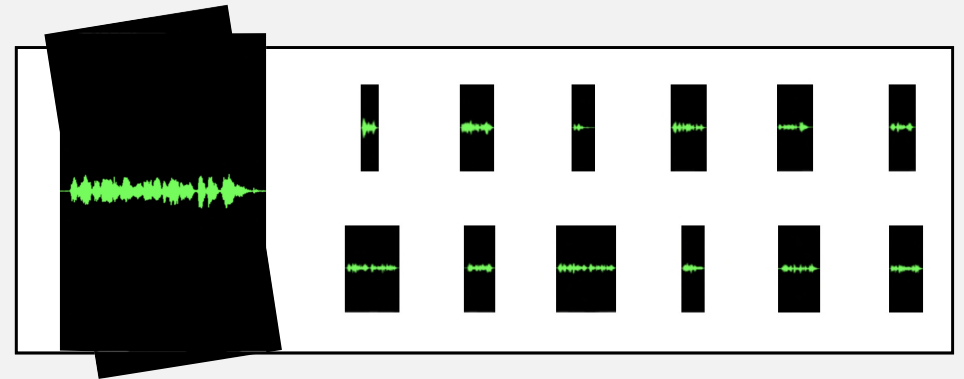
## Poster



Conversation records



Sound Recordings

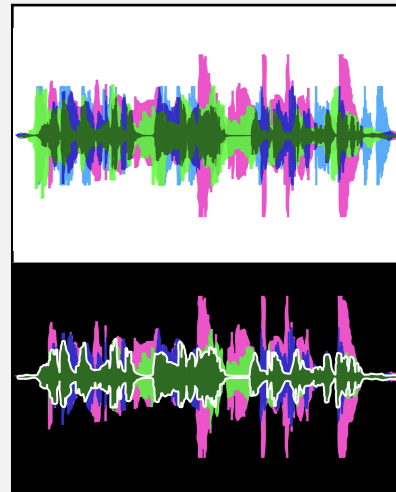


Extract sound waves

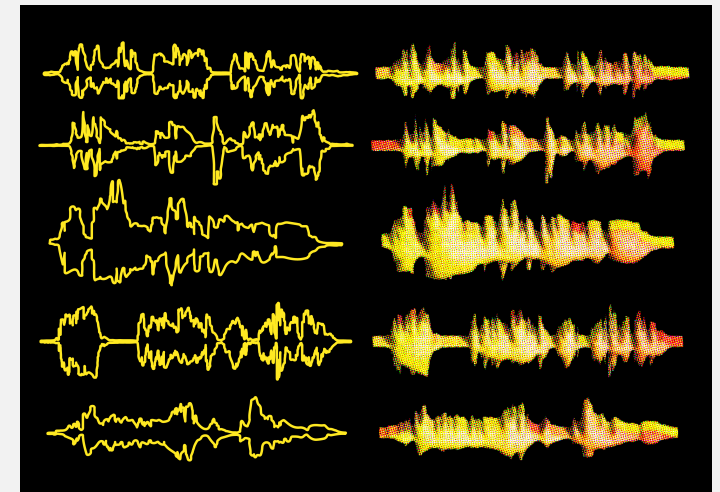
## Collage Experiment



Extract the advantages



Collage and reconstruct



Result