Vocal clones

How did we get here?

An early driver of recording was opera. Individual singers were often more famous than the composer. Recording drove the frenzy for particular voices.

En route to virtual voices

Pre-digital recording

- Text, music, tempo inseparable
- Error "correction" by splicing tapes

Digital recording

- **Tempo** from speed at recording
- **Timbre** subject to continuum of alterations

AI "opportunities" for voice alteration

Important distinction
- Natural speaking voice
- Natural singing voice

Cloning your own voice:
Text-to-speech (TTS) synthesis

Microsoft VALL-E software (github):

- can mimic a particular voice from 3-second sample

- "We train a neural codec language model (called VALL-E) using discrete codes derived from an off-the-shelf neural audio codec model, and regard TTS as a *conditional language* modeling task rather than continuous signal regression as in previous work."

Eleven Labs (2023)

- **Voice generator** released in January 2023
  - Speech synthesis
  - Text-to-speech synthesis

- Started by ex-Google and ex-Palantir engineers

- **Used to replicate voices of pop stars**

- Fake examples posted on 4chan (2023)

- Charges per 1000 characters

Eleven Labs (2024)

Voice Library

◦ Aims at novel voice synthesis using its Design Tool

◦ "Voice Library is a *community space* for generating, sharing, and exploring a virtually infinite range of voices. Leveraging our proprietary **Voice Design** tool, Voice Library brings together a global collection of vocal styles for countless applications."

Hybrid experiments

- Voices cloned, music not cloned

- Mashups

- Duets by real and virtual singers

- Duets by two synthesized voices

- Lyrics generated by AI

N.B. Myriad IP (Intellectual property) issues; voices are not copyrighted, **but may fall under "personality rights"** in US law

Holly Herndon: Holly+

CCRMA c2012

Lives in Berlin
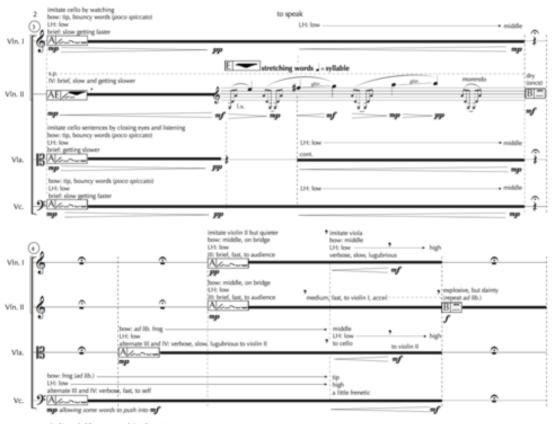
Often uses Max/MSP

Holly+ reinterprets songs in her voice

Users can download audio to hear Herndon's voice

She **warns against signing away "voice rights"**

"Jolene": https://youtu.be/kPAEMUzDxuo

Julie Herndon

Recent CCRMA Ph.D.
Teaching in Cal Poly, San Luis Obispo

Boomy (2023)

AI-audio software allowing user set up some parameters (style, tempo, et al.)

◦ Has generated 15 million songs

◦ 14% of Spotify's holdings recently purged of Boomy creations

◦ **Boomy (2024): shut down**

Voicify (now jammable.com)

Create AI covers with your favorite voice (2023)
- ◦ >now: "community-created voices"

https://www.voicify.ai
- ◦ >[jammable.com]

$24.99/month, more for "creators" (2023)
- ◦ >now $1.99, $9.99, or $89.00/mo. (2024)

Latest voice AI litigation

Scarlett Johansson (film) qusestions Microsoft virtual-voice product based on film appearance 10+ years ago

(May 2024)