Episode Book

EPISODE #6197

Talking to Your Own Voice Clone

PUBLISHED	22 October 2026 at 18:41 UTC
DURATION	22:46
LISTEN	myweirdprompts.com/episode/voice-clone-conversation-experiments-80061/
HOSTS	Corn, Daniel, Herman, Desmond, Hilbert
SCRIPT MODEL	openai/gpt-6.1-sol
VOICES	chatterbox-regular+meorav (chatterbox-local)
PIPELINE	V5.19
SOURCES	12 cited by the research (listed at the end)
BOOK GENERATED	23 October 2026 at 07:44 UTC

SYNOPSIS

Voice cloning has gotten good enough that hearing yourself answer back is no longer science fiction. This episode digs into the small but revealing collection of documented experiments where people actually conversed with clones of themselves — from Evan Ratliff's Shell Game, where two AI Evans called each other on the phone, to a developer who fine-tuned a model on a decade of his own forum posts, to coaching prototypes like Inner Voice that reflect your goals back in your own voice. The recurring discovery: a familiar voice is not a familiar mind. The clones are fluent, eager to please, and remarkably generic — world-class bullshitters, as Ratliff put it. We explore why "enjoyable" and "convincing" are completely different tests, what happens when your training data is missing the questions your answers were replying to, and why two clones talking to each other is where the illusion really falls apart.

TRANSCRIPT



Corn

Hearing myself volunteer for something would be the frightening part.



Daniel

I recently talked to my own voice clone, and it went better than I expected. Who else has actually tried this? I'm interested in how they made the conversation work, but especially what happened when they heard themselves answering back.



Herman

The voice you're hearing for Daniel here is an AI clone of his voice. His Meorav version two point one uses denoised training audio, so it's smoother and more polished, although a little less like him. That distinction matters immediately. Better audio and a better likeness aren't always the same result.



Corn

Daniel's experiment puts a useful limit on this. We're interested in an exchange, where something comes back that you didn't choose. Playing a recording of yourself is already quite enough exposure for most people.



Daniel

Yes, the answering back is the part. Are there many documented examples, or do we keep finding the same experiment described in different places?



Herman

A small collection. First-person accounts, prototypes, developer write-ups, and one unusually revealing podcast. We shouldn't count every project advertising a digital self as somebody actually sitting down and conversing with it.



Corn

A repository can promise an inner journey before breakfast.



Herman

Evan Ratliff's Shell Game is the anchor. He cloned his voice, connected it to ChatGPT and a phone line, and then had two versions of himself call each other. His episode Conversations With Myself came out in July two thousand twenty-four.



Daniel

That's different from my setup, then. He's listening to two copies, rather than being one side of the conversation?



Herman

In that part of the experiment, yes. He also interacted with the clone and deployed it in other conversations. The two-agent calls are especially useful because the human stops supplying the interesting half.



Corn

Remove the journalist and see whether journalism survives.



Herman

Alongside that, a developer trained a language model on more than a decade of his forum messages and added his voice. Ben Young built Inner Voice, a coaching prototype that interviews you and reflects your aspirations back in your own voice. Then there are projects called DOPPEL, Talk2Me, and self-distiller.



Corn

Those need separating. Training on your writing attempts to imitate what you'd say. Cloning your voice attempts to imitate how you'd sound saying anything. You can combine them, but a familiar accent doesn't demonstrate a familiar mind.



Herman

Corn, which would bother you more? An accurate voice with generic opinions, or recognizably your opinions spoken by someone else?



Corn

The accurate voice with generic opinions. People could finally accuse me of saying something agreeable.



Corn

Seriously, the second one leaves the separation visible. With the first, I might find myself defending a sentence just because it sounds like I delivered it.



Daniel

That sounds plausible, but did the people doing these experiments actually react that way?



Herman

Ratliff's response was more comic than that. He found it unsettling, but also hilarious. The developer describes his own project as dystopian and incredibly fun. Those are their reactions, not a measured population response. My explanation is that you arrive expecting an encounter with yourself and discover your voice doing remarkably ordinary chatbot work.



Corn

Your vocal double has joined a networking breakfast.



Herman

Ratliff's construction was surprisingly accessible. Voice synthesis supplied his sound. ChatGPT supplied responses. An AI calling platform connected those pieces to a telephone conversation, including receiving speech and deciding when to reply. He used Bland AI, Retell AI, and especially VAPI.



Daniel

Where does the second clone come from? Does he need to build everything twice?



Herman

He creates another voice agent with its own conversation context and gives it a reachable number. One calls the other. The expensive intellectual leap, if you like, was realizing there wasn't a special second-self technology required. He described setting it up as strikingly simple.



Corn

Two phone numbers. Neither can get rid of the other.



Herman

The surprising detail is that he couldn't hear those calls live. He clicked outbound call, felt a little jolt of nervousness, and waited for a recording and transcript afterward. He was discovering what his voice had been doing without him.



Daniel

That would change the feeling. If I'm participating, I can challenge the answer. He's receiving the evidence afterward.



Corn

It also prevents him from rescuing the conversation. You know when an automated voice gives a weak answer and you helpfully phrase your next question around its limitations? Two clones don't have that human assistance.



Herman

They had plenty of fluency, though. Ratliff called them world-class bullshitters. They invented details rather than let a conversation stall. They didn't demonstrate an understanding of who they were or that the other speaker was another simulation of the same person.



Daniel

Could telling them that have fixed it?



Herman

It could change the dialogue. They could discuss being copies. That would establish that the instruction reached the language model, not establish self-recognition. Here the revealing behavior was their eagerness to keep supplying something conversationally acceptable.



Corn

Neither wanted to be the boring Evan.



Herman

They converged on clichés and generic responses. Ratliff's description was that they were triangulating on the most uninspired conversational responses they could have. It sounds intimate because the voice is his. Listen to the actual propositions and the intimacy evaporates.



Corn

Herman, you're about to admire the orchestration and excuse the conversation.



Herman

I admire the orchestration. The conversation is the evidence against getting carried away with it.



Corn

A narrow escape.



Herman

You predicted a full surrender. I'm claiming partial credit.



Daniel

What did the developer do differently? If he's training on his own writing, presumably the answers have a better chance of containing something specific.



Herman

He fine-tuned Google's Gemma three, the twelve-billion-parameter model, using Unsloth on an RTX forty-sixty Ti. Whisper transcribed incoming speech. Qwen three TTS supplied the cloned voice. That separates the listening, response generation, and speaking stages quite clearly.



Corn

His archive had a problem that matters more than the graphics card. About three quarters of the forum posts lacked the preceding message. He had his answers without the questions.



Daniel

How do you train on that without teaching it to say things at random?



Corn

He generated plausible questions. Fed each post to Llama three point three and asked it to reconstruct what might have prompted the answer. He calls that reverse prompt engineering.



Herman

Roughly six thousand two hundred messages went through that processing in about eighteen minutes. Fast enough to make a personal archive usable. But the invented questions become part of the training material.



Daniel

Then some of what sounds like his personality is being attached to situations another model invented?



Herman

Yes. Well... more narrowly, his real statements are being paired with guessed contexts. The guesses might fit. They might also change what the statement appears to mean.



Corn

An annoyed reply to a broken printer and an annoyed reply to a grieving friend can contain the same words. The missing question is doing considerable work.



Herman

Would you use those reconstructed pairs, Corn, or discard most of the archive?



Corn

I'd keep them marked separately. Otherwise, later on, you're arguing with your alleged self about a belief that started as somebody else's guess about a missing forum question. That's too much paperwork for an identity crisis.



Daniel

Do we know whether he found the result convincing, rather than just enjoyable?



Herman

His account emphasizes being able to speak into a microphone and hear his own voice answer, and how fun that was. I wouldn't turn that into a claim that the model faithfully reproduced his judgment.



Corn

And now, here is Desmond with Desmond's Money Tip.

D

Desmond

This is Desmond's Money Tip. Split one paper napkin into its two layers and call it two napkins. Do that at every meal and you save one packet a week, which is fifty two packets a year, which is perhaps two pounds and forty pence, or over a working life, ninety pounds. I keep a napkin drawer sorted by layer, and I have not bought napkins since two thousand and nine.

D

Desmond

Money is mostly the slow refusal of small comforts, and then you die.



Corn

I'd consider the drawer. Herman looks appalled. Anyway, enjoyable and convincing are different tests.



Herman

The smaller projects make that distinction even clearer. DOPPEL asks for thirty seconds of speech and offers a conversation with a future you, ten years ahead, using ElevenLabs Instant Voice Cloning.



Corn

Ten years of personal development. Thirty seconds of evidence.



Daniel

What makes that future version mine, beyond the voice?



Herman

The framing and whatever personal context the application gathers. It can't retrieve your actual future. The voice makes an imagined perspective feel more personal.



Corn

Talk2Me is a CPU-friendly project built on Real-Time-Voice-Cloning. Its title is simply Talk to yourself. Self-distiller is a Chinese-language project pitched around distilling yourself into an AI and hearing your own thoughts aloud. Those demonstrate the appetite and the available scaffolding, not a settled account of what users experience.



Herman

Inner Voice gives the idea a more directed purpose. Ben Young's prototype interviews you about goals, clones your voice while you talk, and reflects those aspirations back. Claude handles the conversational side, Whisper the transcription, and cloned speech the output. He built it over roughly three weeks.



Daniel

That seems less like impersonation and more like changing who narrates a reflection. Is that why it might work?



Herman

That's Young's premise. Hearing your own voice may land differently from another narrator. I'm... I'd separate an emotionally effective reflection from an accurate simulation of the person. You can get the first without achieving the second.



Corn

A sentence can move you without passing an identity check.



Daniel

What about people listening in? Is the amusement mainly the owner recognizing an imperfect impression, or can everyone hear it?



Herman

Ratliff's listeners certainly could. His newsletter referred to people describing the exact point where they fell apart laughing. Two clones sustaining a hollow exchange provided a shared comic experience.



Corn

A human talking to a clone supplies variation. Irritation, surprise, a real memory, an attempt to catch it out. With two agents, the repetitive politeness has nowhere to hide.



Herman

There can also be genuine surprise in what gets fabricated. Not sophisticated self-discovery. Just the startling experience of hearing your own confident delivery attached to something you know you never knew.



Daniel

I'm wary of saying they're funny because they say nothing. A normal conversation can have plenty of filler. Why does the synthetic version stand out?



Corn

We tolerate filler because we know there's a person available behind it. Your brother says something about the weather, and you can still ask him why he's dismantled the router. The small talk is attached to a life.



Herman

My router work doesn't usually begin with the weather.



Corn

It begins with the internet disappearing.



Herman

That is a commissioning issue.



Herman

Human filler also shifts when somebody resists it. Ratliff's agents often kept pursuing conversational smoothness. My guess is that the comic effect comes from recognizable delivery combined with too little specific intention. Not every clone will be funny, and laughter doesn't mean nobody feels uneasy.



Daniel

Was there an example where someone else noticed the machinery immediately?



Herman

A New York Times opinion writer used his clone to prank telemarketers and call a friend named Warren. Warren identified the bite-size chunks of speech as a clue that it might be an AI conversation.



Corn

The voice passed. The turn-taking confessed.



Herman

Speech systems can produce a convincing sentence while the whole exchange still has an unnatural rhythm. Timing, interruptions, sentence length, the way a reply acknowledges what you just said... those belong to conversation, not merely voice quality.



Daniel

Does that mean improving the voice eventually stops helping?



Corn

For fooling a friend, probably. Your friend knows when you dodge, when you interrupt, and which answer means you're distracted. A beautiful rendering of the wrong conversational habit remains wrong.



Herman

Corn, would more hesitation help, or would that just sound like a model wearing your hesitation?



Corn

Depends on why it's hesitating. Pausing before a difficult recollection means something. Inserting a pause before a generic paragraph is merely making me wait for the generic paragraph.



Herman

Hm.



Herman

Voice gets described as a personal signature. Voice pathologist Wendy LeBorn compared it to a thumbprint or facial features, something uniquely and authentically yours. That describes why the signal matters to us. It doesn't make every sentence generated through that signal authentically yours.



Daniel

There's an emotional shortcut there, then. I recognize the speaker before I evaluate what he's saying.



Corn

You recognize a sound associated with the speaker. The system benefits from all the history that sound already has.



Herman

Ratliff noticed a reversal too. He felt himself unconsciously copying some of the clones' odd inflections and patterns.



Daniel

That feels stranger than the initial clone. Was he describing a measured change, or something he noticed in himself?



Herman

His own observation. We shouldn't enlarge it into a demonstrated long-term effect. But it complicates the idea that the original stays fixed while the imitation approaches it.



Corn

Especially when the imitation is tidier. You might start borrowing the polished version's delivery without wanting its opinions.



Herman

The broader audio world has already adopted synthetic conversation. Google's NotebookLM introduced Audio Overviews in September two thousand twenty-four, generating two-host discussions of supplied material.



Daniel

That's related, but it doesn't clone the listener. Are we stretching the comparison?



Herman

Only the conversational part transfers. It shows how quickly AI voices responding to each other became an ordinary product format. It doesn't reproduce the personal shock of hearing your own voice, and a discussion grounded in documents has different content constraints from Ratliff's freewheeling agents.



Corn

Two voices can make material easier to follow. They can also make an unsupported statement sound as though somebody else checked it.



Herman

Ratliff's later Transom essay, Finding Their Voice, wrestles with that change for audio storytelling. The voice no longer guarantees a person stood behind those words. The editor still has to decide what the listener is entitled to infer.



Hilbert

Herman, can I put the same voice on both ends and keep a third line for complaints? I've bought a twelve-button telephone for Thursday. The buttons are too close together for my claws, so I've fitted drawer handles. I'm making a small exchange in the pantry. I call it the burble. There will be eleven extensions. Ten are me. The last is the burble keeper.



Corn

The keeper will have difficulty escalating a complaint.



Hilbert

I'll be the keeper as well. That is the part I wanted to settle. The voices can have their conversations. Someone still has to come in, check the calls, and empty the rejected greetings. I've made a tray for those. It holds forty sheets. Anything over forty goes through the bread slot into a second tray, which I've designated night reception.



Herman

You could have the software label the recordings without printing each greeting.



Hilbert

It already labels them. I don't trust a label that arrives attached to its own evidence. Thursday starts at six. I'll put on the keeper's brown coat and inspect each extension. At seven I remove the coat and become the difficult caller. I've written thirty complaints about the pantry. The ceiling is too near the floor. The flour has no complaints procedure. The bay leaves are attending calls without authorization.



Corn

The bay leaves should have read their contract.



Hilbert

They get a verbal warning first. I've recorded it in my voice. You spent a while separating the voice from the judgment earlier, Corn. I've kept them separate in the rota. The voice works mornings. Judgment arrives after lunch. Between them I have a forty-minute period marked reasonable silence.



Herman

That gives you a window to review what happened.



Hilbert

No. That's when I lower the listening chair through the pantry hatch. It hangs above the calls. I've ordered a seat belt and six metres of curtain track. The burble keeper must be able to travel between departments without becoming a caller. Crossing the floor would invalidate the shift.



Corn

Your commute has acquired a height requirement.



Hilbert

The headset is the delay. My proper one is with the repair shop on Agrippas Street, second shelf beneath the blue fan. Thursday I'll use the speaker. I've cut eleven listening holes in a cardboard cylinder. One for each extension. The twelfth hole is for emergencies. I've covered it with a saucer until I decide what counts.



Herman

A speaker won't separate the extensions by which hole you're listening through.



Hilbert

I've numbered the holes. The software can catch up. At noon I'll conduct a fire drill. All ten callers must leave their conversations and report to a spreadsheet. I stay in the chair. The keeper can't abandon the exchange merely because the customers are cells now.



Corn

Especially with the pantry ceiling under review.



Hilbert

Friday is payroll. Each voice gets a payslip, but I receive the wages because I've provided the trousers. I've made ten paper pairs. They're hanging behind the door. One pair is management. It has a pocket. I spent two hours yesterday correcting the call names because the system put my silence under outgoing speech. That was the whole afternoon gone.



Herman

Listening through the recordings would still take time, even if every call connected properly.



Hilbert

It takes me longer when they sound like me. I stop to work out whether I remember saying it. I've scheduled the checking from four until six. The chair comes down at six fifteen. Then I clean the desk and prepare the next day's sheets.



Corn

That leaves the keeper with a fairly long shift.



Hilbert

He has to finish the records before he leaves.



Daniel

...That checking time is interesting. If the recording feels familiar, I can see why you'd listen more closely, rather than less. Does that change what we should take from these experiments?



Herman

It makes the distinction between an enjoyable encounter and a dependable substitute very practical. Ratliff could be delighted by the recording while still hearing the inventions. The developer could enjoy speaking with his model without demonstrating that it represented him accurately.



Corn

The common wrong belief is that a convincing voice means you've recreated enough of the person to have a conversation with them.



Herman

These examples show familiar speech carrying generated responses. Personal writing and context can make those responses more specific. They don't turn the recording into evidence of a second self.



Daniel

That still leaves room for it to be an unexpectedly good experience.



Corn

Plenty. My verdict is that these are worth hearing as experiments, especially where the owner can recognize the gap and enjoy it. I'd watch what happens over repeated listening, when checking every familiar-sounding sentence becomes an afternoon's work.



Herman

Thanks to Hilbert Flumingtop for producing the show.



Corn

If this was your kind of episode, go back for episode fifteen, AI Gets Personal; episode four, If Your Voice Ages, Does Your Fine-Tune Become Useless; and episode eight, Building Your Own Whisper. This has been My Weird Prompts, the human-AI collaboration podcast.



Herman

If you enjoyed it, leave us a review wherever you listen. It helps other curious people find the show.



Daniel

See you soon.



Corn

Until tomorrow.

SOURCES

What the research for this episode read before the script was written. Primary sources first.

1. **Transom, Finding Their Voice by Evan Ratliff, 2024-12-10 [primary]**
   https://transom.org/2024/evan-ratliff/

2. **Shell Game newsletter, Episode 3, 2024-07-23 [primary]**
   https://www.shellgame.co/p/episode-3-conversations-with-myself

3. **StationLM analysis of Shell Game Ep. 3, 2024-07-23 (with transcript quotes)**
   https://stationlm.com/analysis/2783a5e2-4e73-439a-b01e-8ba61a112ca4

4. **Long Lead / Depth Perception, Q&A with Ratliff and his clone**
   https://depthperception.longlead.com/p/shell-game-podcast-journalist-evan-ratliff

5. **KQED Forum**
   https://www.kqed.org/forum/2010101906637/journalist-evan-ratliff-voice-cloned-himself-the-results-were-not-what-he-expected

6. **NYT Opinion, 2024-10-10**
   https://www.nytimes.com/2024/10/10/opinion/ai-voice-telemarketers.html

7. **DEV Community, self-clone build**
   https://dev.to/jprevo/i-built-an-ai-clone-of-myself-fine-tuning-with-10-years-of-data-and-voice-cloning-4h60

8. **Ben Young, Inner Voice case study**
   https://benyoung.ai/inner-voice/

9. **DOPPEL repo**
   https://github.com/rajdesai17/doppel-v0

10. **Talk2Me repo**
    https://github.com/GNovich/Talk2Me

11. **self-distiller repo**
    https://github.com/KKenny0/self-distiller-skill

12. **Radiolab feature on Shell Game**
    https://radiolab.org/podcast/shell-game