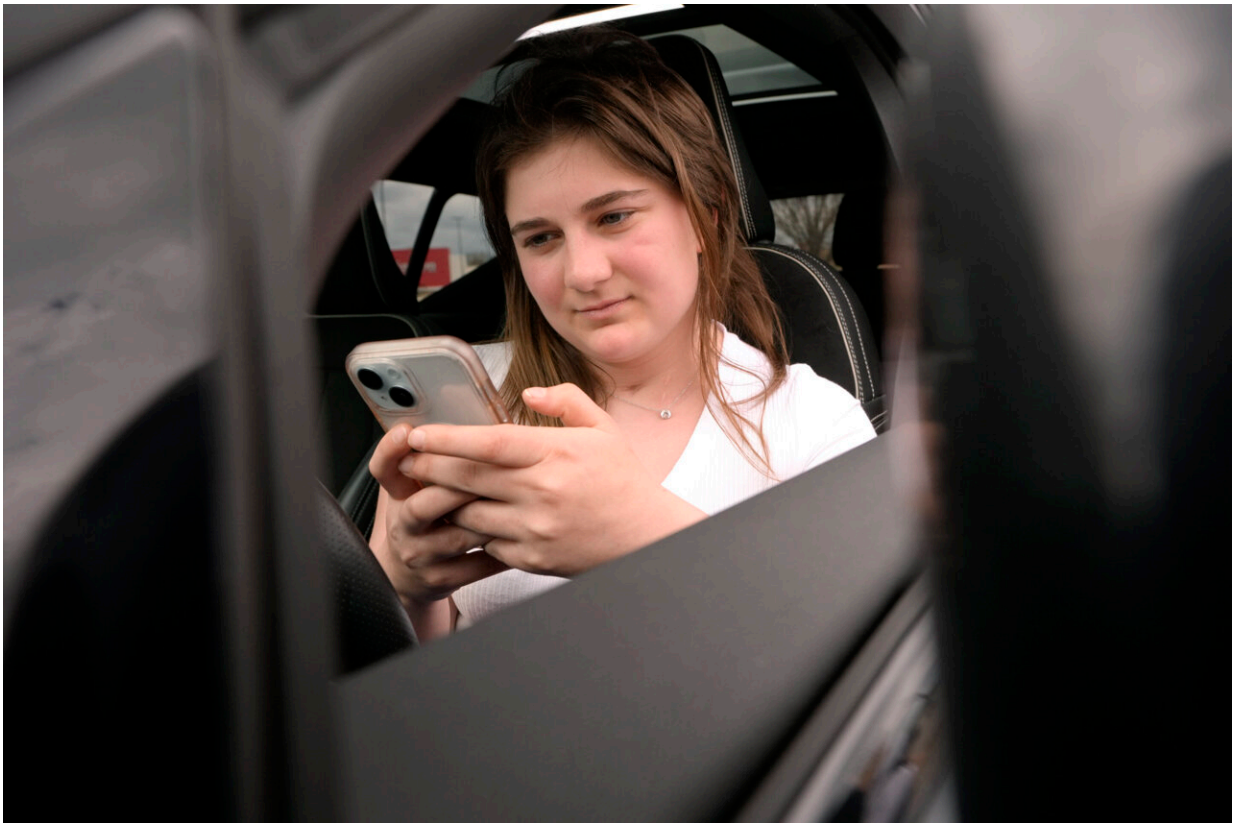
Illness took away her voice. AI created a replica she carries in her phone

May 13 2024, by Matt O'brien



Alexis Bogan, whose speech was impaired by a brain tumor, uses an AI powered smartphone app to create a audible drink order at a Starbucks drive-thru on Monday, April 29, 2024, in Lincoln, R.I. The app converts her typed entries into a verbal message created using her original voice. Credit: AP Photo/Steven Senne

The voice Alexis "Lexi" Bogan had before last summer was exuberant.

She loved to belt out Taylor Swift and Zach Bryan ballads in the car. She laughed all the time—even while corralling misbehaving preschoolers or debating politics with friends over a backyard fire pit. In high school, she was a soprano in the chorus.

Then that voice was gone.

Doctors in August removed a life-threatening tumor lodged near the back of her brain. When the breathing tube came out a month later, Bogan had trouble swallowing and strained to say "hi" to her parents. Months of rehabilitation aided her recovery, but her speech is still impaired. Friends, strangers and her own family members struggle to understand what she is trying to tell them.

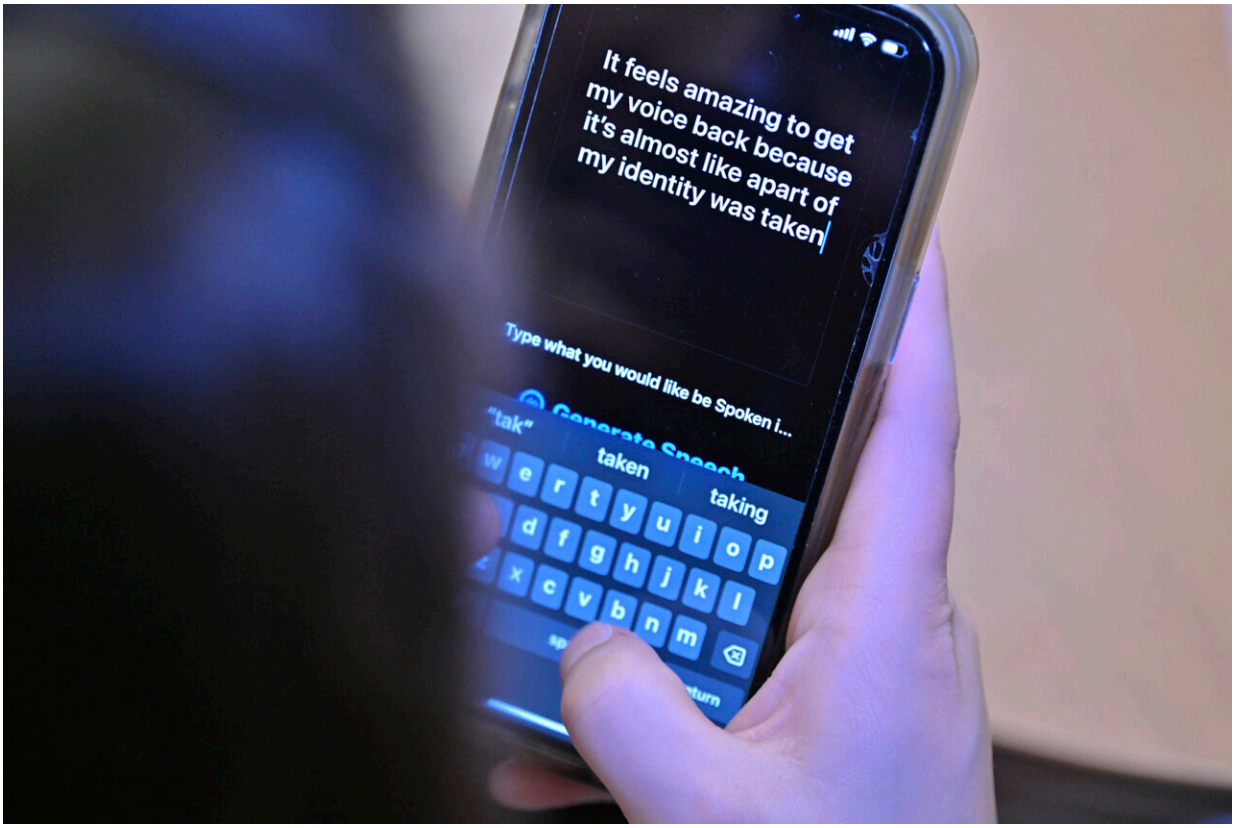
In April, the 21-year-old got her old voice back. Not the real one, but a voice clone generated by artificial intelligence that she can summon from a phone app. Trained on a 15-second time capsule of her teenage voice—sourced from a cooking demonstration video she recorded for a high school project—her synthetic but remarkably real-sounding AI voice can now say almost anything she wants.

She types a few words or sentences into her phone and the app instantly reads it aloud.

"Hi, can I please get a grande iced brown sugar oat milk shaken espresso," said Bogan's AI voice as she held the phone out her car's window at a Starbucks drive-thru.

Experts have warned that rapidly improving AI voice-cloning technology can amplify phone scams, disrupt democratic elections and violate the dignity of people—living or dead—who never consented to having their

voice recreated to say things they never spoke.



Alexis Bogan types a response to a reporter's question with an app which approximates her lost voice, Thursday, March 11, 2024, at Rhode Island Hospital in Providence, R.I. Doctors treating Bogan, whose speech was impaired by a brain tumor, used a voice-cloning tool from OpenAI to recreate her previous voice. Credit: AP Photo/Josh Reynolds

It's been used to produce deepfake robocalls to New Hampshire voters mimicking President Joe Biden. In Maryland, authorities recently charged a high school athletic director with using AI to generate a fake audio clip of the school's principal making racist remarks.

But Bogan and a team of doctors at Rhode Island's Lifespan hospital group believe they've found a use that justifies the risks. Bogan is one of the first people—the only one with her condition—who have been able to recreate a lost voice with OpenAI's new Voice Engine. Some other AI providers, such as the startup ElevenLabs, have tested similar technology for people with speech impediments and loss—including a lawyer who now uses her voice clone in the courtroom.

"We're hoping Lexi's a trailblazer as the technology develops," said Dr. Rohaid Ali, a neurosurgery resident at Brown University's medical school and Rhode Island Hospital. Millions of people with debilitating strokes, throat cancer or neurogenerative diseases could benefit, he said.

"We should be conscious of the risks, but we can't forget about the patient and the social good," said Dr. Fatima Mirza, another resident working on the pilot. "We're able to help give Lexi back her true voice and she's able to speak in terms that are the most true to herself."

Mirza and Ali, who are married, caught the attention of ChatGPT-maker OpenAI because of their previous research project at Lifespan using the AI chatbot to simplify medical consent forms for patients. The San Francisco company reached out while on the hunt earlier this year for promising medical applications for its new AI voice generator.

Bogan was still slowly recovering from surgery. The illness started last summer with headaches, blurry vision and a droopy face, alarming doctors at Hasbro Children's Hospital in Providence. They discovered a vascular tumor the size of a golf ball pressing on her brain stem and entangled in blood vessels and cranial nerves.

"It was a battle to get control of the bleeding and get the tumor out," said pediatric neurosurgeon Dr. Konstantina Svokos.

The tumor's location and severity coupled with the complexity of the 10-hour surgery damaged Bogan's control of her tongue muscles and vocal cords, impeding her ability to eat and talk, Svokos said.

Alexis Bogan, whose speech was impaired by a brain tumor, uses mobile phone with an app that features a voice-cloning tool to order a drink at a Starbucks drive-thru Monday, April 29, 2024, in Lincoln, R.I. Doctors treating Bogan are recreating her original voice using a voice-cloning tool from OpenAI. Credit: AP Photo/Steven Senne

"It's almost like a part of my identity was taken when I lost my voice," Bogan said.

The feeding tube came out this year. Speech therapy continues, enabling her to speak intelligibly in a quiet room but with no sign she will recover the full lucidity of her natural voice.

"At some point, I was starting to forget what I sounded like," Bogan said. "I've been getting so used to how I sound now."

Whenever the phone rang at the family's home in the Providence suburb of North Smithfield, she would push it over to her mother to take her calls. She felt she was burdening her friends whenever they went to a noisy restaurant. Her dad, who has hearing loss, struggled to understand her.

Back at the hospital, doctors were looking for a pilot patient to experiment with OpenAI's technology.

"The first person that came to Dr. Svokos' mind was Lexi," Ali said. "We reached out to Lexi to see if she would be interested, not knowing what her response would be. She was game to try it out and see how it would work."

Bogan had to go back a few years to find a suitable recording of her voice to "train" the AI system on how she spoke. It was a video in which she explained how to make a pasta salad.

Her doctors intentionally fed the AI system just a 15-second clip. Cooking sounds make other parts of the video imperfect. It was also all that OpenAI needed—an improvement over previous technology requiring much lengthier samples.

Dr. Rohaid Ali plays a video from a high school project made by his patient Alexis Bogan on Thursday, March 11, 2024, at Rhode Island Hospital in Providence, R.I. Doctors treating Bogan, whose speech was impaired by a brain tumor, used the recorded sample of her speech and a voice-cloning tool from OpenAI to recreate her previous voice. Neurosurgeon Dr. Konstantina Svokos, right, looks on. Credit: AP Photo/Josh Reynolds

They also knew that getting something useful out of 15 seconds could be vital for any future patients who have no trace of their voice on the internet. A brief voicemail left for a relative might have to suffice.

When they tested it for the first time, everyone was stunned by the quality of the voice clone. Occasional glitches—a mispronounced word, a missing intonation—were mostly imperceptible. In April, doctors

equipped Bogan with a custom-built phone app that only she can use.

"I get so emotional every time I hear her voice," said her mother, Pamela Bogan, tears in her eyes.

"I think it's awesome that I can have that sound again," added Lexi Bogan, saying it helped "boost my confidence to somewhat where it was before all this happened."

She now uses the app about 40 times a day and sends feedback she hopes will help future patients. One of her first experiments was to speak to the kids at the preschool where she works as a teaching assistant. She typed in "ha ha ha ha" expecting a robotic response. To her surprise, it sounded like her old laugh.

She's used it at Target and Marshall's to ask where to find items. It's helped her reconnect with her dad. And it's made it easier for her to order fast food.

Bogan's doctors have started cloning the voices of other willing Rhode Island patients and hope to bring the technology to hospitals around the world. OpenAI said it is treading cautiously in expanding the use of Voice Engine, which is not yet publicly available.

A number of smaller AI startups already sell voice-cloning services to entertainment studios or make them more widely available. Most voice-generation vendors say they prohibit impersonation or abuse, but they vary in how they enforce their terms of use.

Alexis Bogan, center, and her mother Pamela Bogan, right, react to hearing a recreation of her lost voice from a prompt typed by Dr. Fatima Mirza, left, on Thursday, March 11, 2024, at Rhode Island Hospital in Providence, R.I. Doctors treating Bogan, who's speech was impaired by a brain tumor, used a voice-cloning tool from OpenAI to recreate her previous voice. Credit: AP Photo/Josh Reynolds

"We want to make sure that everyone whose voice is used in the service is consenting on an ongoing basis," said Jeff Harris, OpenAI's lead on the product. "We want to make sure that it's not used in political contexts. So we've taken an approach of being very limited in who we're giving the technology to."

Harris said OpenAI's next step involves developing a secure "voice

authentication" tool so that users can replicate only their own voice. That might be "limiting for a patient like Lexi, who had sudden loss of her speech capabilities," he said. "So we do think that we'll need to have high-trust relationships, especially with medical providers, to give a little bit more unfettered access to the technology."

Bogan has impressed her doctors with her focus on thinking about how the technology could help others with similar or more severe speech impediments.

"Part of what she has done throughout this entire process is think about ways to tweak and change this," Mirza said. "She's been a great inspiration for us."

While for now she must fiddle with her phone to get the voice engine to talk, Bogan imagines an AI voice engine that improves upon older remedies for speech recovery—such as the robotic-sounding electrolarynx or a voice prosthesis—in melding with the human body or translating words in real time.

She's less sure about what will happen as she grows older and her AI voice continues to sound like she did as a teenager. Maybe the technology could "age" her AI voice, she said.

For now, "even though I don't have my voice fully back, I have something that helps me find my voice again," she said.

Citation: Illness took away her voice. AI created a replica she carries in her phone (2024, May 13) retrieved 16 August 2024 from https://techxplore.com/news/2024-05-illness-voice-ai-replica.html