

FINDING A VOICE

This season, the first recipients of an innovative new technology for the speech-impaired will begin to talk in tones all their own.

By Gemma Z. Price Illustration by Shout

“
At a conference,
Dr Rupal Patel
had watched a
little girl and an
adult man have
a conversation—
with the same
voice. It looked
and sounded
all wrong.

”



“MY NAME IS Erin Flack, I am 9 years old, and I live in Needham, Massachusetts. My little sister Maeve has cerebral palsy and uses an ECO speaking device we call her talker.” This was the email that landed in Dr Rupal Patel’s inbox in March 2014, shortly after the broadcast of her TED Talk about VocaliD, a new project to create custom voices for speech-impaired people. Attached to the message was a photograph of three young girls, each blonde and smiling. “Me, my mom and my other sister, Meghan, realised that Maeve’s voice she has programmed in her talker sounds nothing like the voice of an 8-year-old and we probably should get a new voice. . . . I saw your video and I thought I would love to give my voice to someone who can’t talk, just like Maeve,” Erin wrote.

In her TED Talk, Rupal Patel, PhD, CEO and founder of VocaliD, described how at a 2002 conference she’d watched a little girl and an adult man have a conversation—with

the same voice. It looked and sounded all wrong. Since the first synthetic voice device prototypes were developed in the mid-1970s—think the Equalizer software adopted by Stephen Hawking in the ’80s—technology has advanced by leaps and bounds. But the range of voices available to users has not. In the U.S. alone, 2.5 million people use speech generating devices (SGDs), and they still have only a handful of voices to choose from.

As a speech scientist, Patel understands that no matter how little of their speaking ability remains, each individual retains a unique vocal identity: They can still modulate their “source” (the pitch, tempo and loudness of the vibrations they generate using their voice box), even if they can’t properly control their “filter” (their tongue, lips and mouth) to produce consonants and vowels. Hence, she thought of combining a source sample from a speech-impaired individual with a filter sample from a matching

“ Since the first synthetic voice device prototypes were developed, technology has advanced by leaps and bounds. But the range of voices available to users has not. ”

healthy talker—a donor of similar age, background, size and so on—to create a new synthetic voice as unique as a fingerprint.

Unlike such procedures as donating a kidney, donating your voice is painless: You simply have to record a few hours of speech. When VocaliD was still a research project, all donors had to visit Patel’s Boston-based recording studio. As the project and donor numbers grew, this became unfeasible. Patel then founded VocaliD and launched its first product, The Human Voicebank, a crowdsourced effort to collect the voices of people

of all ages, ethnicities and cultural/linguistic backgrounds from all over the world. (Currently the programme is only in English, but other languages are also in the works, starting with research into speech synthesis in Korean and Japanese.) The launch of The Human Voicebank web portal in December 2014 leveraged the increasing quality of microphones on laptops, tablets and smartphones so that donors could record their voice anywhere, on their own time.

Completing a recording means reading 3,500 sentences aloud. These aren’t designed to cover every phrase that the recipient will say, but to cover all the combinations of vowels and consonants in a language. Once a recording is made, the VocaliD software can parse the utterances into different sound combinations and populate a data set, called a voice bank. The VocaliD algorithm can draw from the voice bank to synthesise a natural-sounding voice, infused with the recipient’s “vocal DNA.” While the donor must record a fairly large data sample, the recipient only has to extend or repeat a vowel sound, meaning that anyone who uses a speech generating device can now have a voice of their own.

When Patel invited her TED Talk audience to visit the website to learn more, several thousand people from all walks of life, from children to professional voice talents, came forward in the first couple of months; today more than 14,000 people are active on The Human Voicebank platform. The team hopes to bank 10,000 voices by the end of 2016. Once a voice is banked successfully, it can be matched with multiple



above: Dr Rupal Patel shares her work in a December 2013 TED Talk.

recipients, and the VocaliD team gets to work. “It takes us around 20 to 30 hours per person to create a voice,” Patel says. “These are the human hours needed after an algorithm has mixed donor speech with recipient speech.”

The first seven recipients receive their voices in December 2015. Maeve Flack is among them, with her sister Erin as her donor. The next 50 recipients will receive their voices in December 2016, thanks to VocaliD’s “In Demand” campaign on crowdsourcing site Indiegogo. And another 50 personalised voices are now available on the site for \$1,249 each—significantly less than the cost of production, which is about \$7,000 to \$10,000 per voice.

VocaliD’s future goals include extending the software compatibility from Android to iOS and Windows (the platform on which most speech devices are built), adding a range of accents, and automating the tweaking process so that anyone—not just speech scientists—can finesse a new voice, enabling VocaliD to deliver faster. The team is also working on offering a personal connection between donors and recipients, while being respectful of both parties’ wishes regarding use of their personal information.

Hearing Maeve’s words in a VocaliD video makes it clear what motivates Patel and her team. “My voice isn’t me; it sounds like a robot,” Maeve says of her current voice. “I want it to be me.” The first thing Maeve plans to say with her new voice? “I feel happy.”

Gemma Z. Price, based in Saigon and San Francisco, writes for Condé Nast Traveller, Departures and Robb Report.

SPEAK UP

To donate your voice to **The Human Voicebank**, log on to vocalid.co/voicebank and begin recording. All you need is a computer with Google Chrome, internet connectivity, a microphone and a quiet room.