

# How People Who Can't Speak Can Sound Like...Themselves

---

## **For those who rely on electronic communication devices, new technology aims to deliver voices that are more natural and less robotic**

A wave of new technologies is giving people like Max Plansky, who are unable to speak due to a debilitating condition, a more personal synthetic voice. Photo/Video: Denise Blostein/The Wall Street Journal

Amy Dockser Marcus

Updated Feb. 15, 2016 10:11 p.m. ET

0 COMMENTS

Max Plansky is 16 years old, and like any teenager, he is trying to find his own voice.

But for Max, the effort is complicated by the fact that he needs to use a communication device to talk.

At nine months, Max was diagnosed with cerebral palsy, a neurological disorder. His vocal cords vibrate, and Max can make sounds. But the area of his brain affected, which controls his muscles, also controls his ability to speak. When Max wants to talk, he chooses words and sentences on an electronic device, which then speaks in a computerized voice often known as "Perfect Paul."

Perfect Paul is a common option on devices like the one Max uses. But it doesn't sound anything like a 16-year-old boy. It sounds more like an adult male robot.

But only recently have advances in computers and other technologies made it possible to analyze the characteristics and more ethereal aspects of human voices

and mix them together to the point where an artificial voice doesn't sound so, well, artificial.

Most of these efforts are commercial and geared toward improving the synthetic voices that give you directions or speak to you on the phone when you try to make a flight reservation. But researchers also are starting to apply the technology to help people, like Max, who due to medical conditions are forced to rely on electronic communication devices. They want a voice that not only sounds more natural, but that manages to capture who they are.

### **'A window'**

“Speech is not just a means of communication—it is a window into the soul,” says Matthew Aylett, chief scientific officer of an Edinburgh-based company, CereProc, that sells technology for personalizing synthetic voices.

To capture inflection and tone in a synthetic voice is an enormous challenge, Dr. Aylett says. People speak faster, louder or at a higher pitch when they are upset, or slower, softer and deeper when they are sad. The stronger the emotion, the harder it is to simulate, he says. “Telling a joke is also tough.”



---

HOOP IT UP | Max Plansky goes public with his new voice in the Northeastern University men's basketball locker room. Photo: Matthew MODOONO/Northeastern University

Another thing that is hard to capture, says Dr. Aylett, is the myriad ways a voice can be used to convey different meanings of the same word. "When the waiter asks if you have decided what you want," he says, "it means something different if you say 'Yes!' or 'Yessss.' "

Building Max a personalized voice is not only a technical issue; it also involves questions about his identity. What sounds might capture the different aspects of Max's personality? Moreover, if Max is to have a synthesized voice, should it sound like a 16-year-old speaking, or should it sound more like who Max might become later?

These are issues that Rupal Patel, a speech-technology professor on leave from Northeastern University, has been thinking about for years. Dr. Patel is chief executive of VocaliD, a Belmont, Mass., company that custom-builds voices for the speech-impaired. She became interested in personalizing synthetic voices, she says, after attending an assistive-technology conference where she heard a young girl and an older man conversing by means of the same synthetic voice.

"How come this young girl has the same voice as this adult male?" Dr. Patel recalls wondering.

### **The right voice**

Dr. Patel's company is focused on finding ways to extract an individual's unique sounds and blend them with the closest match from donor voices stored in a voice bank she is building. The idea, she says, is to capture the personality of the recipient and the clarity of the donor.

Last fall, she and VocaliD began work on one of their first customized voices, for Max. It is a process that can still be painstaking and slow. VocaliD begins by recording bits and pieces of vocalizations and sounds from Max. It then draws on contributions to the voice bank, some of which were made through the efforts of a nonprofit, You're

With Us!, started by Max's father, Michael Plansky.

As with many startups, costs are initially high. Max's voice will be one of the first created by the company. The family agreed to pay \$10,000—which they raised with help from a generous uncle and aunt and other donors—in return for three voices Max could choose from and ensured delivery by the end of 2015. With more experience and infrastructure in place, VocaliD is now charging customers \$1,249 plus an annual fee of \$240 to tune and modify the voice over time.

Max recognizes that even with a more natural-sounding voice, it is still going to be difficult for him to communicate. The controls for his new voice will be the same as for his old—a switch mounted on the headrest of his wheelchair, and menus of words and sentence that he must choose from. “I don't think it will make communication easier,” he says. “But I do think I will like how my new voice sounds more.”

A big part of Max's life these days revolves around sports, particularly the Northeastern University Huskies men's basketball team. At the university gym, he has his own locker (M. Plansky, No. 25) and an endless supply of team T-shirts. He attends practice and cheers from the bench during games. Last year, when the team won a conference championship, Max was presented with his own ring by the captain and the point guard.

Dr. Patel says these things are never far from her mind as she and Geoff Meltzner, VocaliD's director of R&D, prepare the voices for Max to choose from. Played for a reporter, the first voice is clear but still a little robotic. The second is higher and boyish. Dr. Patel says it reminds her of Max as he is now, with his smooth face and wide smile. The third is deeper, as if it belongs to someone a little older, the person Max might be someday.

On a December morning at his school, Max must choose. He listens to the three voices and makes his preference known. Just as important, however, is how others will hear Max's voice. Later that day, he and a select group head to the Northeastern gym, where Max has decided he will publicly use his new voice for the first time after the basketball team finishes its practice.

Dr. Patel feels nervous, she says. “The voice tells secrets and has biases,” she explains. When people hear a voice, they make judgments about whether someone is

intelligent, extroverted, introverted or attractive. It isn't easy to figure out how to acoustically represent any of these qualities. "How can we know this is Max's voice?" she wonders, as everyone arrives at the gym.

## The unveiling

The locker room fills with players, Max's family and friends, Dr. Patel and her colleagues, among others. The players flank Max, towering over him. Then the buzz in the room settles for a moment.

It is Max's turn to speak: "Getting a voice from VocaliD is going to be a great opportunity for me," he says. People start to clap and cheer.

Max has chosen the older-sounding voice, the one that speaks to the future.

"He sounds compassionate," Mr. Plansky says. "You feel his soul."

To Zach Stahl, the Huskies' captain, he sounds "cheerful." "Max lights up when we see him," Mr. Stahl says.

At the start of the project, Max mused that the voice being made for him might end up sounding like his father's. But when asked about it now, Max says, "I sound like me."

Dr. Patel once described the technology she is working on by saying, "A new voice is born." Maybe that is why in the locker room, after trying out his voice, Max does what many people do when they witness a birth. He stops talking and begins to cry.

*Ms. Dockser Marcus is a Wall Street Journal reporter in Boston. Email her at [amy.marcus@wsj.com](mailto:amy.marcus@wsj.com)<sup>[1]</sup>.*

---

## Giving Voice

Several efforts are under way to provide more natural synthetic voices for people who rely on communication devices.

Here's how one approach, from VocaliD,  
a Massachusetts startup works

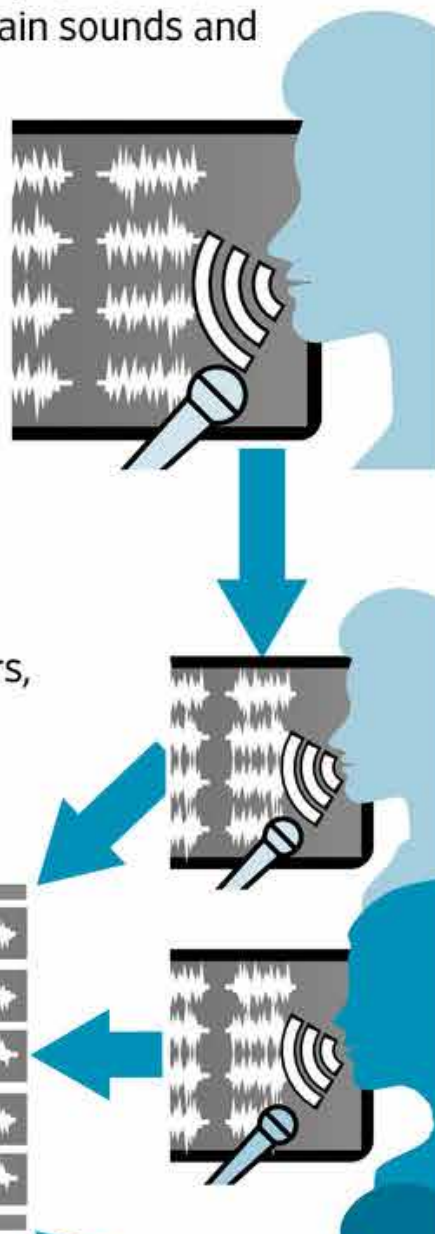
a Massachusetts startup, works.

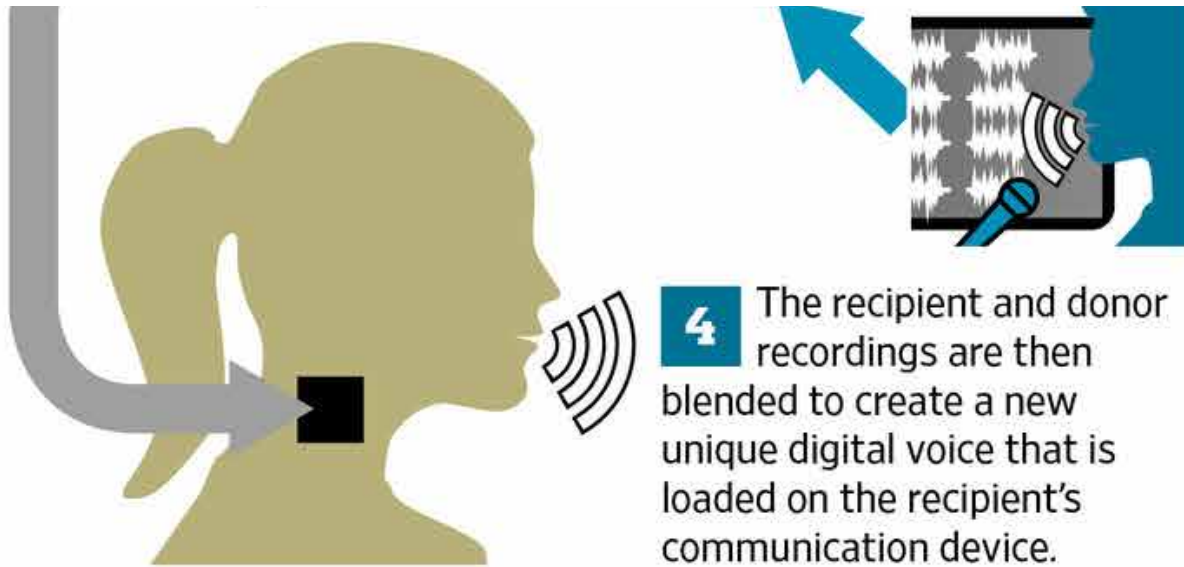


**1** The person who is getting the personalized voice sends a two-to three-second sample of his or her vocalization. The digital recordings are made on a computer or other recording device

**2** Speech donors contribute their voices to the company's voice bank. They can log onto a website from home or another quiet location and record sentences that contain sounds and sound combinations that make up the English language. There are 3,500 sentences. Recordings can take up to seven hours to complete over several days.

**3** The company uses proprietary algorithms to search the voice bank and find a matched donor. The algorithm considers region, age and gender, among other factors, as well as vocal characteristics such as tone, clarity and sound to try to find a match.





Source: VocaliD

ENLARGE

THE WALL STREET JOURNAL.

1. <mailto:amy.marcus@wsj.com>