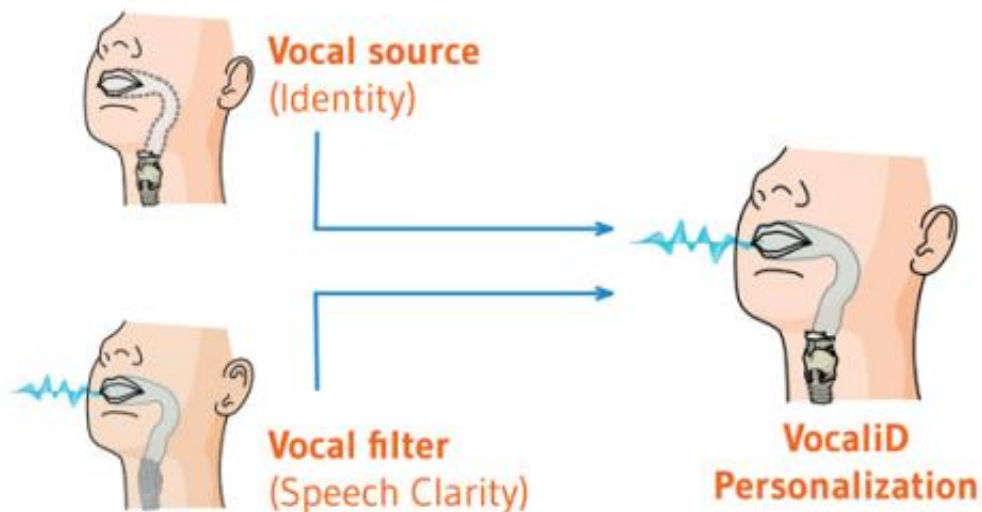


Researcher gives subjects their voice

February 21 2013, by Angela Herring



Credit: Rupal Patel.

Stephen Hawking and a 9-year-old girl with a speech disorder most likely use the same synthetic voice. It's called Perfect Paul and it's easy to understand, especially in acoustically chaotic environments like classrooms full of children. While new, more natural-sounding voices are available, [Perfect Paul remains the most oft-used synthetic voice](#) in the community of disordered speakers.

But Perfect Paul conveys none of the personality inherent in vocal identity, explains Rupal Patel, an associate professor of computer science and speech language pathology and audiology.

"What we're trying to do is improve the quality," she said, "but also increase the personalization of those voices, by not just making it a little kid's voice, but making it that little kid's voice."

Backed by a grant from the National Science Foundation, Patel and her research team are developing ways to create personalized synthetic voices that resemble users' vocal identities while remaining as understandable as those of the healthy donors.

In the first iteration of the project, which Patel calls [VocaliD](#) (pronounced vocality, for Vocal Identity), her team computationally merged the acoustics of a sustained vowel sound from a child with a speech disorder, like this:

with the acoustics of a full sentence spoken by a healthy speaker of the same demographic, like this:

The result is a clear, synthetic voice with the personality of the end user:

These voices have already elicited great responses from parents; one said, "If [my son] had been able to talk, this is what he would sound like." However, the early version of VocaliD used a difficult-to-scale approach that is not easily reproducible. Patel said, "We'd like to be able to allow users to create new voices as they mature in the same way a natural voice evolves."

With the support of another grant from the National Science Foundation, her team is currently adding physiological information on top of the acoustics. "When you hear speech, it's a combination of your source and your filter," Patel said. The source, she explained, derives from the voice

box in the larynx whereas the filter is determined by the shape and length of the vocal tract.

Vocal characteristics—such as pitch, breathiness, and loudness—all emerge from the vocal folds in the larynx and give rise to vocal identity. Modulating those features by changing the shape of our mouths and moving our tongues gives rise to distinct vowel and consonant sounds, which, Patel said, are typically impaired in disordered speech.

Using data from a set of sensors placed on participants' tongues and mouths, the researchers will determine the most efficient way to approximate the physical aspects of the disordered speaker's vocal tract. They can then add this information into the voice-synthesis software to create voices that will grow and change as the users mature.

The academic community has long accepted the source-filter theory of speech, but more work needs to be done in order to understand it, according to Patel, especially as researchers develop more advanced speech technologies for security and other applications.

Patel's work in particular also aims to inform basic research questions such as, "How much do both the source and filter contribute to the identity of a speaker's output?"

Patel's software is compatible across assistive technology platforms, including mainstream touch-pad devices, a feature she hopes will increase its adoption within the community. Patel speculates that assistive communication devices will eventually appeal to healthy people as a new way of learning, communicating, and interacting.

"The iPad revolution is helping to break down barriers and increasing the emphasis on user interface issues," said Patel, who has been working to improve assistive communication technologies for more than 16 years.

"Lots of kids, both healthy and impaired, are using screens to interact now."

Provided by Northeastern University

Citation: Researcher gives subjects their voice (2013, February 21) retrieved 20 April 2024 from <https://medicalxpress.com/news/2013-02-subjects-voice.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.