Speech synthesis research brings new hope to the impaired

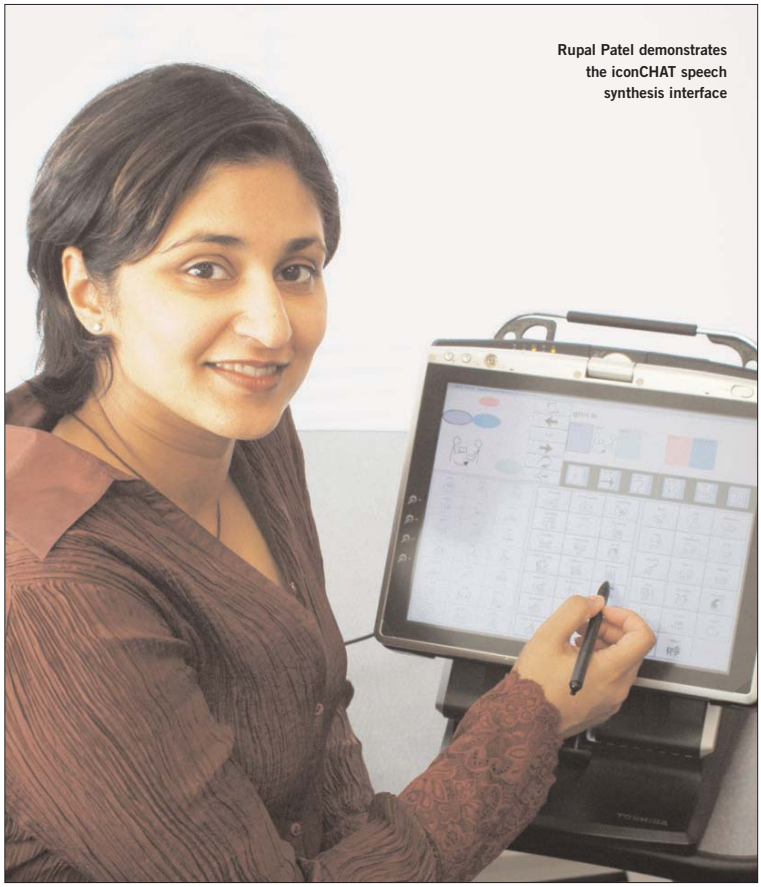
Are you talking to me?

BY SUSAN SALK

It seems unfair that the computerized voice in an automobile navigation system speaks more concisely and legibly than individuals who rely on speech synthesizers — sounding no better than the computer HAL from "2001: A Space Odyssey"to communicate their thoughts and desires.

research

Sentences from synthesizers are as basic as a child's. And the voice is often flat and unintelligible, especially in situations with loud background noises.



Rupal Patel demonstrates the iconCHAT speech synthesis interface

CRAIG BAILEY

But Rupal Patel , an assistant professor of speech language pathology and audiology, has heard the cries for help. She is attempting to build a better speech synthesizer for loud backgrounds, one that mirrors intonation, volume and rhythm of people trying to be heard against the backdrop of traffic or restaurant noise.

In January, she parlayed an $8,000 provost's office seed grant into $100,000 of National Science Foundation funding for her project, "Loudmouth: Toward Intelligible Speech Synthesis in Everyday Noise Situations."

Working in her newly dedicated 750-square-foot laboratory in102 Forsyth — the old Husky Card quarters has never looked so good with its now freshly painted red walls, new computers and two high-tech audio-testing booths — Patel hopes to advance the cause of assistive technology. The Communications Analysis and Design Lab (CadLab) is where her mission to improve the lives of the speech impaired will get cranking.

"I was really struck by the need for better technology when I attended a conference in which speech impaired people were using automated devices in a huge auditorium, and it was impossible for them to hear each other," Patel said. "What good is it if you have one of these devices and you still can't be heard?'

She will approach the problem by first studying the melody of speech prosody in non-impaired people having conversations over background noise, paying close attention to which words they emphasize and whether they elongate the pronunciation of others. Thinking about how she would convey her intentions to someone in a loud environment, she reasons, "If I wanted to go to the store, I would not yell out each of those words like, 'I. Want. To. Go. To. The. Store.' No. I would say something like 'I want to go to the store,'" she said.

The change in inflection and speaking styles to accommodate an acoustic environment, technically called the Lombard effect, often finds that people slow down their speech, speak louder, and enunciate nouns and verbs, while deemphasizing consonants, she said.

Recording the speech patterns is the first step toward her goal of creating a device that adapts itself to background noise and pronounces sentences with the strength, loudness and clarity of a regular speaking voice, she said. "We don't just want to crank up the volume on the synthesizers," she said. "Currently speech synthesizers are not capable of mimicking the Lombard effect."

Working hand-in-hand with that project, her lab filled with speech language and computer science students are working on all manner of projects, of which a new image-based communication aid called iconCHAT is a star. The computer tool, fitted to a touch-tablet computer, uses icons to and a touch wand to allow the speech impaired to create sentences. Although there is something similar on the market, Patel's method improves on the usability and quality of the sentences, she said. Her novel approach to sentence construction is more intuitive, and therefore faster, she added.

"We don't just want to crank up the volume on the synthesizers."

The system offers two-dimensional cartoon displays to more closely mirror the way a person thinks, she said. Because it will be attached to global positioning satellites, the computer will adapt its speech to time and location, she said, explaining that if a user is at a McDonald's restaurant, icons reflecting menu choices will appear on the screen.

A team of undergrad and grad students from both speech language pathology and audiology and computer engineering is busily working on all aspects of this research. Among the standouts are undergraduate student Elyes Yaich, computer science; graduate student Rajiv Radhakrishnan, computer science; graduate student Kevin Schell, speech and hearing; graduate speech and hearing students Carrie Watkins and Alexia Salata; and graduate student Eldar Sadikov, computer science.

Looking back on the life aspirations that landed her on Northeastern's campus in June 2003,with about $1 million in grant funding to her credit, Patel figures she will one day bring to fruition her two loves: Science and fashion. "I always loved science, but I was torn between that and fashion design," she said. "One day, though, I will invent a wearable computer for assistive technology—that is my dream."