What makes your voice yours? Researchers take steps to characterize and quantify voice quality

December 2 2016

What are the characteristics of the way you say, "hello," (or anything else for that matter) that makes you recognizable over the phone? Despite the increasing amount of literature on personal voice quality, very little is actually known about how to characterize the sound of an individual speaker.

Two researchers from UCLA in Los Angeles, California, Patricia Keating and Jody Kreiman, are joining forces (as they have done many times in the past) to apply acoustics tools to their linguistics research, investigating this question. Keating and Kreiman will present preliminary findings of their research at the 172nd Meeting of the Acoustical Society of America and the 5th Joint Meeting with Acoustical Society of Japan, held Nov. 28-Dec. 2, 2016, in Honolulu, Hawaii.

Essentially, Keating and Kreimen want to find out how to measure what people sound like. "There's no way to quantify what that means," Kreiman said. "When you change something physical, can you predict what that will sound like?"

An individual person's voice may vary over time because of their emotional state, health, the context of the conversation, or a host of other factors that make quantifying this measurement particularly difficult.

A large body of evidence from phonetics, cognitive psychology and neuropsychology indicates that listeners organize all this intra-talker variability into a prototype for each talker—an "average" representation—and a set of deviations from that prototype. Even a single syllable can carry enough information to distinguish one voice from another, but it's not yet clear what specifically are the most important identifying characteristics within such a prototype, or how much each characteristic must vary before the voice becomes unrecognizable.

"Voice quality is going to wander," Keating said. "We are looking at the point when you stop sounding like yourself and start sounding like someone else."

Keating and Kreiman digitally analyzed recordings from fifty women, all native speakers of English, who read five sentences twice on three different days. This analysis looked at multiple acoustic parameters for the vowel and consonant sounds making up the read sentences, such as fundamental frequency, intensities of harmonic frequencies relative to one another, and how they compare to the underlying noise levels within the voice.

These sentences provided each characteristic with a quantitative average and range, the collection of which formed a potential identifying voice

profile of sorts. By comparing all of the speakers to this set of characteristics—a particular person's voice profile—using a random set of their sample sentences, it could be tested for accuracy in distinguishing the correct speaker and compared to how well other sets of characteristics act to distinguish a particular voice.

This work expands on previous work the two have successfully completed with a sample of just three speakers. The larger sample size offers more insight to understanding which characteristics, and by what margin, make a recognizable voice unrecognizable. This is why the set of samples was comprised of similar speakers, all female and native English speakers.

"Who should be confusable and under what circumstances?" Kreiman asked. "How much of an acoustical change is perceptible?" Looking ahead, answering these questions may help in generating predictions about confusability in the context of both human listeners, who tend to be able to discern recognizably in a matter of seconds, and computer algorithms, that typically require samples closer to a minute in length.

  More information: Poster 5aSC7, "Acoustic similarities among female voices," by Patricia Keating is at 8:00 a.m.-12:00 p.m. HAST, Dec. 2, 2016 in Room Coral 3.

Provided by Acoustical Society of America