

Machines have nothing on mum when it comes to listening

October 8 2015



Credit: University of Montreal

More than 99% of the time, two words are enough for people with normal hearing to distinguish the voice of a close friend or relative amongst other voices, says the University of Montreal's Julien Plante-

Hébert. His study, presented at the 18th International Congress of Phonetic Sciences, involved playing recordings to Canadian French speakers, who were asked to recognize on multiple trials which of the ten male voices they heard was familiar to them. "Merci beaucoup" turned out to be all they needed to hear.

Plante-Hébert is a [voice](#) recognition doctoral student at the university's Department of Linguistics and Translation. "The auditory capacities of humans are exceptional in terms of identifying familiar voices. At birth, babies can already recognize the voice of their mothers and distinguish the sounds of foreign languages," Plante-Hébert said. To evaluate these auditory capacities, he created a series of voice "lineups," a technique inspired by the well-known visual identification procedure used by police, in which a group of individuals sharing similar physical traits are placed before a witness. "A voice lineup is an analogous procedure in which several voices with similar acoustic aspects are presented. In my study, each voice lineup contained different lengths of utterances varying from one to eighteen syllables. Familiarity between the target voice and the identifier was defined by the degree of contact between the interlocutors." Forty-four people aged 18-65 participated.

Plante-Hébert found that the participants were unable to identify short utterances regardless of their familiarity with the person speaking. However, with utterances of four or more syllables, such as "merci beaucoup," the success rate was nearly total for very familiar voices. "Identification rates exceed those currently obtained with automatic systems," he said. Indeed, in his opinion, the best speech recognition systems are much less efficient than auditory system at best, there is a 92% success rate compared to over 99.9% for humans. Moreover, in a noisy environment, humans can exceed machine-based recognition because of our brain's ability to filter out ambient noise. "Automatic speaker recognition is in fact the least accurate biometric factor compared to fingerprints or face or iris recognition," Plante-Hébert said.

"While advanced technologies are able to capture a large amount of speech information, only humans so far are able to recognize familiar voices with almost total accuracy," he concluded.

Provided by University of Montreal

Citation: Machines have nothing on mum when it comes to listening (2015, October 8) retrieved 26 April 2024 from <https://medicalxpress.com/news/2015-10-machines-mum.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.