

Self-learning computer software can detect and diagnose errors in pronunciation

July 20 2016

An inexpensive, versatile and personalized system for recognizing and correcting mispronounced words could improve language learning. The A*STAR-devised system gradually picks up the most common speech mistakes made by an individual, and potentially could be applied to any language.

"The majority of research in this field focuses on one language, or one type of native-language speaker," explains Nancy Chen at the A*STAR Institute for Infocomm Research, who led the effort along with Ann Lee at the Massachusetts Institute of Technology. "We wanted our system to be more general."

Computers typically 'learn' by recognizing patterns hidden in large data sets, such as the tendency of native Mandarin Chinese speakers to express 'v' sounds in English as 'b'. Most current speech recognition software learns these rules from training data—compiled recordings of language beginners that have been marked by a linguistics expert for phonetic mistakes. "But the process of having humans transcribe how sounds are mispronounced is time-consuming and labor-intensive, and doesn't scale well from language to [language](#)," says Chen. Instead the researchers developed an unsupervised learning system that could train itself.

Lee had previously created a rudimentary model that groups phonemes into distinct acoustic units—'a's, 'e's, and 'i's—by measuring the differences between the speech sounds. The model then sifts mistakes

and stores them as mispronunciation patterns to seek out.

To improve the model's ability to recognize mispronounced phrases, the A*STAR–MIT team introduced two techniques. First, instead of storing every possible error, the system only considers the most likely errors when assessing sound bites. "Unsupervised learning is a noisy process, so it helps to only consider estimated guesses that you are more confident with," explains Chen.

The second technique involves checking errors not just against a standard native speaker's voice, but also against the learner's own voice. By accounting for the learner's unique vocal characteristics, the system avoids detecting errors where they do not exist. "Smartphone apps can collect a lot of data specific to a user, which allows us to build a compact speech recognizer tailored to an individual," Chen elaborates.

The researchers tested their upgraded system on native English learners of Mandarin, and found that it halved the number of unlikely errors identified by the earlier model and reduced the number of undetected errors to levels comparable with a trained learning system.

Chen's team is currently advancing supervised and unsupervised learning techniques to also assess melody in speech, which affects the meaning of words in tonal languages like Mandarin.

More information: Lee, A., Chen, N. F. & Glass, J. Personalized mispronunciation detection and diagnosis based on unsupervised error pattern discovery 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 6145–6149 (2016). [DOI: 10.1109/ICASSP.2016.7472858](https://doi.org/10.1109/ICASSP.2016.7472858)

Provided by Agency for Science, Technology and Research (A*STAR),
Singapore

Citation: Self-learning computer software can detect and diagnose errors in pronunciation (2016, July 20) retrieved 11 September 2024 from <https://phys.org/news/2016-07-self-learning-software-errors-pronunciation.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.