

One class in all languages

June 14 2019

Advances in communication technology have had a major impact in all sorts of industries, but perhaps none bigger than in education. Now, anyone from around the world can listen live to a Nobel Prize laureate lecture or earn credits from the most reputable universities via internet access. However, the possible information to be gained from watching and listening online is lost if the audience cannot understand the language of the lecturer. To solve this problem, scientists at the Nara Institute of Science and Technology (NAIST), Japan, presented a solution with new machine learning at the 240th meeting of the Special Interest Group of Natural Language Processing, Information Processing Society of Japan (IPSJ SIG-NL).

Machine [translation](#) systems have made it remarkably simple to ask for directions in a foreign country. Sometimes, the systems can make amusing and innocent errors, but overall, they achieve coherent communication, at least for short exchanges only a sentence or two long. In the case of a presentation that can extend past an hour, such as an academic lecture, they are far less robust.

"NAIST has 20 percent [foreign students](#) and, while the number of English classes is expanding, the options these students have are limited by their Japanese ability," explains NAIST Professor Satoshi Nakamura, who led the study.

Nakamura's research group acquired 46.5 hours of archived lecture videos from NAIST with their transcriptions and English translations, and developed a [deep learning](#)-based system to transcribe Japanese

lecture speech and subsequently translate it into English. While watching the videos, users see subtitles in Japanese and English that match the lecturer's speech.

One might expect the ideal output would be simultaneous translations that could be done with live presentations. However, live translations limit the processing time and thus the accuracy. "Because we are putting videos with subtitles in the archives, we found better translations by creating subtitles with a longer [processing time](#)," he says.

The archived footage used for the evaluation consisted of lectures from robotics, speech processing and software engineering. Interestingly, the word error rate in speech recognition correlated to disfluency in the lecturers' [speech](#). Another factor from the different error rates was the length of time speaking without pause. The corpus used for the training was still insufficient and should be developed more for further improvements.

"Japan wants to increase its international students and NAIST has a great opportunity to be a leader in this endeavor. Our project will not only improve machine translation, it will also bring bright minds to the country," he continued.

Provided by Nara Institute of Science and Technology

Citation: One class in all languages (2019, June 14) retrieved 27 June 2024 from <https://phys.org/news/2019-06-class-languages.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.