AI SHAPING THE EVOLUTION OF MULTILINGUAL SPEECH TRANSLATION TECHNOLOGY

VoiceTra is a speech translation app that uses artificial intelligence to allow speakers of thirty-one languages to communicate instantly. VoiceTra's fast and accurate translation is expected to solve the communication problems for travelers from abroad.

TAMAKI KAWASAKI

VOICETRA is a multilingual speech translation app capable of instant translation between thirty-one languages, developed by the National Institute of Information and Communications Technology (NICT). Speech translation systems are being developed elsewhere in the world, with several well-known web services already in existence. However, one of the biggest advantages VoiceTra has is that it does not use English as an intermediate step in the translation process. Instead, the app translates directly from Japanese into thirty other languages, allowing for a much higher level of accuracy.

Taking into account the wide range of tourists in Japan, it supports major European languages and Asian languages as well, which is also one of its unique selling points.

"VoiceTra is currently the only app that supports two-way speech translation between Japanese and Myanmarese," states Kiyotaka Uchimoto, director of the Planning Office at the Advanced Speech Translation Research and Development Promotion Center of NICT. The number of visitors to Japan is expected to keep increasing as the 2020 Tokyo Olympic and Paralympic Games approach. To create a society where visitors would not have to worry about language barriers, the Ministry of Internal Affairs and Communications announced its Global Communication Plan in 2014. Since then, many private enterprises have joined this nationwide initiative, speeding up NICT's research and development to the point where they were able to carry out field experiments of their multilingual speech translation technology in real-world settings, such as hospitals, shops, restaurants and tourist destinations.
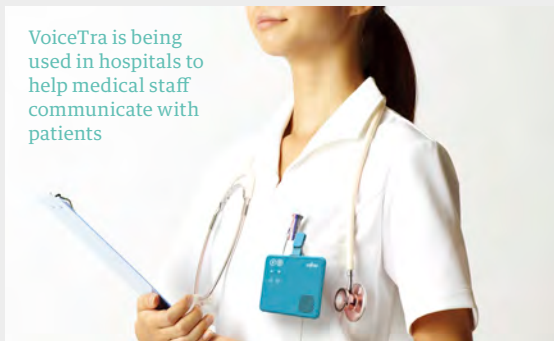
VoiceTra is available for free to individual users. It offers a simple and easy-to-use interface, and accurately recognizes the differences between various speakers such as their inflections and speaking styles. Moreover, the translation process is nearly instantaneous. The system utilizes a neural network (deep learning) based on human neural pathways to accelerate its learning process. The accuracy increases as the number of users rises with more information at the disposal of machine learning algorithms, leading to a natural and nonliteral translation. Emergency services, railway operators, taxi companies and hospitals have customized the VoiceTra technology and are using it to communicate with visitors from abroad. The smartphone app has also marked a total of three million downloads.

This wonder of technology is the culmination of nearly three decades of study, despite a rocky start due to skepticism regarding the merits of such research. "Nevertheless, we were convinced that this speech translation technology was necessary for the future of Japan," says Uchimoto, thinking back over the road that led him to this point. NICT's speech recognition technology has held up well against rival technologies from prestigious universities and noted research labs around the world at the International Workshop on Spoken Language Translation (IWSLT), an international competition, placing first in the world for three years in a row.

Besides being used in train stations and other public transport, medical facilities and in other situations in Japan, VoiceTra is now going beyond just a smartphone app in the hands of private enterprises. For instance, a hands-free ID card-sized device developed by one company targets doctors and nurses in hospitals. The device translates voices detected from the front (i.e., the patient's voice) into Japanese, and voices detected from the top of the device (i.e., the doctor or nurse's voice) into a designated language. Further research is being done at NICT on technologies that detect and translate multiple languages simultaneously. "The ultimate goal of our research is simultaneous interpretation that gives the most appropriate translation according to the context. To help us build the necessary database, we hope that as many people as possible will try out VoiceTra when they visit Japan," says Uchimoto. ⨍

VoiceTra is being used in hospitals to help medical staff communicate with patients

VoiceTra official site:
http://voicetra.nict.go.jp/en/index.html