

# 19

## Listening languages around the world simultaneously

- Multilingual speech identification & recognition technology -

### Abstract

We are developing the technology to **listen languages around the world simultaneously**, composed of 3 components;

1. **Intelligent Microphone:**

Separate speech spoken simultaneously, being able to **separate up to 7 speakers**.

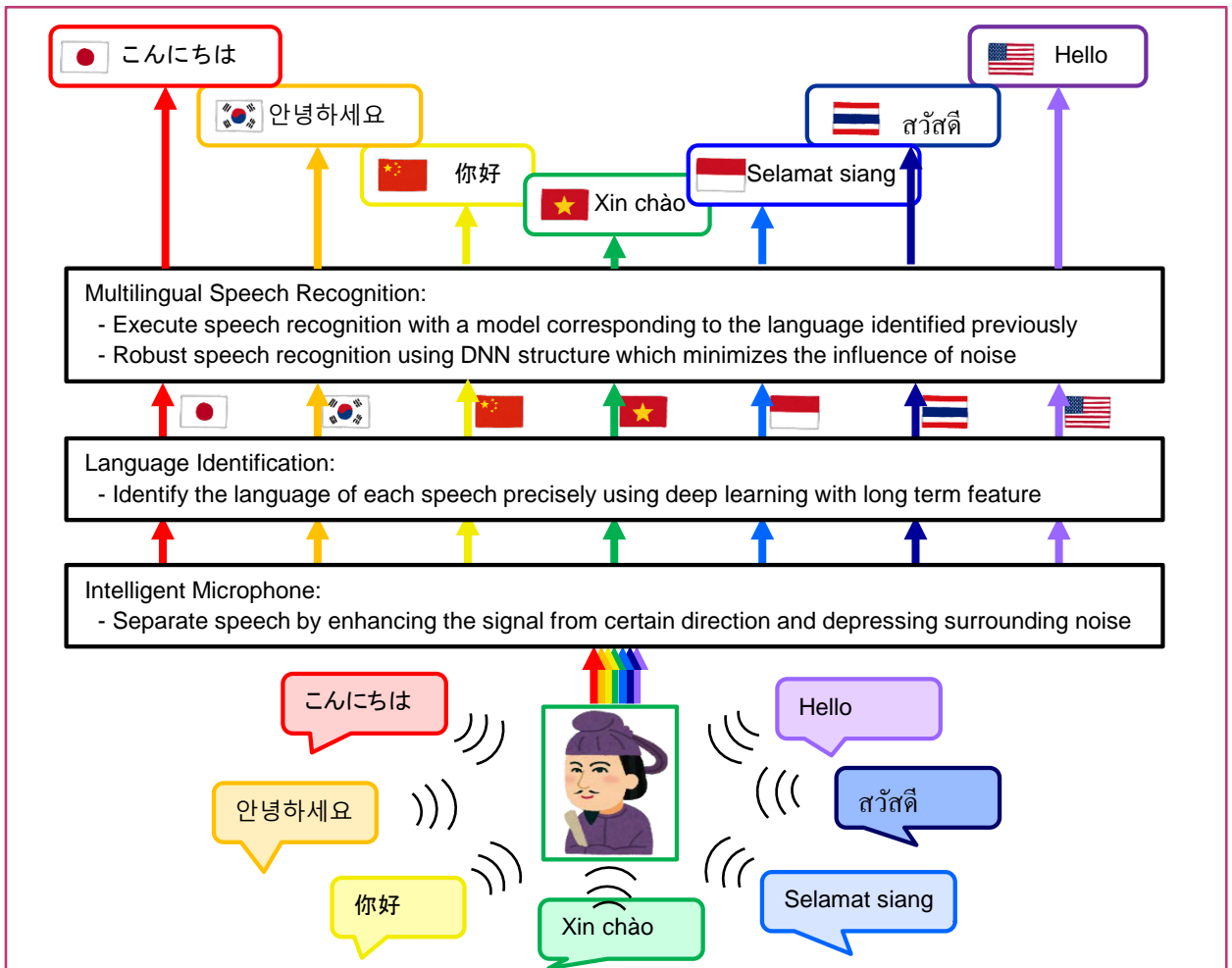
2. **Language Identification:**

Identify the language of speech precisely using LSTM framework

3. **Multilingual Speech Recognition:**

Support **10 languages in Asia-Pacific region**, using DNN structure robust for noisy environment.

We are planning to **increase the supporting language** and applying it to **natural speech interface** for foreign visitors.



### Reference

- [1] T. Yoshioka, N. Ito, M. Delcroix, A. Ogawa, K. Kinoshita, M. Fujimoto, C. Yu, W. J. Fabian, M. Espi, T. Higuchi, S. Araki, T. Nakatani, "The NTT CHiME-3 system: Advances in speech enhancement and recognition for mobile multi-microphone devices," in *Proc. ASRU2015*, pp. 436-443, 2015.
- [2] R. Masumura, T. Asami, H. Masataki, Y. Aono, "Parallel Phonetically Aware DNNs and LSTM-RNNs for Frame-by-Frame Discriminative Modeling of Spoken Identification," in *Proc. ICASSP2017*, 2017.

### Contact

**Hirokazu Masataki** Media Intelligence Laboratories  
Email [masataki.hirokazu\(at\)lab.ntt.co.jp](mailto:masataki.hirokazu(at)lab.ntt.co.jp)