28 I can speak like a native
～Speaking rhythm conversion rules by English speech corpus～

Abstract

Because of speaking rhythm differences between Japanese and English, most native Japanese speakers have difficulty communicating well in English. We have developed a technique of automatically correcting the halting English rhythm of native Japanese speakers by approximating the natural rhythm of native English speakers. However, our previous speaking rhythm conversion technique needed the same sentences to be spoken by a native speaker. In this study, we devised rules for converting the speaking rhythm of native Japanese speakers into that of native English speakers using English speech corpus. We hope that this technique will eventually alleviate the burden involved in communication using non-native languages.

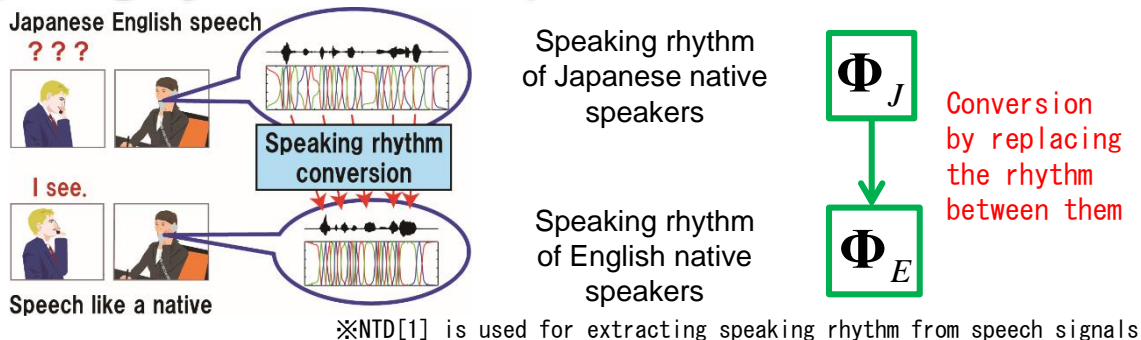## Speaking rhythm differences between Japanese and English

Stress-timed rhythm (English)   Even intervals among stressed syllables

Humpty dumpty sat on a wall.   (Stress in red)

Mora-timed rhythm (Japanese) Equal duration of every mora

ハンプティ ダンプティ サット オン ア ウォール

Speaking rhythm conversion 'Speak like a native'

Speaking rhythm of Japanese native speakers

$$\Phi_J$$

Conversion by replacing the rhythm between them

Speaking rhythm of English native speakers

$$\Phi_E$$

※NTD[1] is used for extracting speaking rhythm from speech signals

Previous method   Required a sample of the same sentence read by a native

Proposed method

Speaking rhythm conversion without a sample of the same sentence

Machine learning technique creates rules for converting the speaking rhythm of Japanese speakers into that of English speakers.

$$\Phi_E \approx \boxed{F}\Phi_J$$

Conversion rules

1. Modeling of Japanese native speakers' English (KATAKANA EIGO)
2. Phoneme-based conversion rules

Related works

[1] S. Hiroya, "Non-negative temporal decomposition of speech parameters by multiplicative update rules," *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 10, pp. 2108-2117, 2013.
[2] S. Hiroya, "Speaking rhythm extraction and control by non-negative temporal decomposition," *NTT Technical Review*, Vol. 11, No. 12, 2013.

Contact

Sadao Hiroya   Sensory and Motor Research Group, Human Information Science Laboratory
E-mail : hiroya.sadao(at)lab.ntt.co.jp