

# 英語音声ファイルから発音が近い和文の自動生成 - 空耳自動生成システム -

榎間 祐太 堀 幸一郎 水谷 雄樹 濱川 礼  
中京大学 情報科学部 情報科学科

## 1. はじめに

本論文では、英語音声ファイルから日本語の空耳を自動的に生成するシステムについて述べる。空耳とは、「外国語の音を何らかの意味を持った日本語に聞こえるように工夫する言葉遊びの一種」である。

現在テレビ番組やネット上で空耳事例が多くあるが、面白いものはそれほど多くはない。それは、視聴者から投稿された空耳を紹介し評価するテレビ番組「タモリ倶楽部」で過去 17 年間での最高評価が全体の 2.5% しかない例から鑑みることができる[1]。そこで、我々はユーザが面白い空耳を思い付くことを支援することを目的とし、空耳自動生成システムの開発を行った。

## 2. 本システムの特徴

本システムは、英語の音声ファイルを入力することで、日本語の空耳をテキストで出力する。英語を使用するのは、日本の義務教育で学習され、日本人が日常で触れる機会が多いからである。

空耳には「音の位置がずれる」、「変形していく単語」、「音階の錯覚」、「雑音の七変化」、「背後や頭上を通る音」などがある[2]。我々はその中で「音の位置がずれる」、「音階の錯覚」を元に「発音が小さく聞き取りづらい部分を考慮しない」という特徴が、「変形していく単語」を元に「文字の区切りを都合よく解釈する」という特徴があると考えた。本システムではこの2つの特徴を考慮して、日本人が思い付く空耳に近い空耳を生成するために、それぞれ「音声編集機能」と「漢字変換機能」を実装した。また、生成された空耳の信頼性を検証するための機能として「音声比較機能」を実装した。

## 3. システムの全体構成

本システムは、「音声編集・認識部」、「漢字変換部」、「音声比較部」の3つの機能から構成されている[図 1]。

Automatic generation of Japanese which pronunciation is near from English voice file  
-Mishearing automatic operation generation system-  
Yuta Sakakima, Koichiro Hori, Yuki Mizutani and Rei Hamakawa  
Chukyo University Department of information science

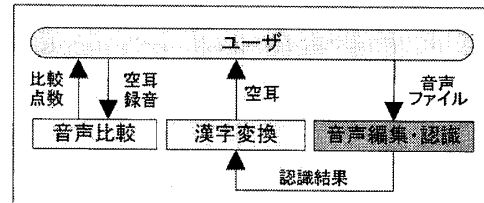


図 1 システム構成図

## 4. 音声編集・認識部

2 章でも述べた「発音が小さく聞き取りづらい部分を考慮しない」という特徴を表現するために、小さな発音が一定区間以上連続する箇所を音声データからあらかじめ削っておく。そして、小さな発音を削った音声データを、日本語音声認識エンジンによって音声認識を行うことで空耳を生成している。

また、音声を削る度合い(カット率)を変化させることで認識結果が変化する。どの認識結果が最も良いかは聞き手によって変化するため、英語音声の認識で良い出力結果が得られることが多かった3種類のカット率を設定し、カット A, B, C とした。この3つの結果を空耳として出力する。

[表 1]にカット率の具体的な数値を示す。音声波形の中で最大の振幅を max とし、max の a% 以下の値が b/16000s 続いた区間の振幅を 0 とし、音声をカットする。

表 1 音声のカット率ごとの具体的な数値

カット率	a (%)	b (1/16000s)
カット A	2	100
カット B	6	200
カット C	10	400

[図 2]に音声編集を行った際の認識結果を示す。

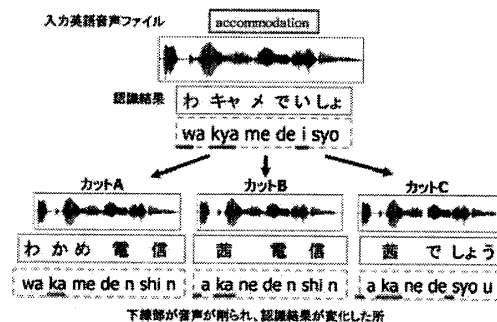


図 2 音声編集による認識結果の変化

