

人を楽しませる音声の合成に向けて

# 伝統話芸 落語音声の合成

加藤集平, 高木信二, 山岸順一, Xin WANG

## どんな研究？

人間は声を使って聞き手を楽しませることが出来ます。本研究では、伝統話芸である落語を音声合成で実現し、人を楽しませることを目指しています。

## 何がわかる？

単なる情報伝達を超えた音声合成を行うことで、話芸のメカニズムを解明し、機械と人間との新たなコミュニケーションを実現できると期待されます。

## 状況設定

音声合成の品質はここ数年で飛躍的に発展し、今では人間の声と区別が付かないほどになっています。しかし、これは文をただ読み上げる場合の話であり、感情などの多様な表現、ひいてはそれらを駆使して人を楽しませることはまだできません。本研究では、人を楽しませる音声合成の実現に向けて、代表的な話芸の一つである落語の音声合成を目指します。

## 研究内容

### 落語音声データベースの構築

収録環境: 録音スタジオ

演者 : 柳家三三 (落語協会、真打)

演目 : 古典落語25演目 (およそ13時間)



収録の様子

### 落語のしゃべり方を再現する音声合成モデルの構築

構成 : DNNの音響モデル + WaveNetボコーダー

入力 : quinphone (音の種類) + 追加のラベル

追加のラベル: 発話者の属性 (役など)、状態 (驚き、困惑など) ほか

### 今後の課題

音素長の推定、出来上がった音声の評価方法など

### 使用想定例

スマートスピーカーへの組み込み？



組み込み



出典: <https://clova.line.me>

🎧 開発途中ですが、サンプル音声があります。



連絡先: 加藤集平 / 国立情報学研究所 コンテンツ科学研究系

TEL: 03-4212-2576

FAX: 03-4212-2150

Email: [skato@nii.ac.jp](mailto:skato@nii.ac.jp)