

物真似音声の分析

甲南大学理工学部情報システム工学科

北村達也

t-kitamu@konan-u.ac.jp

概要

- プロの物真似タレントの音声进行分析.
- 分析対象(都々逸)
 - 対象話者の音声
 - 物真似タレントによる物真似音声と地声音声
- 結果
 - 音節継続時間長: 似せていない
 - 平均基本周波数: 高く誇張
 - 基本周波数の変化パターン: 似せている
 - スペクトル: 似せている
 - 声帯音源特性(H1-H2): 似せている

研究の動機と目的

- 物真似タレントは
 - 音声の中のどの音響特徴量を近づけているのか？
 - 音声生成系の制御はどのようなものか？
- 物真似を似ていると知覚する要因は何か？
→ 個人性制御や声質制御に関する有益な知見

先行研究

- 鈴木 (1968)
 - 桜井長一郎氏が宇野重吉氏と柳家金語楼氏の物真似 (声帯模写)した音声の分析.
 - 聴感上はよく似ている音声でもそのスペクトログラムには様々な違いがある.
- Zetterholm (2001), (2002), (2006)
 - 基本周波数, 継続時間長, ホルマント周波数の分析, 音素ラベリング, 方言を真似した音声の分析, 1人の対象話者を2人の物真似タレントが真似した音声の分析
- Clermont & Zetterholm (2006)
 - ホルマント周波数の変化パターン

音声データ

- 話者: 対象話者U, 物真似タレントC
- 話者C: 物真似発話と地声発話
- 文章: 「一度でいいから見てみたい, 女房(にようぼ)がへソクリ隠すところ」
- 収録環境: 比較的静かな部屋(残響, 暗騒音あり)
- 標本化周波数16 kHz, 量子化16 bit.

話者U



話者C物真似



話者C地声



分析方法

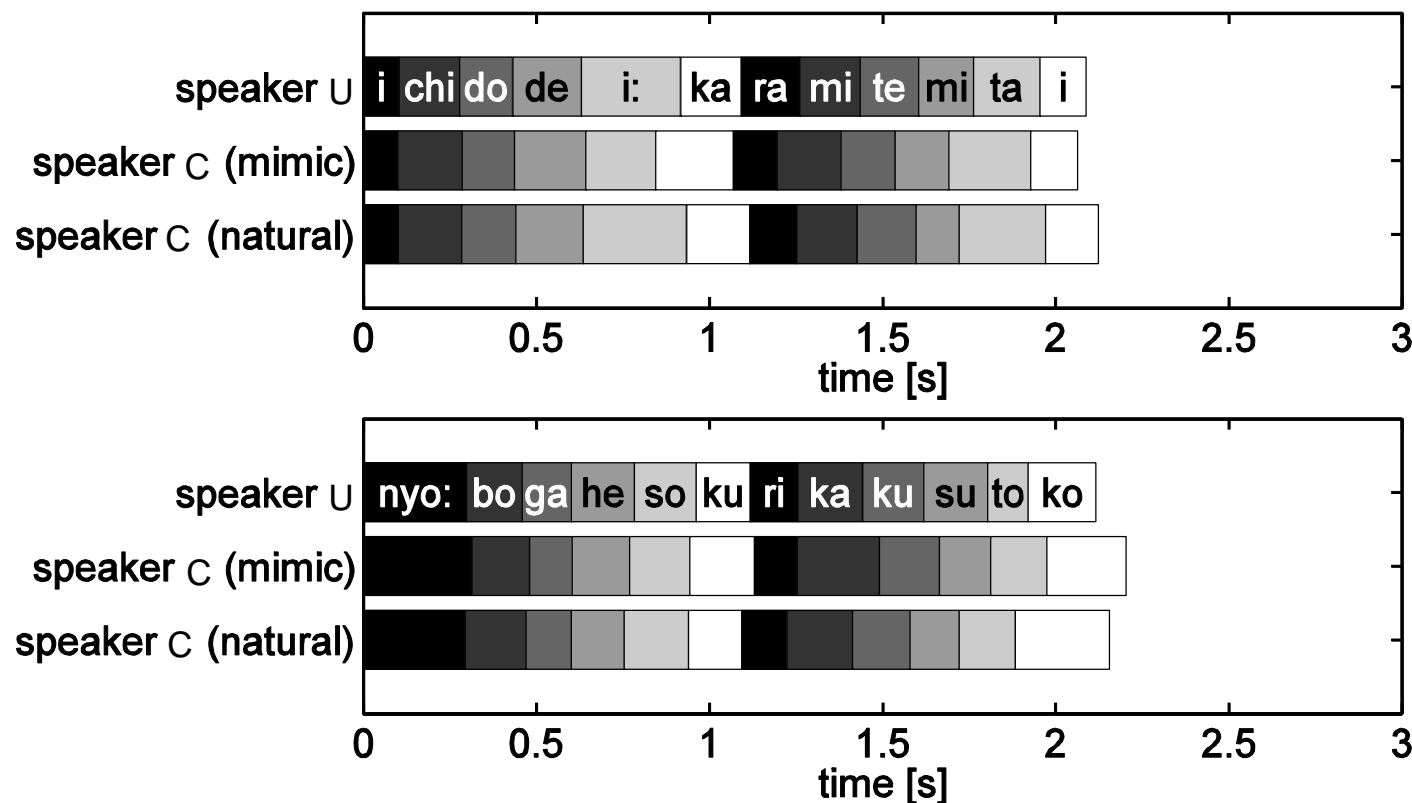
- 音節継続時間長
 - 目視でラベリング
- 基本周波数
 - Wavesurferの基本周波数抽出機能 (Talkin, 1995)
 - 目視で修正
- DFTスペクトル
 - 概形
 - ホルマント周波数
 - H1-H2
- 除外: パワー(残響のため)

結果(1): 音節継続時間長

- 相関係数

- 対象話者の音声-物真似音声: 0.758
- 対象話者の音声-地声音声: 0.847

音節継続時間長を近づけているとはいえない(誇張している可能性はある)。

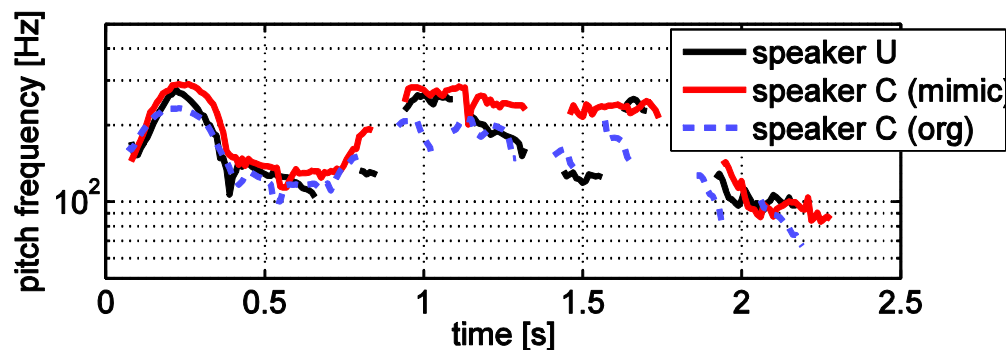
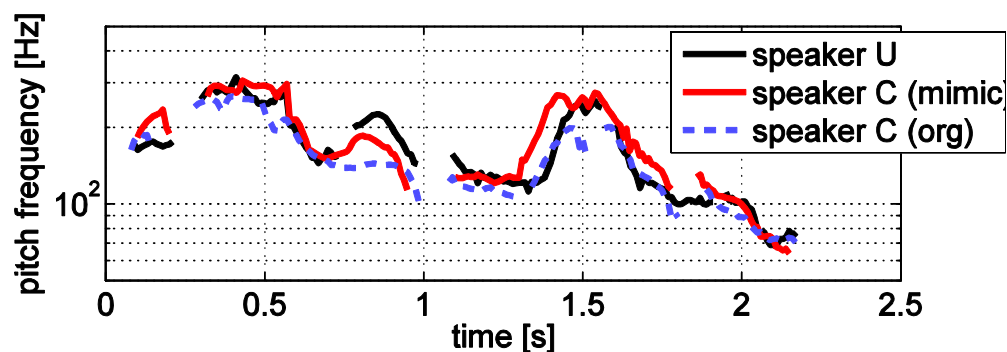


音節継続時間長. (上)前半部, (下)後半部.

結果 (2) : 基本周波数

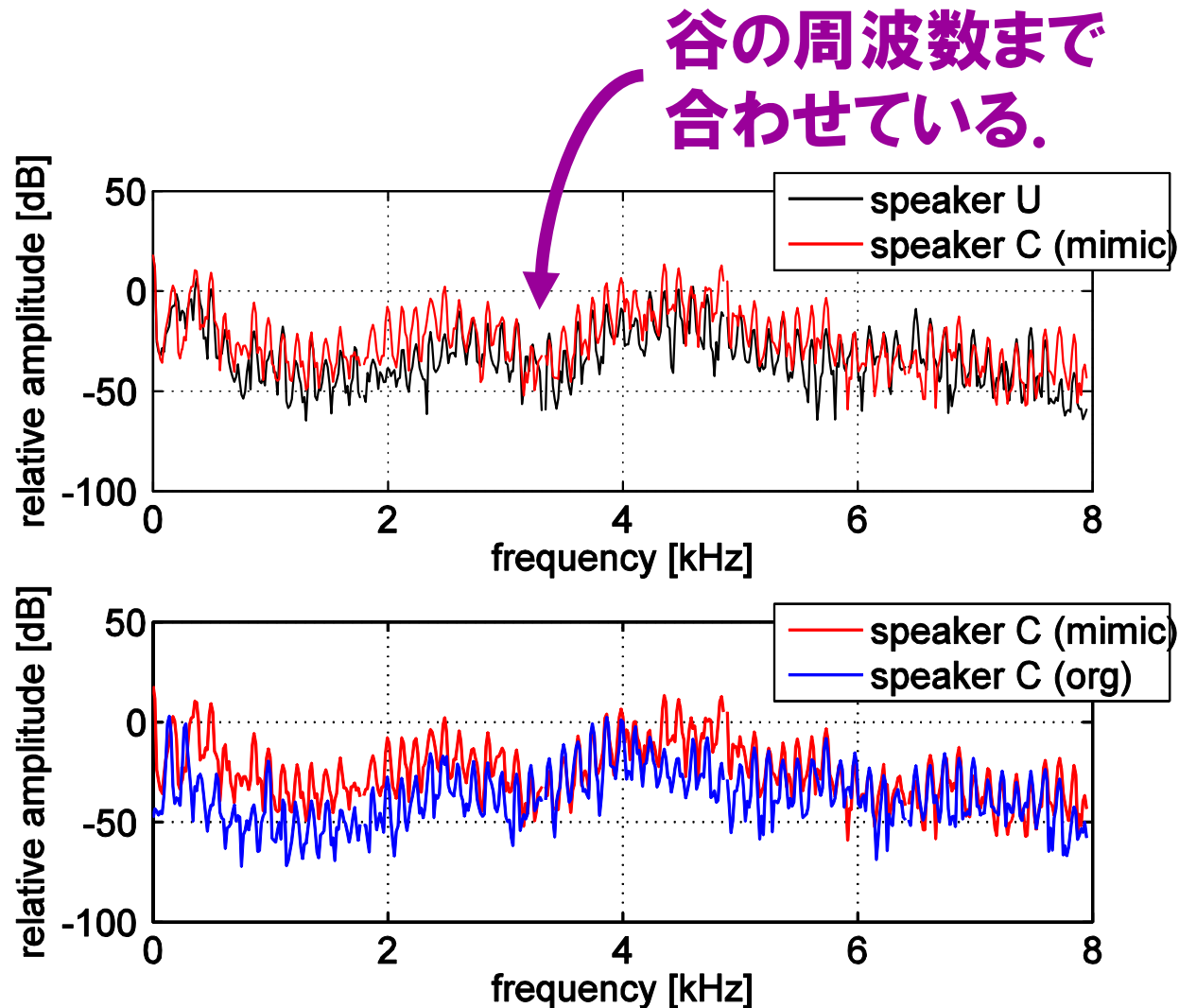
- 平均値
 - 話者U: 167.2 Hz
 - 話者C物真似: 185.1 Hz
 - 話者C地声: 152.0 Hz
- 変化パターン: (主観的には)よく似ている.

誇張しているのではないか.



基本周波数. (上)前半部, (下)後半部.

結果 (3) : DFTスペクトル



F1付近(約1 kHz以下)と約2.6 kHz以上の帯域が、主観的にはよく似ている。

前半部「いい」のDFTスペクトル。

結果 (4) : H1-H2

- DFTスペクトルの第1, 第2調波の振幅差
- 声帯音源特性の指標 (Hanson et al., 2001)
- 「いい」
 - 話者U: -8.03 dB
 - 話者C物真似: -8.12 dB
 - 話者C地声: 3.82 dB
- 「あ」
 - 話者U: -2.22 dB
 - 話者C物真似: -11.55 dB
 - 話者C地声: 18.3 dB

} 極めて近い

} 符号は一致

**声帯音源特性も
制御している。**

考察

- 平均基本周波数を誇張。
 - Zetterholm (2001)でも同様の傾向。
- 基本周波数の変化パターンを似せる。
 - 変化パターンにも音声の個人性が存在 (Akagi & Ienaga, 1997)。
- スペクトル形状を似せ, ホルマント周波数を近付ける。
 - 声道形状の制御
- 声帯音源特性を近付ける(甲高いしわがれ声)
 - 声帯の制御
- 音節継続時間長, F2は近くない。
 - 知覚上の寄与が小さい? 変えたくても変えられない?

まとめ

- プロの物真似タレントの音声进行分析.
- 結果
 - 音節継続時間長: 似せていない
 - 平均基本周波数: 高く誇張
 - 基本周波数の変化パターン: 似せている
 - スペクトル: 似せている
 - 声帯音源特性(H1-H2): 似せている
- 基本周波数, 声帯音源特性, 声道音響特性を制御している.

物真似 ≠ 「声帯」模写