

P21 母音の鼻音化が声紋認証に与える影響について

野間勇雅, 吉村博幸(千葉大学)

1. 研究背景

声紋認証とは音声によって個人識別をするバイOMETRICSの一種[1,2]

声紋認証では主に母音が対象[3]
しかし実際には日本語は母音を単独で発声しない場合もある

子音が母音に与える影響を明らかにしより自然な発声状態で認証可能にする

2. 研究目的

鼻音化が声紋認証にどのような影響があるのかを解明することが目的

鼻音に影響を受ける母音の鼻音化に着目

単独の母音と鼻音化した母音をケプストラム分析

/a/と/ma/の/a/のEER(等誤り率)を比較

3. 解析手順

①データ入力

音声コーパス「PASIL-DSR」からA~Fの男性6名の/a/と/ma/を使用

振幅の絶対値の最大値で割り正規化

全体の録音時間から発声時のみを目視で特定

32分割した11~22項目を抽出

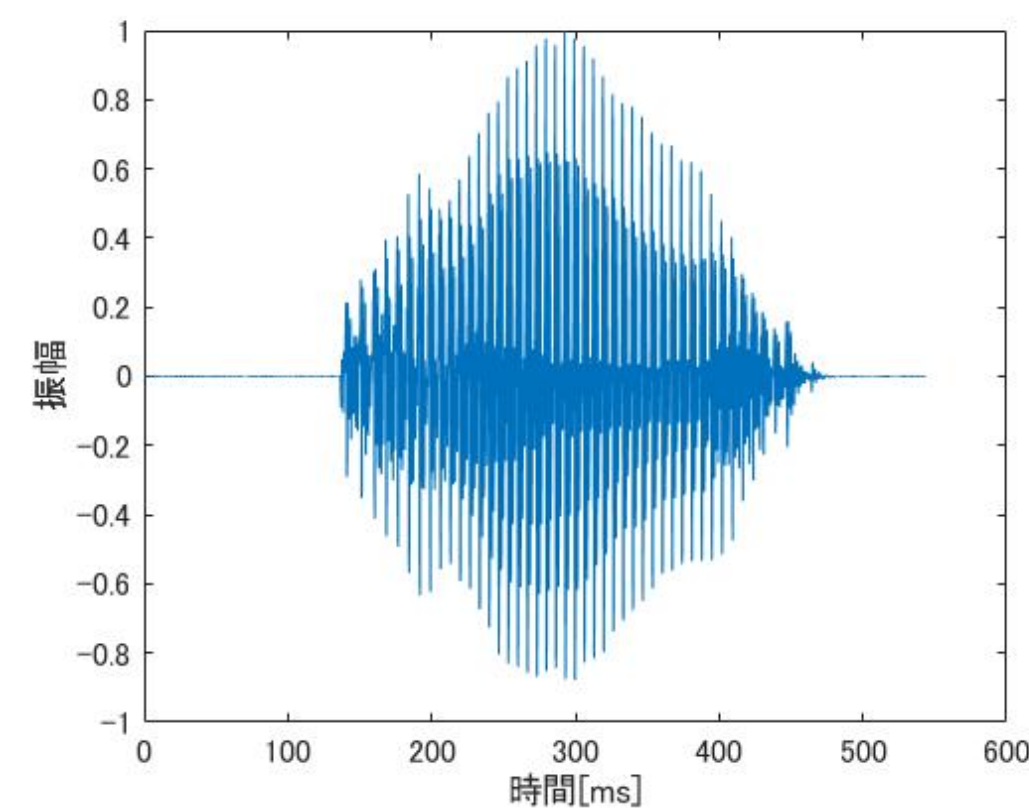


図1 /a/と発声した時の音声波形(Aの場合)

②前処理

40msの範囲をトリミング

フーリエ変換の前処理で切り出した波形にハニング窓をかける

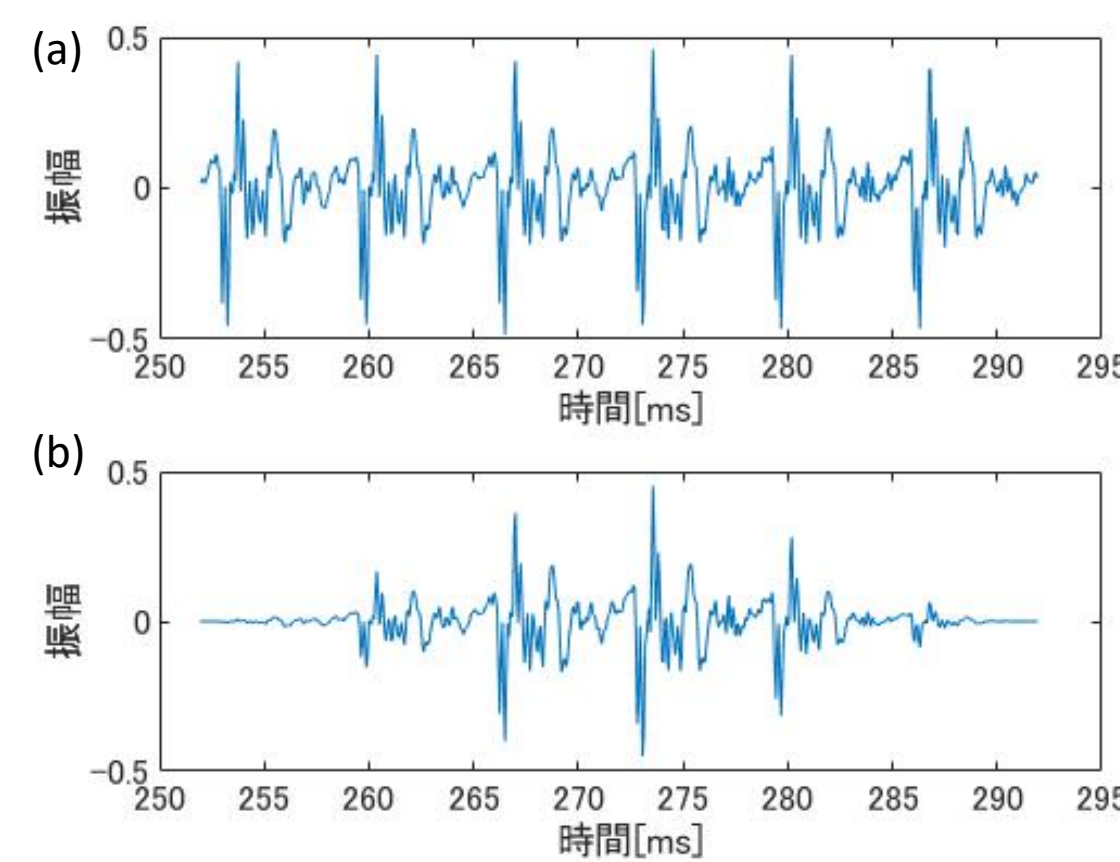


図2 音声波形(図1)に対して前処理を行った結果 (a)40msに切り出した場合 (b)窓処理を施した場合

③フーリエ変換と対数変換

サンプリング周波数2048Hzで離散フーリエ変換

対数変換で声帯情報と声道情報の成分を積から和に変換

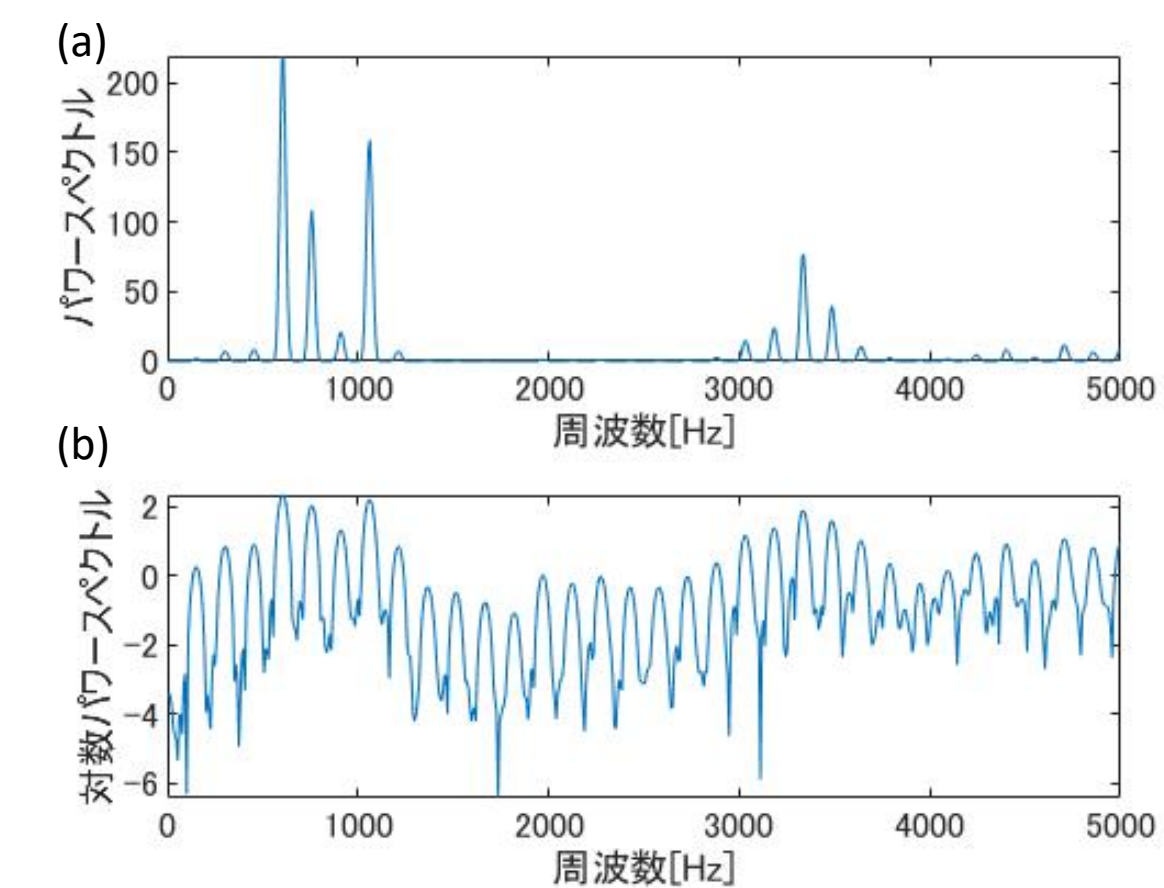


図3 図2(b)にフーリエ変換と対数変換を行った結果 (a)パワースペクトル (b)対数パワースペクトル

④リフタリング

図3(b)の対数振幅スペクトルを離散フーリエ変換

ローパスフィルタをかけて声道情報のみを抽出

離散フーリエ変換を行い対数を除き二乗

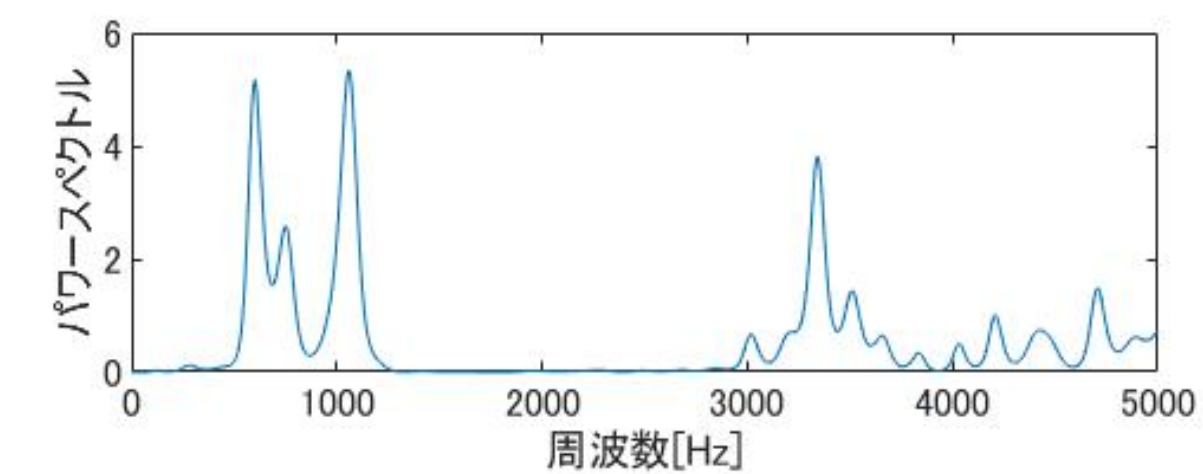


図4 図3(b)をリフタリングした後のパワースペクトル

4. 解析結果

表1から/a/のEERは既存研究[2]と同程度

表2からEERには個人差あり

表2から、A,B,D,Eの/ma/の/a/のEERは/a/のEERの約2倍
Cの場合、最小の1.11倍

Fの場合、最大の 1.07×10^3 倍。ただし、/a/のEERは6名中で最も低い。

EERの変化が最大と最小の2名(FとC)の/a/の平均パワースペクトルに差異あり(図6, 7)

表2 人物ごとのEER

	A	B	C	D	E	F
/a/	2.30×10^{-3}	4.53×10^{-2}	1.30×10^{-1}	5.76×10^{-2}	4.79×10^{-2}	1.06×10^{-4}
/ma/の/a/	5.00×10^{-3}	9.88×10^{-2}	1.44×10^{-1}	1.42×10^{-1}	7.27×10^{-2}	1.13×10^{-1}

表1 EER

/a/(図5)	4.99×10^{-2}
/ma/の/a/	1.21×10^{-1}

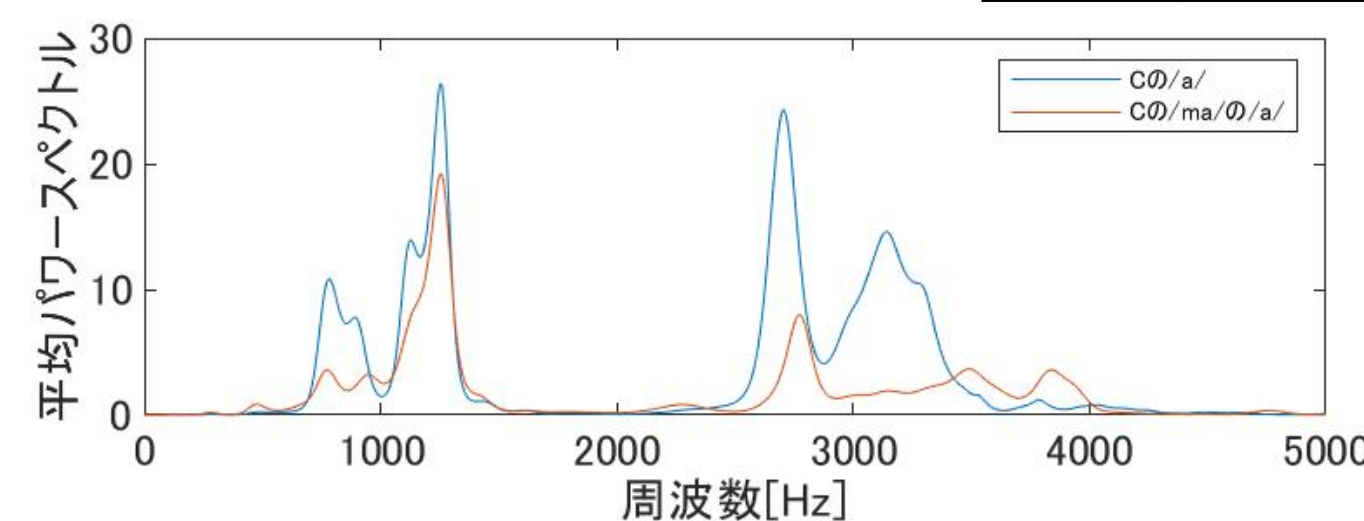


図6 Cの/a/の解析結果を平均したパワースペクトル

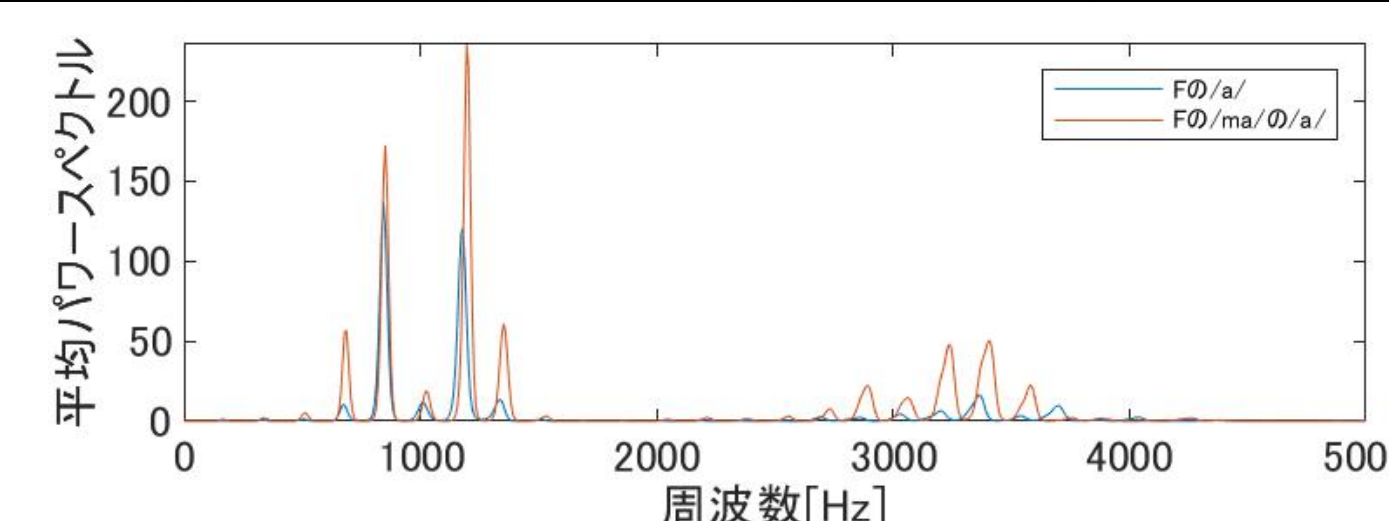


図7 Fの/a/の解析結果を平均したパワースペクトル

⑤EERの算出

図4から相互相関関数の最大値を算出しヒストグラムと正規分布をプロット

以下を満たす閾値を設定

(閾値以上の他人分布=閾値以下の本人分布)

この時の面積が評価指標であるEER

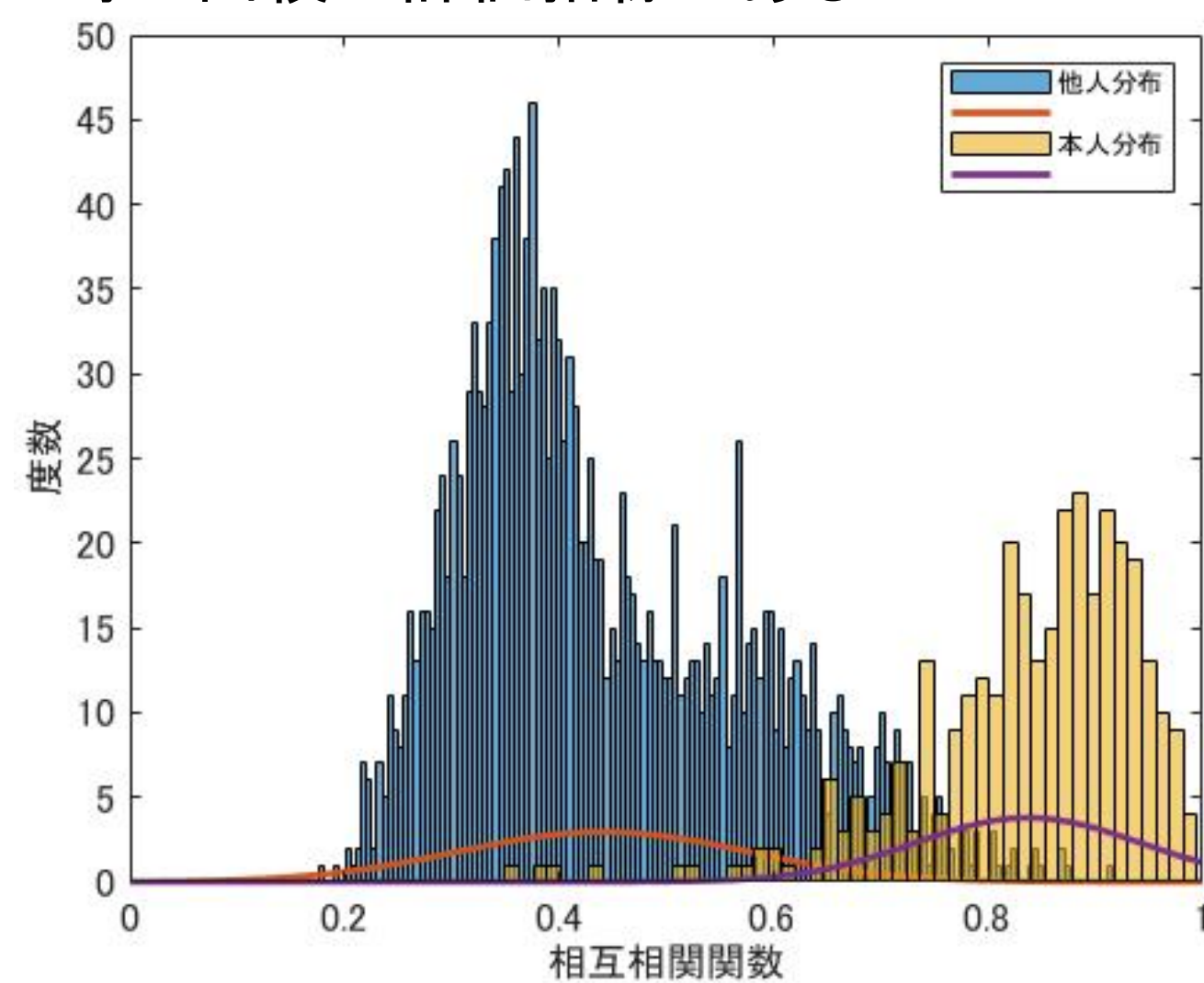


図5 本人分布と他人分布

5. 考察

Cの/a/は2500~3500Hz付近にピーク
(単独の母音が鼻音化した母音の特徴を有する)→/m/の影響が小さい

Fの/a/はピークが鋭く2500~3500Hzにはほとんどピークは存在しない
→/m/の影響を強く受けた

6. まとめ

鼻音化によってEERは上昇するが個人差あり

影響を受けにくい人物は単独の母音と鼻音化された母音が似ている

—今後の課題—

- ・データ数の増加による信頼性の向上
- ・男性だけでなく女性でも同様の解析
- ・/n/など他の鼻音を対象に解析