

# 音声の話者内変動が個人性知覚に与える影響\*

北村 達也, パーハム・モクタリ (ATR 人間情報科学研究所)

## 1 はじめに

我々は、基本周波数や音韻などの変化が音声の個人性知覚に与える影響について調査を行っている。音声の個人性は主に基本周波数やスペクトルおよびその時間変化パターンに現れる。母音のスペクトル包絡における個人性は約 1.7 kHz 以上の周波数帯域に顕著に現れると報告されている [1]。しかし、その個人性情報が母音間で共通しているのか、または母音ごとに異なるのかについては明らかになっていない。音声には声帯音源や声道音響特性などの生得的な身体特徴に起因する情報が含まれているため、母音に依存しない個人性情報が含まれていると考えられる。そこで、本研究ではこの母音間に共通する個人性情報の話者識別への寄与を明らかにするために聴取実験を行った。実験では未知話者の単母音を用い、任意の母音から抽出した個人性情報によって他の母音の話者を識別することができるか否かを調べた。母音により既知話者の識別に対する有効性が異なるという報告があるため [2][3]、母音間の差異についても考慮し、さらに音声の継続時間長および平均基本周波数の影響も調査した。

## 2 実験条件

### 2.1 刺激音

ATR 音声データベース [4] の男性話者 4 名 (MAU, MMS, MSH, MTT) による日本語 5 母音 (タスクコード SY) から作成した。これらの話者は母音の平均基本周波数と継続時間長が近いという条件で選択した。この音声データから次の 3 種類の刺激音を作成した。

刺激音 1 振幅のピーク値を正規化した音声。

刺激音 2 刺激音 1 の継続時間長を 2 倍にした音声。

刺激音 3 刺激音 2 の平均基本周波数を全音声データの平均値 (129.3 Hz) にした音声。

刺激音 1 の平均継続時間長は 234 ms である。継続時間長の違いにより話者識別に用いられる特

徴量が異なる可能性があるため、刺激音 2 では音声区間を均等に 2 倍にした。刺激音 3 では基本周波数の影響を調べるために各刺激音の平均基本周波数を一定にしたが、基本周波数の時間変化パターンは変更していない。

### 2.2 実験参加者

正常聴力を有する成人 10 名 (男性 7 名, 女性 3 名)。

### 2.3 実験方法

ABX 法により行った。A, B, X の順で刺激音を呈示し、X の話者が A と B のどちらの話者であるかを強制選択させた。A と B は同じ母音 (/a/ または /i/) を呈示し、X では 5 母音のいずれかを呈示した。本研究では A, B で呈示する母音を参照母音と呼ぶ。A の話者は MAU と MTT のみ、B および X の話者は全話者とした。以上の条件で ABX の組み合わせを作成し、ランダムに並べ替え、実験参加者に呈示した。1 つの ABX の組み合わせは実験中に 1 回呈示される。なお、順序効果を打ち消すために BAX の組についても実験を行った。全試行数は 600 である。

実験参加者はヘッドホン (Sennheiser HDA200) により刺激音を両耳受聴し、パーソナルコンピュータのディスプレイ上の A または B のボタンをマウスで選択することにより回答を行った。聴き直しは 3 回まで許した。各実験参加者は聴き易いレベルにて受聴した。

## 3 結果と考察

各刺激音に関する話者識別率 (X の話者を A または B から正しく選択した率) を図 1 に示す。これらは実験参加者間で平均した値である。いずれの刺激音でも参照母音と同じ母音の話者識別率が参照母音と異なる母音の話者識別率よりも高く、前者が 90 % 以上で後者は 60% から 75 % の範囲であった。この傾向は参照母音が /a/ と /i/ どちらの場合でも同じであった。この結果は、ある母音から抽出した個人性情報を用いてその母音の話者を識別することはできるが、他の母音の

\* Effects of intra-speaker variation of speech sounds on perception of speaker characteristics. by KITAMURA Tatsuya and MOKHTARI Parham (ATR Human Information Science Laboratories)

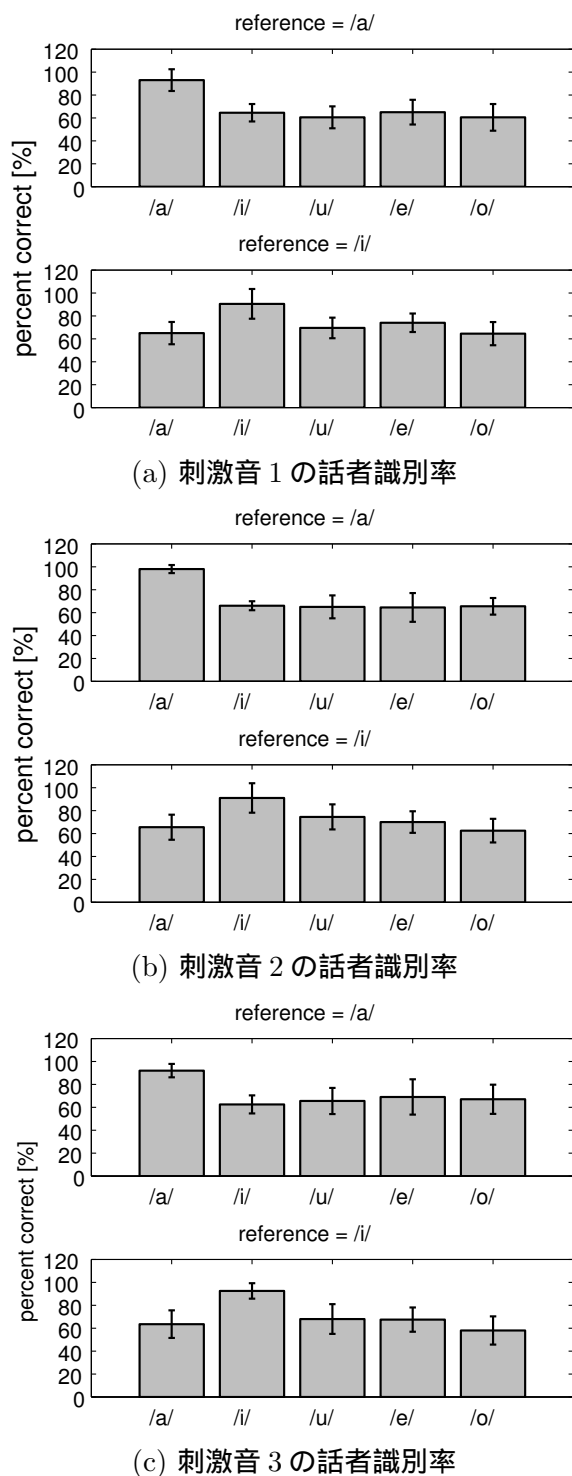


Fig. 1 各刺激音の平均話者識別率および標準偏差．各図の上段は参照母音が/a/の結果，下段は参照母音が/i/の結果．

話者を識別することは困難であることを示している．

刺激音の種類による影響を調べるために分散分析による検定を行った．参照母音と同じ母音の話者識別率に関しては ( $F(1, 38; 0.05) = 4.10$ )，刺激音 1 と刺激音 2 の間 ( $F(1, 38) = 1.09$ )，刺激

音 2 と刺激音 3 の間 ( $F(1, 38) = 0.75$ ) のいずれにも有意差がなかった．参照母音と異なる母音の話者識別率に関しても ( $F(1, 158; 0.05) = 3.90$ )，刺激音 1 と刺激音 2 の間 ( $F(1, 38) = 0.56$ )，刺激音 2 と刺激音 3 の間 ( $F(1, 38) = 0.80$ ) のいずれにも有意差がなかった．従って，全体としては継続時間長と平均基本周波数のいずれも話者の判断に影響が無かったといえる．

ただし，話者によっては平均基本周波数の影響がみられた．話者 MTT の平均基本周波数は，/a/ が 137.4 Hz，/i/ が 119.1 Hz，/u/ が 118.1 Hz，/e/ が 129.0 Hz，/o/ が 119.4 Hz であり，/a/ だけ相対的に高い．そして，話者 MTT の参照母音が/a/ の場合の平均話者識別率は，刺激音 1 では 49%，刺激音 2 では 50%，刺激音 3 では 66% であった．この結果は，刺激音 1 と 2 では，実験参加者が基本周波数を主な手がかりにして話者識別をしていたため，/a/ 以外の母音に対する話者識別率が低かったと考えられる．一方，刺激音 3 では基本周波数が話者識別の手がかりとして使えなかったことで逆にスペクトルなどその他の情報を用いられ話者識別率が向上したと考えられる．

#### 4 おわりに

本研究では，任意の母音の個人性情報を用いて他の母音の話者を識別することができるか否かを聴取実験によって調査した．今回の実験条件の下では，参照母音が/a/，/i/ どちらの場合でも他の母音の話者を識別するのは困難であるという結果が得られた．しかし，未知話者の母音/a/，/i/，/u/ を用いて個人性を学習すると母音/e/ と/o/ の話者を識別できるようになるという報告もあるため [6]，さらなる検討が必要である．

謝辞 本研究は平成 17 年度総務省 SCOPE-R により実施したものである．ご助言いただいた上智大学言語学専攻 網野加苗氏，ATR 人間情報科学研究所の諸氏に感謝の意を表す．

#### 参考文献

- [1] 北村ら，音響誌 53(3) 185–191, 1997.
- [2] 西尾，言語生活 158, 36–42, 1964.
- [3] 松井ら，音講論 (秋), 379–380, 1993.
- [4] 武田ら，ATR Tech. Rep. TR-I-0028, 1988.
- [5] Kawahara *et al.*, Speech Commun. 27, 187–207, 1999.
- [6] 北村ら，音講論 (秋), 365–366, 1999.