

# 未知話者識別における子音および後続母音の影響\*

◎ 網野 加苗, 荒井 隆行 (上智大・理工)

## 1 はじめに

音声に含まれる話者の個人性をヒトがどのように知覚するかを調べることは、法科学や認知科学の観点から興味深い。著者らの一連の研究<sup>[1-3]</sup>において、聴取する単音節内に鼻音が含まれる場合に話者識別の正答率が高いことが分かっている。本研究では、前回までには扱っていなかった /a/ 以外の母音も含め、刺激音の話者識別正答率に与える影響について調べた結果を報告する。

## 2 実験

### 2.1 音声資料および聴取者

本実験で使用した音声は、電子協日本語共通音声データ<sup>[4]</sup>より、4名の男性話者の単音節発話を使用した。単音節の一覧を表 1 に示す。意図された単音節として聞こえることを、実験者および音声学の知識を持たない日本語母語話者によって確認した。

話者の選定基準は、比較的雑音の少ない環境で録音を行った東京方言話者であることとした。音声データは、標準化周波数 48kHz、量子化精度 16bit で提供されているものをそのまま用いた。

実験に参加した聴取者は、日本語を母語とする健聴者 15 名である。いずれも実験前に本実験の話者の音声を聞いたことがない。

### 2.2 手続き

実験を始める前に、「保留」「変換」「改行」という単語の発話を用いて、聴取者に 4 名の話者の音声を覚えてもらった。これらの単語には、表 1 に示した音節は含まれていない。

これらの単語を、話者 4 名を覚えられるまで何度でも聞いてもらい、同単語を用いた練習試行で 90%以上の正答率が得られるまで学習を繰り返した。1 問ごとにフィードバックを与え、全員 10 分以内に学習を終了した。

本実験では、練習試行とは異なり単音節によって話者を判断してもらうことを予め伝えた上で、576 試行 (4 話者×48 音節×3 発話)を行った。所要時間は平均 40 分であった。

## 3 結果と考察

### 3.1 子音の影響

本実験の結果は、話者識別の正答率によって評価する。表 1 最右列に各子音における話者識別正答率を示し、図 1 (a) にそのグラフを示す。

表 1 および図 1 から、話者識別の正答率は全て偶然確率 (25%) 以上であることが分かる。子音では、/n/ /nʲ/ で正答率が高く、次いで /d/ /z/ が高い。鼻音では、口蓋音 /n/ /nʲ/ の方が両唇音 /m/ よりも有効であり、同様の傾向が接近音 /w/ /j/ 間でも見られた。さらに、子音がない場合に最も正答率が低い。これらの傾向は全て今までの実験結果<sup>[1-3]</sup>と一貫している。

### 3.2 母音の影響

母音別の話者識別正答率を図 1 (b) に示す。この図から分かるように、/o/ /ɯ/ /a/ /e/ /i/ の順に正答率が高いという結果になった。

この結果は、広母音 /a/ の方が狭母音 /i/ より有効であるという報告<sup>[5]</sup>にも反せず、また順序は異なるが後舌母音 /ɯ/ /o/ で高い正答率を示したという報告<sup>[6]</sup>にも共通する。

### 3.3 考察

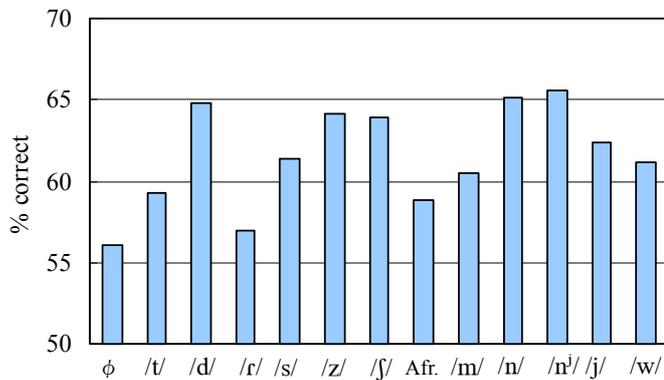
表 1 に示した子音間で分散分析を行ったところ、その差は 5%水準では有意ではなかった ( $p = 0.058$ )。次に阻害音-共鳴音間および鼻音-非鼻音間で検定を行ったところ、後者でのみ有意差が見られた ( $p = 0.045$ )。このことから、[+sonorant] より [+nasal] の素性が話者識別に重要であると言える。鼻音の有効性については、鼻腔・副鼻腔形状の個人差によって説明できる。

後舌・前舌母音間および広・狭母音間で検定を行ったところ、前者においてのみ差が有意であり ( $p = 0.003$ )、[+back] という素性が有効であると言える。後舌母音が有効であるということは、第 2 フォルマントの周波数が低い方が話者を識別しやすいということを意味するが、子音においても歯茎音 /s/ /n/ よりも、それに対応する口蓋化音 /ʃ/ /nʲ/ で正答率が高く、舌を後退させることと話者の個人性の現われには関連がある可能性がある。

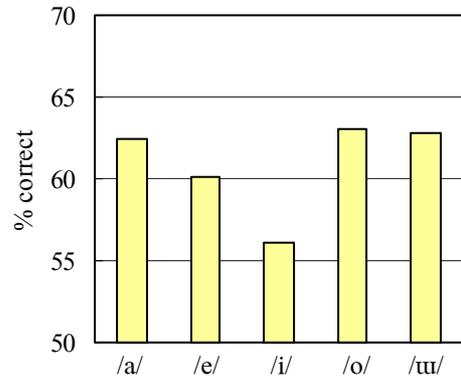
\* Effects of the consonants and the following vowels in monosyllabic stimuli on identification of previously unknown speakers, by AMINO, Kanae and ARAI, Takayuki (Sophia University).

Table 1 Stimulus monosyllables

Consonant		/a/	/e/	/i/	/o/	/ʊ/	% correct
None	ϕ	/a/	/e/	/i/	/o/	/ʊ/	
Stops	/t/	/ta/	/te/	-	/to/	-	
	/d/	/da/	/de/	-	/do/	-	
Tap / Flap	/ɾ/	/ra/	/re/	/ri/	/ro/	/ru/	
Fricatives	/s/	/sa/	/se/	-	/so/	-	
	/z/	/za/	/ze/	-	/zo/	-	
	/ʃ/	/ʃa/		/ʃi/	/ʃo/	/ʃʊ/	
Affricates	/tʃ/ /tʃ/	-	-	/tʃi/	-	/tʃʊ/	
	/dʒ/ /dʒ/	-	-	/dʒi/	-	/dʒʊ/	
Nasals	/m/	/ma/	/me/	/mi/	/mo/	/mu/	
	/n/	/na/	/ne/	/ni/	/no/	/nu/	
	/nʲ/	/nʲa/	-	-	/nʲo/	/nʲʊ/	
Approximants	/j/	/ja/	-	-	/jo/	/ju/	
	/w/	/wa/	-	-	-	-	



(a) Consonants



(b) Vowels

Figure 1. Results of perceptual speaker identification according to (a) consonants and (b) vowels

#### 4 まとめ

本研究では、ヒトによる話者識別実験を行い、刺激音中の子音およびその後続母音が識別正答率に与える影響を調べた。その結果、今までの実験の結果に一貫して、鼻子音を含む音節で話者識別正答率が高く、口音のみの音節との差は有意であった。母音では、/o/ /ʊ/ /a/ の順に後舌母音で正答率が高く、前舌母音との差は有意であった。子音でも歯茎音よりもそれを口蓋化させた後部歯茎音で正答率の上昇が見られたことから、舌の後退と話者の個人性に関連があると思われる。

#### 謝辞

本研究は文部科学省私立大学学術研究化推進事業上智大学オープンリサーチセンター「人間情報科学研究プロジェクト」の助成を得た。

#### 参考文献

- [1] Amino *et al.*, Acoust. Sci. Tech., 27 (4), 233-235, 2006.
- [2] Amino *et al.*, IEICE Tech. Rep., 105, 109-114, 2006.
- [3] Amino *et al.*, Acoust. Sci. Tech., 28 (2), 128-130, 2007.
- [4] 社団法人 電子情報技術産業協会  
[http://www.sunrisemusic.co.jp/dataBase/fl/dataBase00\\_fl.html](http://www.sunrisemusic.co.jp/dataBase/fl/dataBase00_fl.html)
- [5] 北村, 赤木, 音響誌 (E), 16(5), 283-289, 1995.
- [6] 蒔苗ら, 音講論 (秋), 97-98, 2007.