

子音部・母音部の入れ替えが個人性知覚に及ぼす影響について*

◎網野加苗, 荒井隆行 (上智大)

1 はじめに

音声には発話の内容を表す言語学的情報以外に、話者個人に関する情報も含まれている。音声の個人性の研究アプローチとして、聴取による話者識別実験を行うことがあげられる [1]。音声の個人性知覚には、刺激の長さや内容による差があることが知られており [2]、母音 [3, 4] や鼻音 [5, 6] を刺激として呈示した際に話者識別正答率が高くなることが報告されている。

また、筆者らの前研究 [7] では、知覚における話者間の混同と刺激音の音響特性の関連について調べ、話者識別率の高かった「鼻子音+母音」という音節では子音区間も含む音節全体で話者間距離が大きいのに対し、「口子音+母音」の音節ではその母音部のみで知覚との相関が見られることを示した。

一方、子音部・母音部の音韻性及び個人性への貢献を調べた研究 [8] では、子音・母音をそれぞれ無音で置き換えた刺激を用いて語認識及び話者認識実験を行い、その結果、母音は個人性に貢献し、子音は音韻性に貢献すると結論付けている。

本研究では、子音部・母音部の話者間の入れ替えが個人性知覚に与える影響を調査した。

2 音声資料

全ての録音は防音室で行われた。録音条件は表 1 に示した通りである。話者には各 CV 単音節を 5 回ずつ発話してもらい、刺激の作成にはそのうちの 2 発話分 (2 トークン) を使用した。音声資料からターゲットとなる CV 音節を目視によって切り出し、それぞれについて子音部・母音部の境界を決定した。境界の決定には Furui [9] による D パラメータを用いた。

境界が求まった後、音節ごとに話者内・話者間で子音

部・母音部の入替を行い、表 2 に示すように刺激音を作成した。刺激音の音節は 8 つの子音と /a/ から成る単音節である。各刺激の繰返し回数は 2 回であった。なお、今回の実験では異なる子音間での話者の入替は行っていない。

3 聴取実験

3.1 手続き

実験は録音と同じ防音室で行われた。聴取者は日本語を母語とする健聴者 13 名で、全員本実験の話者とは知り合っていない。

聴取者は、まず 2 名の話者のサンプル音声を聞いて話者の学習を行い、このサンプル音声を用いた練習で正答率が 100% になってから本番の実験に移った。サンプル音声は実験で用いた刺激音とは別の音声である。練習に要した時間は全員 5 分以内であった。本実験の試行数は 352 回で、刺激はランダムな順序で呈示された。実験全体に要した時間は 20 分程度であった。

3.2 結果と考察

図 1 から図 3 に、表 2 に示した刺激種別ごとの結果を示す。図 1 は原音声の、図 2 は話者内入替音声の話者識別正答率であり、図 3 は話者間入替音声で母音部の話者と答えた回答率である。




表 1. 録音条件

話者	男性 2 名 (東京方言, 20 歳代)
機器	SONY DAT レコーダ (TCD-D8) SONY マイク (ECM-959A)
実験語	キャリア文中の CV 単音節 /da/ /ma/ /na/ /nja/ /ra/ /sa/ /ta/ /za/
標準化周波数	48 kHz
量子化精度	16 bit

表 2. 刺激作成における子音ごとの組合せパターン

↓ C・V →	Sp1-token1	Sp1-token2	Sp2-token1	Sp2-token2
Sp1-token1	Sp1-C1 + Sp1-V1	Sp1-C1 + Sp1-V2	Sp1-C1 + Sp2-V1	Sp1-C1 + Sp2-V2
Sp1-token2	Sp1-C2 + Sp1-V1	Sp1-C2 + Sp1-V2	Sp1-C2 + Sp2-V1	Sp1-C2 + Sp2-V2
Sp2-token1	Sp2-C1 + Sp1-V1	Sp2-C1 + Sp1-V2	Sp2-C1 + Sp2-V1	Sp2-C1 + Sp2-V2
Sp2-token2	Sp2-C2 + Sp1-V1	Sp2-C2 + Sp1-V2	Sp2-C2 + Sp2-V1	Sp2-C2 + Sp2-V2

刺激音の種別

-  原音声 (入替操作を行わない音声): 表中の 4 つ (各話者 2 トークン)、+ 表にない 6 つ (各話者 3 トークン)
-  話者内入替音声 (同話者のトークン間での入替音声): 表中の 4 つ (各話者 2)
-  話者間入替音声 (子音と母音で異なる話者の音声): 表中の 8 つ

* The effects of inter-speaker consonant- and vowel-swapping on perceptual speaker identification, by AMINO, Kanae and ARAI, Takayuki (Sophia Univ.).

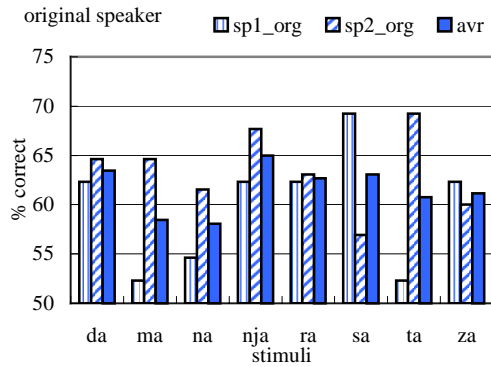


図 1. 原音声の話者識別正答率

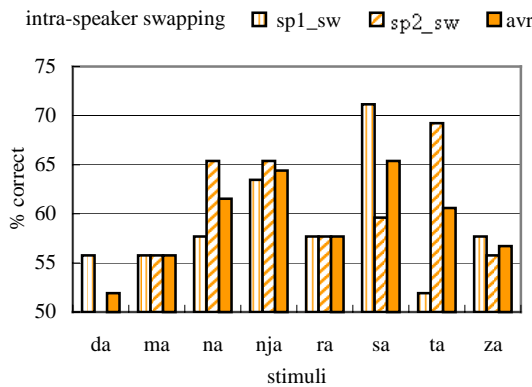


図 2. 話者内入替音声の話者識別正答率

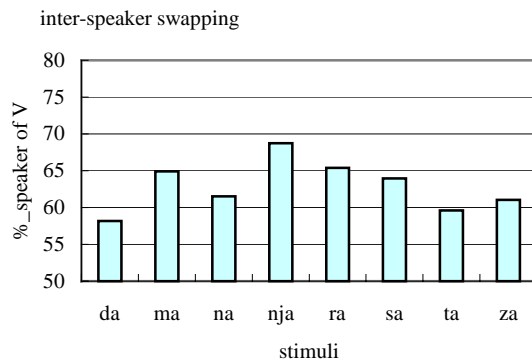


図 3. 話者間入替音声で母音部の話者と答えた割合

本実験の聴取者は 13 名であったため、刺激音の子音別の評価回数 (各棒グラフにおける刺激数) は、原音声で 130 回 (10×13)、話者内入替音声で 52 回 (4×13)、話者間入替音声では 208 回 (16×13) であった。

図 1 より、原音声では全子音において話者識別の正答率が偶然確率 (50%) 以上であったことが分かる。正答率は /nja/ で最も高く、次いで /da/, /sa/ の順であった。図 2 の話者内入替音声では、図 1 と比較して /da/ の正答率の下降が目立つが、話者内入替による大きな影響は特に見られない。話者識別正答率における子音の順序は、今までの実験 [5-6] のように、全ての鼻音で正答率が高いという結果とは異なるが、実験のタス

クが異なるため単純には比較できない [10]。また、原音声、話者内入替音声ともに /sa/, /ta/ では話者間のばらつきが大きく、一方の話者が特徴的であるために他方と区別された可能性が捨てきれない。

子音部と母音部で異なる話者の音声を用いた話者間入替音声では正答は存在しないが、図 3 で全ての刺激において母音部の話者と答えた割合が 50% を超えていることから、CV 単音節における子音部と母音部の入替では、個人性知覚は母音部の影響をより受けやすいことが分かる。その理由としては、母音部はよりエネルギーが大きく、ソノリティも高いことがあげられる。また、持続時間の違いも関わっていると考えられる。

子音の種類に関しては /nja/, /ra/, /ma/ の順に母音部の影響を受けやすく、調音方法別には、若干の相違はあるものの、今までの実験 [5-6] における話者識別正答率の順序と同様、鼻音、摩擦音、破裂音の順になっていると言える。ここで子音による効果は有意で、特に /nja/ と /da/ の間に有意な差が見られた ($p < 0.01$)。

今回は /nja/ に関しては話者識別正答率及び母音部の話者の回答率が共に高かったが、「鼻子音+母音」の音節で特に母音部の話者と知覚される傾向が強いということは、鼻音化された母音に個人性が多く含まれる可能性を示唆している。今後は、本研究では話者間の入替を行わなかった、異なる子音間での入替実験を行うことで、鼻音化・非鼻音化母音における個人性に関して比較することができるだろう。

4 まとめ

単音節の子音部・母音部で話者を入れ替えて個人性知覚の実験を行った結果、話者の個人性の判断は母音部の影響をより受けやすいことが分かった。特に子音部の調音方法別に見ると、その回答率は今までの話者識別実験における識別正答率と同じ順序になっており、鼻音化環境にある母音に特に話者の個人性が現れている可能性が示された。

謝辞

本研究は日本学術振興会特別研究員奨励費 (17・6901) の助成を得た。

引用文献

- [1] O'Shaughnessy, "Speech Communication," Addison-Wesley Publishing Company, 2000.
- [2] Nygaard, "The Handbook of Speech Perception," Blackwell Publishing, 2005.
- [3] 西尾, 言語生活, 158, 36-42, 1964.
- [4] 新美, "音声認識," 共立出版, 1979.
- [5] 網野, 信学技報, 104(149), 49-54, 2004.
- [6] Amino et al., Acoust. Sci. Tech., 27 (4), 233-235, 2006.
- [7] 網野, 荒井, 音講論(秋), 273-274, 2006.
- [8] Owren and Cardillo, JASA, 119 (3), 1727-1739, 2006.
- [9] Furui, JASA, 80 (4), 1016-1025, 1986.
- [10] Bricker and Pruzansky, "Experimental Phonetics," Lass (ed.), Academic Press, 1976.