

# ATR 音声データベースセット C の文音声の個人性類似度

川元 広樹<sup>†</sup> 北村 達也<sup>†</sup>

<sup>†</sup> 甲南大学知能情報学部知能情報学科

〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

E-mail: †si971038@center.konan-u.ac.jp, ††t-kitamu@konan-u.ac.jp

**あらまし** ATR 音声データベースセット C の関東出身男性話者 20 名による文音声を対象にして聴取実験に基づき音声の個人性の類似度評価を行った。実験協力者 40 名を対象に 2 話者の文音声を聴かせ、その類似度を評定させた。得られた結果を多次元尺度構成法により処理して話者を知覚空間上に布置した。

**キーワード** 音声の個人性, 文音声, ATR 音声データベースセット C, 多次元尺度構成法

## Similarity of speaker individualities of sentence in ATR speech database set C

Hiroki KAWAMOTO<sup>†</sup> and Tatsuya KITAMURA<sup>†</sup>

<sup>†</sup> Faculty of Intelligence and Informatics, Konan University

8-9-1, Okamoto, Higashinada, Kobe, Hyogo, 658-8501 Japan

E-mail: †si971038@center.konan-u.ac.jp, ††t-kitamu@konan-u.ac.jp

**Abstract** We measured perceptual similarity of speaker individualities for a sentence of twenty male Japanese speakers in ATR speech database set C. Forty participants evaluated perceptual similarity of the sentence of pairs of speakers. We obtained inter-speaker distances by a multidimensional scaling analysis on the basis of the results of the perceptual experiments.

**Key words** Speaker individualities, Spoken sentence, ATR speech database set C, Multidimensional scaling

### 1. はじめに

従来、音声の個人性の知覚要因に関する研究が数多く行われてきたが、ほとんどの場合は研究者もしくは研究組織独自の音声データが用いられている。そのため、得られた結果はその話者セットに少なからず依存することが否定できない。この問題を解消するためには、研究者が共通の音声データベースを用いて研究し、その結果を相互に比較検証できるようにしていく必要がある [1]。

このような折、高度言語情報融合フォーラム (ALAGIN) から ATR 音声データベースが研究目的で公開された。この中にはセット C と呼ばれる男女各 120 名による単語、数字、および音素バランス文の音声が含まれており、それらには音声セグメントラベルが付与されている。この規模の話者数が収録された日本語音声データベースはほとんどなく、現段階で個人性研究の共通基盤とするには十分と考えられる。

そこで、本研究ではこの音声データベース中の関東出身男性話者 20 名による文音声を対象にして、聴取実験に基づき個人性の類似度評価を行った。

### 2. 実験条件

#### 2.1 刺激音

ATR 音声データベースセット C から、19 歳から 39 歳の関東 (東京, 神奈川, 千葉, 埼玉) 出身の男性話者 20 名 (M109, M113, M211, M214, M318, M409, M504, M508, M509, M510, M517, M519, M520, M601, M603, M614, M705, M710, M714, M718) を選択した。音声データに雑音が含まれる話者は除外してある。対象とした文は「冷房では冷えすぎが問題になる」である。標準化周波数 20 kHz, 量子化ビット数は 16 bit で、話者間で最大振幅を正規化した。

#### 2.2 実験協力者

19 歳から 23 歳の聴覚に異常のない男性 34 名, 女性 6 名の計 40 名のデータを分析対象にした。

#### 2.3 聴取実験の方法

2 つの刺激音を 1 組として、話者 20 名のすべての組み合わせで刺激音を提示した。刺激音間の無音区間は 0.3 s である。順序効果を排除するため提示順を入れ替えた刺激対も提示したので、刺激対の数は 400 である。この刺激対を各 1 回提示した。

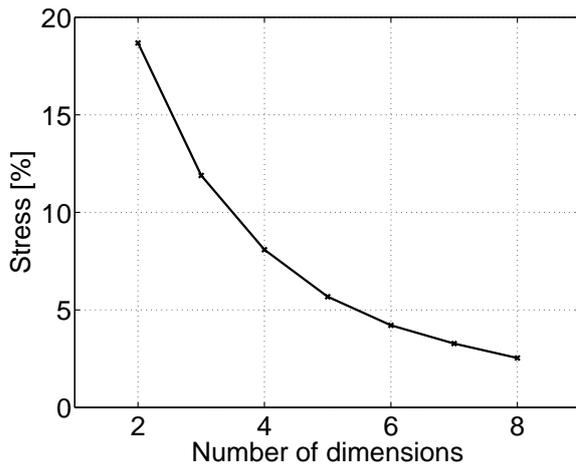


図1 ストレス値 [%]

刺激対はランダムに提示した。100 試行を 1 セットとし、実験協力者 1 名に対して計 4 セットを実施した。

実験協力者は、各刺激対に対し、5 段階 (似ていない、あまり似ていない、やや似ている、似ている、同一人物) [2] で類似性を評定した。聴き直しは 1 度だけ許した。

刺激音は、PC から出力された音声をヘッドフォンアンプ (Fostex HP-A3) にて D/A 変換し、密閉型ヘッドフォン (Sennheiser HDA200) にて提示した。実験協力者は各自の聴きやすいレベルで聴取した。

## 2.4 分析方法

本研究では、上記の 5 段階評定に 1 から 5 の数字を割り当て、実験協力者間で平均した評定値を非計量多次元尺度構成法にて分析した。それぞれ、1: 似ていない、2: あまり似ていない、3: やや似ている、4: 似ている、5: 同一人物、とした。計算には統計処理ソフトウェア *R* を使用した [3]。

なお、2 つの刺激音の話者が同じ刺激対に対して、「同一人物」と評定した回数が他の実験協力者と比較して著しく少ない実験協力者のデータは除外した。本研究では最終的に 40 名の実験結果を分析対象とした。

## 3. 結果と考察

評定値の標準偏差を計算したところ、全 400 対のうち 399 対で 1 未満であった。これは実験協力者間で評定に大きな差異がないことを意味している。従って、本研究の結果はおおむね信頼できると判断した。

MDS 分析におけるストレス値を図 1 に示す。ストレス値とは、 $n$  次元におけるデータ間の心理距離の適合度を示すものであり、その値が小さいほど適合度は高くなる。

MDS により求めた個人性類似度の空間配置を図 2 に示す。図中の数字と話者は表 1 の通り対応している。本研究では、ストレス値が 5 % 以下となる 6 次元を採用した。ストレス値 5 % の適合度は「良い適合」といわれている [3]。

この空間配置においては、データ間の距離に近いほど話者の知覚的類似度が高いと評価される。この結果によると、話者 M109, M113, M601 の組、話者 M318 と M510 のペアなどの

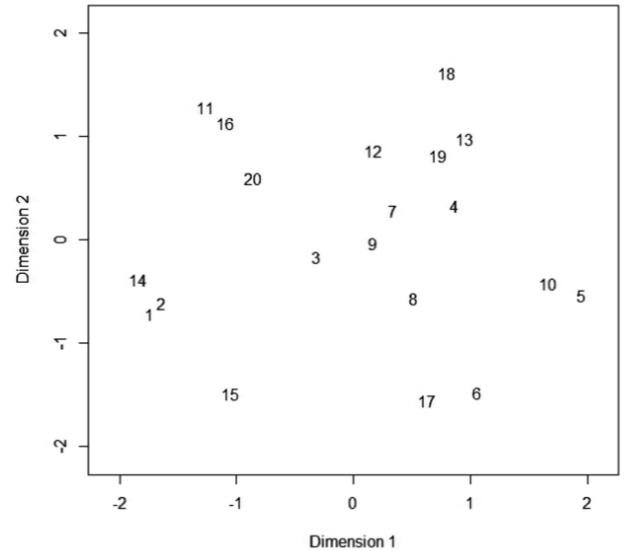


図2 話者 20 名の個人性類似度の iso MDS による空間配置

表 1 図 2 中の番号と話者 ID の対応表

No.	ID	No.	ID	No.	ID	No.	ID
1	M109	6	M409	11	M517	16	M614
2	M113	7	M504	12	M519	17	M705
3	M211	8	M508	13	M520	18	M710
4	M214	9	M509	14	M601	19	M714
5	M318	10	M510	15	M603	20	M718

距離が近い。著者らが聞いた限りでは、実際にこれらの話者の音声は類似しているように感じられた。また、著者らの印象では、第 1 次元は「声の高さ」、第 2 次元は「迫力」に対応するように感じられた。

## 4. おわりに

ATR 音声データベースの C セット内の関西出身男性話者 20 名の文音声を対象にして、聴取実験に基づき、個人性類似度の空間配置を行った。今後、空間配置の各次元に対応する特徴量の検討、他の文でも同様の結果が得られるかに関する検討などが必要である。

実験結果の詳細を web ページにて公開し、音声の個人性に関する研究に役立ててもらう予定である。

### 謝辞

本研究の一部は平成 24 年度科研費基盤研究 (B)(21300071) および平成 24 年度私立大学等経常費補助金の支援を得て行われた。ATR 音声データベース C セットの開発および公開に関わられた皆様に感謝いたします。

### 文 献

- [1] 北村達也, 出水田剛志, 橋亮輔, 声の類似性から個人性知覚を探る, 音講論, 2, 253-256 (2011).
- [2] 出水田剛志, 音声の音響特徴量の動的成分が個人性知覚に与える影響に関する研究, 北陸先端科学技術大学院大学 修士論文 (2012).
- [3] 中村永友, *R* で学ぶデータサイエンス 2 多次元データ解析法, 共立出版 (2009).