

## [ポスター講演] 通る声を持つ話者の声質の心理的評価

中尾 天<sup>†</sup> 北村 達也<sup>†</sup>

<sup>†</sup> 甲南大学知能情報学部 〒658-8501 兵庫県神戸市東灘区岡本 8-9-1

E-mail: <sup>†</sup>s1771067@s.konan-u.ac.jp, <sup>††</sup>t-kitamu@konan-u.ac.jp

**あらまし** 男女各 140 名の文音声から「通る声」を持つ話者を抽出させる聴取実験を行い、その上位の男女各 5 名、下位の男女各 5 名を選択した。それらの話者の音声を対象に 13 の表現語対を用いて印象評定実験を行った。因子分析の結果、第 1 因子として美的因子、第 2 因子として金属性・迫力因子が抽出された。第 1 因子、第 2 因子の平面上に各話者を布置した結果、第 1 因子は男女の違いに対応し、第 2 因子が通る声か否かに対応した。従って、通る声は金属性・迫力因子を持つ声質といえる。

**キーワード** 声質, SD 法, 因子分析, 金属性因子, 迫力因子

## Psychological evaluation of popping-out voice quality

Takashi NAKAO<sup>†</sup> and Tatsuya KITAMURA<sup>†</sup>

<sup>†</sup> Faculty of Intelligence and Informatics, Konan University 8-9-1 Okamoto, Higashinada-ku, Kobe, Hyogo 658-8501 Japan

E-mail: <sup>†</sup>s1771067@s.konan-u.ac.jp, <sup>††</sup>t-kitamu@konan-u.ac.jp

**Abstract** The present study evaluated the “popping-out” voice quality through auditory tests using a semantic differential method. Participants selected speakers who have popping-out voice from 280 speakers (140 male and 140 female) of ATR speech database set C, and then evaluated speech sounds of speakers who have and do not have popping-out voice by thirteen bipolar adjective pairs. Results of a factor analysis suggested that the popping-out voice quality has metallic and powerful qualities.

**Key words** Voice quality, Semantic differential method, Factor analysis, Metallic voice quality, Powerful voice quality

### 1. はじめに

我々は「通る声」という声質があることを経験的に知っており [1], 通る声を持つ話者が存在することも知っている。スーパードictionary [2] の「通る」の項目にも「声や音が遠くまで伝わる」という語義が含まれている。「通る声」のこのような特性を生み出す音響的な原理やその心理的な効果を明らかにすることによって、公共空間等における効果的かつ効率的な情報伝達に活用できる可能性がある。天野ら [3] は通る声を「ポップアウトボイス」と呼び、その技術的応用を目指して検討を始めている。

横山と井上 [4] は男女計 13 名の単母音/e/を対象にして声質の表現語対を用いた評価実験を行った。彼らは、表現語対「通る声-通らない声」が別の表現語対「澄んだ声-濁った声」と相関が高く、この 2 つの表現語対が音声の好ましさと相関が高いことを示した。さらに、音声为好まれる話者のスペクトル包絡では、そうでない話者と比較して第 3 フォルマント周波数 (F3) と第 4 フォルマント周波数 (F4) が近く、その領域のエネルギー

が高いことを報告している。

彼らの報告は、F3, F4 の周波数帯域の振幅が「通る声」の声質と関連することを示唆している。この周波数帯域は、アナウンサーの音声にて振幅が大きい帯域 [5], 歌手のフォルマント (singer’s formant) の現れる帯域 [6], [7] と一致する。また、Krause と Braida [8] は通常の話速で発話された clear speech の長時間平均スペクトル (LTAS) の 1 kHz から 3 kHz の帯域の振幅は通常の音声よりも増加することを報告し、Bele [9] は知覚的に良い声質 (better normal voice quality) は LTAS における 3 kHz から 4 kHz の帯域のピーク (彼女は speaker’s formant と呼んでいる) と関連すると報告している。上記の報告は、通る声を含めこれらの類似した現象・事象が音響、音声生成、音声知覚の各側面において密接に関連することを示唆している。

本研究では、通る声の知覚特性の一端を明らかにするために、印象評定実験を実施する。音声データベースから通る声を持つ話者と持たない話者を抽出し、それらの音声を対象に Semantic differential 法 (SD 法) を用いて印象評価を行う。そして、得られた結果から通る声の特性を明らかにする。

## 2. 予備実験

話者 280 名の文音声を対象にした聴取実験により通る声を持つ話者を抽出した。

### 2.1 方法

#### 2.1.1 刺激音

ATR 音声データベース セット C [10] 内の話者 (男女各 140 名) の文音声「あらゆる現実を全て自分の方へねじ曲げたのだ」を対象にした。このデータセットには、日本各地を出身とする様々な世代の話者の音声が含まれている。音声データは 20 kHz にて標準化され、16 bit にて量子化されている。天野ら [3] のようなノイズを加える処理は施していない。

#### 2.1.2 実験参加者

日本語を母語とし、聴覚に関する問題がないと自己申告した大学生 13 名 (男性 7 名, 女性 6 名) が参加した。参加者には謝礼が支払われた。

#### 2.1.3 手続き

上記の音声データを参加者各自の PC にダウンロードさせ、好みの音量にて聴取させた。音声を聴取する回数に制限を設けず、話者の聴取順序は指定しなかった。通る声に関して以下の説明を与え、話者 280 名の中から通る声を持つと思った話者を回答させた。

「通る声」がどういうものかについてはご自身のイメージで結構ですが、おおよそ「ざわざわしている中で、その人の声を取り立てて大きいわけでもないのによく聞こえそう」というものです。

選択の際には、「どちらかと言えば通る声」というレベルでなく、「とても良く通る声」を選ぶよう指示し、選択する話者数に制限は設けなかった。

### 2.2 結果

男女各 140 名の話者のうち、通る声であると判定した実験参加者が多かった上位 5 名を表 1, 表 2 に示す。話者は ATR 音声データベース セット C 内でのラベルである。各実験参加者が選択した通る声を持つ話者の人数は、最小 6 名, 最大 146 名, 平均 64.54 名 (標準偏差 41.74) であった。

上位 5 名の話者に対して、参加者 13 名中 8 名以上が通る声と判定していた。この結果は、参加者間にある程度共通した通る声に対するイメージが存在することを示している。また、通る声と判定された話者の年齢は特定の年齢層に偏ることなく分布していた。

## 3. 声質の心理的評価

通る声、通らない声を持つ話者の音声を SD 法により評価した。

### 3.1 方法

#### 3.1.1 刺激音

予備実験にて用いた音声データのうち、表 1, 表 2 に示した男女各 5 名の音声と、通る声であると判定した実験参加者がいなかった男女各 5 名の音声をを用いた。後者のうち男性話者は

表 1 通る声であると判定された男性話者上位 5 名。スコアは通る声であると判定した人数 (13 名中)。

Table 1 Top five male speakers who have most popping-out voice.

順位	話者	年齢	スコア
1	M618	42	10
2	M110	31	9
3	M203	21-25	9
4	M311	51	9
5	M405	不明	9

表 2 通る声であると判定された女性話者上位 5 名。スコアは通る声であると判定した人数 (13 名中)。

Table 2 Top five female speakers who have most popping-out voice.

順位	話者	年齢	スコア
1	F604	26	11
2	F602	29	10
3	F314	36-40	9
4	F109	33	8
5	F119	39	8

M316, M406, M506, M520, M616, 女性話者は F311, F319, F418, F504, F615 である。本稿ではこれらの話者を通らない声を持つ話者と呼ぶ。すべての音声データは Praat [11] にて振幅を正規化した。

#### 3.1.2 実験参加者

日本語を母語とし、聴覚に関する問題がないと自己申告した 20 歳から 22 歳の大学生 30 名 (男性 24 名, 女性 6 名) が参加した。参加者には謝礼が支払われた。

#### 3.1.3 手続き

実験に先立ち、書面により実験の説明を行い、実験参加の同意書に署名を得た。

実験参加者は上記 20 名の音声をランダムに聴取し、以下の表現語対を用いて両極 7 段階にて声質を評価させた。繰り返し聴取は許した。

- (1) 澄んだ - 濁った
- (2) きれい - 汚い
- (3) 力強い - 弱々しい
- (4) 鋭い - 鈍い
- (5) 迫力がある - 物足りない
- (6) 固い - 柔らかい
- (7) 高い声 - 低い声
- (8) 澄んだ声 - かすれた声
- (9) 落ち着きのある声 - 落ち着きのない声
- (10) 太い声 - 細い声
- (11) 張りのある声 - 張りのない声
- (12) 男性的な声 - 女性的な声
- (13) 若い感じの声 - 老けた感じの声

これらの表現語対のうち、(1) から (6) は音色因子の中で代表的な美的因子、金属性因子、迫力因子の表現語対である [16]。 (7) から (13) は木戸と粕谷 [15] が提案した声質の表現語対リストより (5) と重複する「迫力のある声 - 弱々しい声」を削除したも



図1 Microsoft Forms による印象評定実験のスクリーンショット  
Fig.1 Screenshot of web form used in voice quality evaluation.

のである。

実験は換気された静かな部屋で行った。刺激音はノート PC (hp ProBook 430 G3) とヘッドフォン (SHURE SRH440) を用いて提示した。その際、1 kHz 純音の等価騒音レベルが 50 dB となる音量に設定した。この値はヘッドフォン用の簡易音圧確認治具 ChomeJig [17] にて計測した。

実験参加者は、各刺激音に対して上記の表現語対を 7 段階で評価した。評価結果の入力、回収には Microsoft Forms を用いた。入力画面の例を図 1 に示す。表現語対の評価順および表現語対の左右の配置はランダム化し、カウンターバランスをとった

### 3.2 結果

通る声、通らない声を持つ男女話者の各表現語対についての評価値を図 2、図 3 に示す。これらの値は話者間、実験参加者間にて平均されている。男性の通る声を持つ話者と通らない声を持つ話者にて差が大きい表現語対は「力強い-弱々しい」(差: 0.99), 「迫力がある-物足りない」(0.77), 「張りのある-張りのない」(0.75)であった。一方で、女性話者では多数の表現語対で通る声と通らない声に対する評価の差異が大きかった。特に、「張りのある-張りのない」(1.80), 「力強い-弱々しい」(1.57), 「迫力がある-物足りない」(1.35), 「鋭い-鈍い」(1.17)において差が大きかった。これらの表現語対は男女でほぼ一致しており、通る声の特徴を表しているといえる。

評定結果に対して統計分析用フリープログラム HAD [12]~[14] にて因子分析を行った。因子数は 3 とし、主因子法にて因子抽出を行い、プロマックス回転にて因子軸を回転させた。得られた結果を表 3 に示す。第 1 因子、第 2 因子の  $\alpha$  係数がいずれも 0.8 を超えているため、これらの因子は信頼できるといえる。

第 1 因子は「澄んだ-濁った」、「きれい-汚い」等の表現語

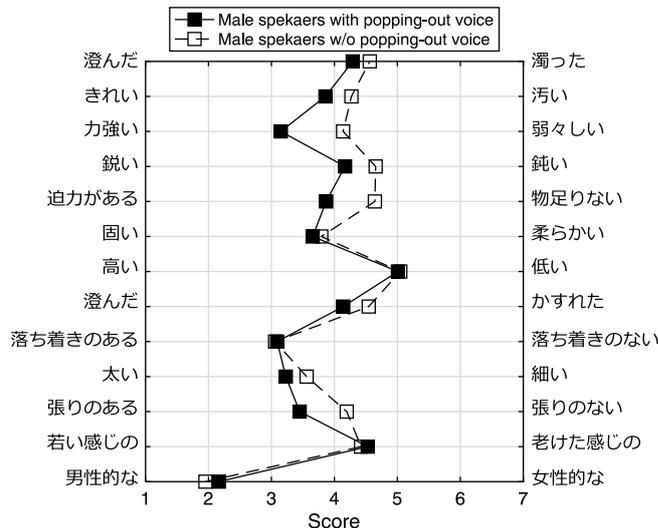


図2 通る声 (■), 通らない声 (□) を持つ男性話者に対する評価値  
Fig.2 Profile analysis for male speakers who have and do not have popping-out voice.

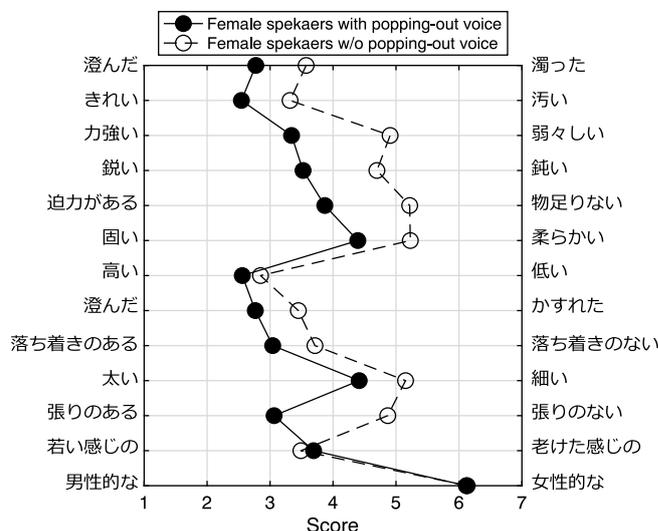


図3 通る声 (●), 通らない声 (○) を持つ女性話者に対する評価値  
Fig.3 Profile analysis for female speakers who have and do not have popping-out voice.

対の因子負荷が高かったため、「美的因子」とした。第 2 因子は「鋭い-鈍い」等の金属性を表す表現語対、「力強い-弱々しい」等の迫力を表す表現語対の因子負荷が高かったため、「金属性・迫力因子」とした。

各話者の因子得点を第 1 因子-第 2 因子平面上にプロットした結果が図 4 である。この図から、第 1 因子は男女を分ける軸であることがわかる。一方、第 2 因子は通る声を持つ話者と通らない声を持つ話者を分ける軸とみなすことができる。つまり、本研究にて得られたデータからは、通る声は金属性および迫力を有すると考えることができる。

## 4. 考察

本研究では通る声を持つ話者と持たない話者の文音声を対象にして、SD 法により印象評定実験を行った。因子分析の結果、

表3 回転後の因子負荷量と因子間相関

Table 3 Results of factor analysis.

表現語対	因子負荷量		
	第1因子	第2因子	第3因子
1. 美的因子			
澄んだ - 濁った	0.989	0.186	-0.210
きれい - 汚い	0.971	0.192	-0.245
澄んだ声 - かすれた声	0.968	0.155	-0.172
高い声 - 低い声	0.894	0.012	0.249
男性的な声 - 女性的な声	-0.842	0.123	-0.630
太い声 - 細い声	-0.713	0.374	-0.237
若い感じの声 - 老けた感じの声	0.406	-0.280	0.254
2. 金属性・迫力因子			
張りのある声 - 張りのない声	0.176	1.013	0.036
力強い - 弱々しい	-0.088	0.922	-0.061
鋭い - 鈍い	0.552	0.863	0.165
迫力がある - 物足りない	0.036	0.838	-0.203
固い - 柔らかい	-0.678	0.769	0.226
3. その他の因子			
落ち着きのある声 - 落ち着きのない声	0.089	0.043	-0.981
因子間相関			
第1因子	1.000	-0.640	0.236
第2因子	-0.640	1.000	-0.686
第3因子	0.236	-0.586	1.000
$\alpha$ 係数	0.913	0.924	—

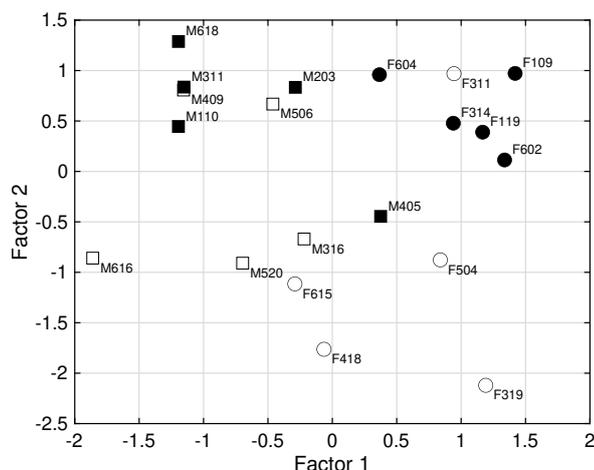


図4 第1, 第2因子に関する各話者の因子得点分布。塗りつぶしたマーカーは通る声を持つ話者, 白抜きのマーカーは通る声を持たない話者に対応する。四角のマーカーは男性, 丸のマーカーは女性を表す。

Fig. 4 Factor score plot for Factors 1 and 2.

通る声と通らない声は金属性・迫力因子によっておおよそ分類できることが明らかになった(図4)。この因子は、図2、図3において差の大きかった表現語対と対応する(表3)。

岩宮[16]によれば、音色の金属性因子はスペクトル構造、特に周波数軸上での成分の分布と対応し、この性質を持つ音は高域に強い成分を有する。つまり、通る声もこのような性質を持つ可能性が高い。このことは、横山と井上[4]の報告と対応するとともに、通る声とアナウンサーの声質[5]、歌手のフォルマ

ント[6],[7]、clear speech[8]、speaker's formant[9]との関連を示唆しており、明瞭な声質や耐雑音性の高い声質に通底する原理の存在を期待させる。

Hanayamaら[18]は、母音/e/を対象に“metallic voice”の分析を行い、この声質の音声では第2, 第3, 第4フォルマントの振幅が増加し、知覚上大きく聞こえると報告している。彼女らの報告は、通る声が金属性因子に加え、迫力因子とも関連するとして本研究の結果と対応する。しかし、本研究では刺激音の振幅を正規化の際に聴覚特性や音響系の特性を考慮しなかったため、迫力因子と通る声の関連については今後も検討が必要である。

## 5. おわりに

本研究では、SD法により通る声と通らない声の印象評定実験を行い、これら2つの声質は金属性・迫力因子の有無によりおおまかに分けられることを報告した。今後、通る声の生成要因の調査を通して、歌手のフォルマント等との関連を明らかにすることが期待される。

## 謝 辞

本研究はJSPS科研費(JP20H0029)の支援により行われた。本研究にて使用した音声データベースは高度言語情報融合フォーラムより提供を受け、簡易音圧確認治具 ChomeJig は和歌山大学システム工学部 入野俊夫先生よりご提供いただきました。また、因子分析については国立国語研究所 籠宮隆之様、甲南大学 知能情報学部 山中仁寛先生よりご助言いただきました。

## 文 献

- [1] 粕谷英樹, 楊長盛, “音源から見た声質,” 音響誌, 51(11), 869–875, 1995.
- [2] スーパー大辞林, 三省堂, 東京, 2006.
- [3] 天野成昭, 河原英紀, 坂野秀樹, 牧勝弘, 山川仁子, “バブルノイズ環境下におけるポップアウトボイスの評価実験,” 音講論 (春) (2021年3月発表予定).
- [4] 横山正人, 井上和夫, “声質感覚に基づく個人性知覚情報の抽出,” 人間工学, 20(1), 41–48, 1984.
- [5] 桑原尚夫, 大串健吾, “アナウンサー音声の音響的特徴,” 信学論 A, J66-A(6), 545–552, 1983.
- [6] J. Sundberg, “Articulatory interpretation of the “singing formant,”” *J. Acoust. Soc. Am.*, 55(4), 838–844, 1974.
- [7] ヨハン・スンドベリ, (訳) 榊原健一, 伊藤みか, 小西知子, 林良子, 歌声の科学, 東京電機大学出版局, 2007.
- [8] J. C. Krause and L. D. Braida, “Acoustic properties of naturally produced clear speech at normal speaking rates,” *J. Acoust. Soc. Am.*, 115(1), 362–378, 2004.
- [9] I. V. Bele, “The speaker’s formant,” *J. Voice*, 20(4), 555–578, 2006.
- [10] 竹沢寿幸, 中村篤, 隅田英一郎, “ATR の会話音声翻訳研究用データベース,” 音声研究, 4(2), 16–23, 2000.
- [11] P. Boersma and D. Weenink, Praat: doing phonetics by computer, <https://www.fon.hum.uva.nl/praat/> (2020年12月30日閲覧)
- [12] 清水裕士, 統計ソフト HAD, <https://osf.io/32cyp/> (2020年12月30日閲覧)
- [13] 清水裕士, “フリーの統計分析ソフト HAD: 機能の紹介と統計学習・教育, 研究実践における利用方法の提案,” メディア・情報・コミュニケーション研究, 1, 59–73, 2016.
- [14] 小宮あすか, 布井雅人, Excel で今すぐはじめる心理統計: 簡単ツール HAD で基本を身につける, 講談社, 東京, 2018.
- [15] 木戸博, 粕谷英樹, “通常発話の声質に関連した日常表現語の抽出,” 音響誌, 55(6), 405–411, 1999.
- [16] 岩宮真一郎, 音のチカラ: 感じる, 楽しむ, そして活かす, コロナ社, 東京, 2017.
- [17] 河原英紀, 榊原健一, 水町光徳, 入野俊夫, “音声収集と聴取における対話的実時間音響計測ツールの応用について,” 聴覚研資, 401–406, 2021.
- [18] E. M. Hayamama, Z. A. Camargo, D. H. Tsuji, and S. M. R. Pinho, “Metallic voice: Physiological and acoustic features,” *J. Voice*, 23(1), 62–70, 2009.