

ポップアウトボイスの個人内変化に伴う 声帯振動特性と声道音響特性の変化*

○北村 達也 (甲南大), 能田 由紀子 (国語研), 榊原 健一 (北海道医療大),
河原 英紀 (和歌山大), 天野 成昭 (愛知淑徳大)

1 はじめに

いわゆる“通る声”を持つ話者の音声は、大きな背景雑音の下でも目立ってよく聞こえ、遠くまで届くことを我々は経験的に知っている。著者らは、この通る声を“ポップアウト”ボイスと呼び [1], これを拡声音声の伝達性向上に活かすためにその知覚特性および音響特性の解明に取り組んでいる。そして、ポップアウトボイスはポップアウトしにくい音声と比較して雑音下でも検出されやすいこと [2], その効果は当該の言語の知識にあまり関係しないこと [3] などを報告している。

本研究では、ポップアウトボイスを持つ話者がポップアウトボイスとそうでない音声 (通常音声) を発声した際の音声および電気グロトグラフィ (Electroglottography, EGG) データから、ポップアウトボイスの生成面の特徴を分析する。以下、ポップアウトボイスを PV, 通常音声を NV と表す。

2 データ

2.1 話者

プロダクションに所属するナレーターのうち、著者らがサンプル音声を聞いて PV を持つと判断した、あるいはプロダクションからポップアウトボイスを持つと推薦された男性 3 名 (M1, M2, M3), 女性 3 名 (F1, F2, F4) を対象にした。加えて、PV を持つ一般の女性話者 1 名 (F3) も対象にした。話者の年齢は 20 代から 50 代に分布している。

2.2 声質と発話資料

以下の発話資料をまず「通る声で」と指示して発話させた後、同じものを「日常用いる声で」と指示して発話させた。通常音声はポップアウトする話者 (F3) の場合、「通る声」として PV を収録し、「日常用いる声」はポップアウトしないよう意識して発声した音声を収録した。

1. 日本語 5 母音
2. 文「今すぐ逃げてください」

3. FW07 [4] の低親密度語 50 語, 高親密度語 50 語
4. 上記 3 を「これが～型」のキャリア文に入れたもの
5. イソップ童話「北風と太陽」の文章

なお、F3 は上記 4 の収録を行っていない。

2.3 収録方法

株式会社キャラのスタジオ (大阪市) にて収録した。音声は通常のコンデンサマイク (RODE, NT2-A) に加え、接話型マイク (DPA, 4166-OL-F-B00-LH) も用いて収録した。同時に、EGG 装置 (Glottal Enterprises, EG2-PCX2) にて EGG データを収録した。NT2-A のデータはスタジオの Pro Tools (Avid) を用いて収録し、接話型マイクと EGG のデータはオーディオインタフェース (Roland, Rubix24) を介して Audacity を用いて PC に収録した。全てのデータを標本周波数 48 kHz, 量子化レベル 24 bit にて保存した。

3 分析方法

3.1 EGG データ

「北風と太陽」の文章音声を対象に分析した。収録された EGG データには低周波成分が重畳されていたため、カットオフ周波数 20 Hz のハイパスフィルタによりその成分を除去した。その後、DECOM 法 [5] により各フレームの声門開放率 (O_q) および基本周波数 (f_o) を求めた。

3.2 音声データ

接話型マイクにより収録した母音/a/を対象にした。WORLD [6] によりのスペクトル包絡を計算し、音声区間の中心部の 10 フレームの平均を求めた。

4 結果と考察

典型例として図 1, 2 に話者 M1, F3 の「北風と太陽」の文章発話時の f_o と O_q の関係を示す。これらの話者の結果は、PV 発声時と通常発声時とで f_o および O_q の範囲が異なることを示唆している。PV 発声時に O_q は減少し、 f_o は上昇する傾向が見られる。

* Changes in glottic and vocal tract acoustic properties due to intra-speaker variation of “pop-out” voices. by KITAMURA, Tatsuya (Konan Univ.), NOTA, Yukiko (NINJAL), SAKAKIBARA, KeNichi (Health Sciences Univ. of Hokkaido), KAWAHARA, Hideki (Wakayama Univ.), and AMANO, Shigeaki (Aichi Shukutoku Univ.)

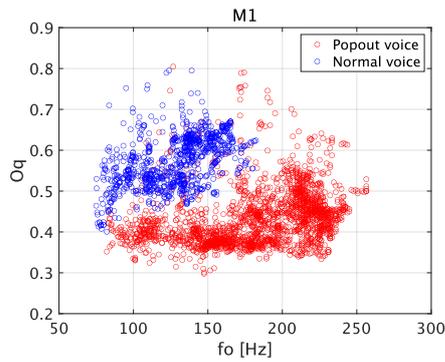


Fig. 1 Voice profile of pitch frequency (f_o) and open quotient (O_q) of speaker M1.

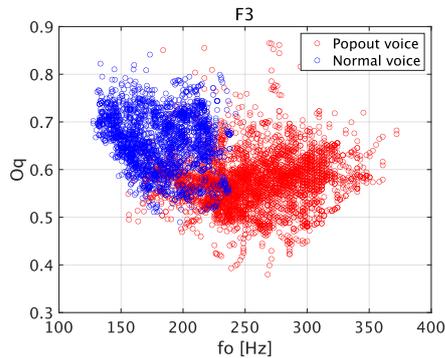


Fig. 2 Voice profile of pitch frequency (f_o) and open quotient (O_q) of speaker F3.

話者ごとに f_o と O_q を独立変数, 声質 (PV/NV) を従属変数とした線形判別分析を行い, 得られた判別式から誤識別率を求めた. あわせて, 各話者の全音声, PV のみ, NV のみの f_o と O_q の相関係数 r_{all} , r_{PV} , r_{NV} を求めた. これらの結果を表 1 に示す. 誤識別率は全体的に低く, f_o と O_q が PV と NV を区別する特徴として有効であることを示している. 一方で, f_o と O_q の間には弱い相関しかない. これらの結果は, PV 発声と NV 発声では f_o と声帯振動パターンが異なることを意味している.

図 3 に話者 M1, F2 の母音 /a/ のスペクトル包絡を示す. これらは PV と NV の差異が大きい例である. 両話者とも PV は NV より音圧が高いためスペクトル包絡全体が上にシフトしているが, それ以外にもスペクトル構造 (フォルマント周波数やスペクトル傾斜を含む) が大きく異なる. これは PV が単なる“大きい声”ではないことを意味し, 声帯振動のみならず調音も NV とは異なることを示唆している.

5 おわりに

本研究では, PV を有するプロのナレーターおよび一般の話者計 7 名を対象に PV と NV の音声と EGG

Table 1 Error rate of linear discriminant analysis and Pearson correlation coefficients between f_o and O_q . r_{all} : PV and NV, r_{PV} : PV, r_{NV} : NV.

Speaker	Error rate	r_{all}	r_{PV}	r_{NV}
M1	2%	-0.210	0.299	0.503
M2	4%	-0.200	0.097	0.355
M3	11%	0.400	0.256	-0.038
F1	13%	-0.030	-0.134	0.110
F2	8%	0.278	0.153	-0.114
F3	4%	-0.416	0.107	-0.161
F4	7%	0.408	0.472	0.386

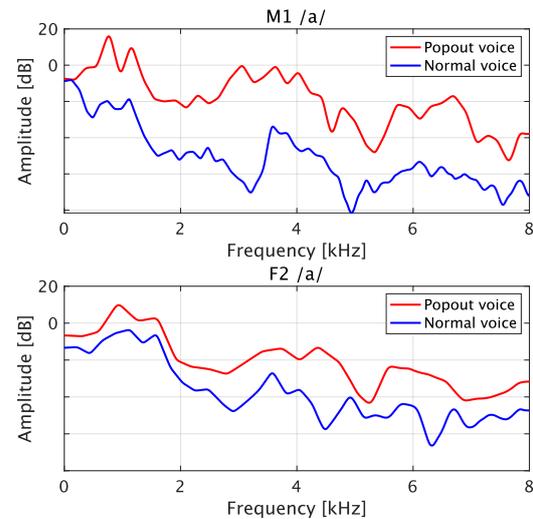


Fig. 3 Spectral envelopes for vowel /a/ of speakers M1 and F2. (Top) M1 and (Bottom) F2.

データを収録, 分析し, PV と NV では声帯振動特性と声道音響特性が異なることを示唆する結果を得た. 今後, MRI を用いた観察により PV と NV の調音上の差異を明らかにする.

謝辞 本研究は JSPS 科研費基盤研究 (A) 「ポップアウト・ボイスの生成・知覚基盤の解明に基づく高性能拡声技術の開発」(JP20H00291) の助成を受けた.

参考文献

- [1] Amano *et al.*, *Acoust. Sci. Tech.*, 43(2), 105–112 (2022).
- [2] 天野ら, 音講論 (秋), 3-9-1 (2022).
- [3] 天野ら, 音講論 (秋), 3-9-2 (2022).
- [4] 近藤ら, 信学技報, 7(432), TL2007-62, 43–48 (2008).
- [5] Henrich *et al.*, *JASA*, 115(3), 1321–1332 (2004).
- [6] Morise *et al.*, *IEICE Trans. Info. & Sys.*, E99-D(7), 1877–1884 (2016).