

創作ノート

「声」の楽器による楽曲制作を通じた、声の違和感に関する研究 Through the music production by the instrument of the “voice”, research on the incongruity of voice

吉村帆生

Fou YOSHIMURA

九州大学大学院芸術工学府

Graduate School of Design, Kyushu University

城一裕

Kazuhiro JO

九州大学芸術工学研究院

Faculty of Design, Kyushu University

概要

本発表では、年齢や性別を知覚させないニュートラルな音声とイントネーションを持つ「声」の楽器と、それをを用いた楽曲の制作およびライブパフォーマンスについてその詳細を述べる。人工的な声の生成における先行事例として、ケンペレンの音声合成装置や明和電機のおタマトーン等を参考に、人工的な構築物である機械が発する声の人間に与える違和感の究明を試みる。

1. はじめに

声に違和感を感じたことはないか。例えば、アニメにおいて本業が声優ではない俳優や素人がキャラクターにあてた声だったり、そもそも声優による所謂「アニメ声」そのものだったり、人によって感じられる場面は多様に存在するだろう。今や世界に羽ばたくバーチャルシンガーである VOCALOID「初音ミク」(クリプトン 2007) もまた、2007 年発売当初、音響エンジニアたちによるニッチな需要で盛り上がりを見せつつも、その機械的な歌声は耳障りで違和感があると非難する声があった (Werde 2003)。つまり、その違和感は慣れとともに消えてしまいがちなのである。そしてそれは最近のことではなく、以前から続く事象でもある。例えば、1887 年に音の記録・再生装置である蓄音機のフォノグラフが登場した時、機械が声を発するという事実聴衆は驚き、興味を示し、違和感を感じた (秋吉 2008)。この感覚は、人工知能を備えたソフトバンクの感情認識ヒューマノイド Pepper や Apple の音声アシスタント Siri といったロボットやスマートフォンから発せられる合成音声が発する身近なものとなった現代において、最早立ち返ることのできないものである。

そこで、本研究では、声の違和感に改めて注目し、慣れきった現状に対して客観的に明示しながら再び投げかけた時にどのような反応になるのかを探る。具体的に

は、前述した両者をつなぐ「声」の楽器の系譜をたどり得られる知見と、自身の学士論文である「声における年齢差と性差の知覚について」(吉村 2017) に基づいて、特定の年齢や性別を知覚させないニュートラルな音声とイントネーションを持つ「声」の楽器を提案する。また、それをを用いた楽曲を制作し、ライブパフォーマンスとして作品を公開することで、さらにそこで得た聞き手の反応を再び楽器の改良や作品へと反映させていく。それにより、聞き手が声の違和感の面白みを再び意識し、適応しながら、やがては慣れてしまうというプロセスを明らかにする。

2. 関連研究

声の装置・楽器の関連研究としては、まず「ケンペレンの音声合成装置」(Kempelen 1791) が挙げられる (Dudley 1950)。これは「人間の口蓋を模した構造を作りだし、手押し式のふいごと管によってそこに空気を送り込み、切切口蓋のなかにとりつけられたフィルターを共鳴させることで、いくつかの母音を生み出した」ものである。「音声を作り出す」ために人間の口腔器官を模したことは、人工音声が発声器官の状態ごと再現するための技術であることを示す原点でもある (Cater 1983; Heudin 2008)。

また、「ボコーダー」(Dudley 1928) は「発話を分析してコンポーネントに分け、音声ではなくコードを電話線経由で送信し、受信側に届いたコードがオリジナルの発話を再現するというテクノロジー」に「コード化されたサンプルを人間の声に近づけて発音するシンセサイザー」が合わさったものである (Dudley 1940)。しかしながら、開発当初はピッチに問題があり、人間の声とは言いがたく、あまりにも気に障る音であった (越智 2017)。

その実験過程で生まれたものに「ヴォーダー」(Dudley 1939) が存在する。それはニューヨーク・クイーンズで



図 2: フレキシブル素材の声道模型

チ制御機能のついた電気式人工咽頭(上見 1995)「ユアトーン³」(DENCOM 2017)を用いる。

3.3. 演奏方法

演奏方法としては、利き手でユアトーンのスライド式スイッチを、逆の手で声道模型を操作する。ユアトーンのモード(会話もしくは歌)や抑揚の有無等はパフォーマンス中でも自由に変更する。また、声道模型はその柔らかさを存分に活かすために、様々な方向から力を加えて変形させる。握り潰したり、絞ったり、と力加減を調節しながら様々な声を作り出していく(図3)。



図 3: 声道模型と人工声帯を用いたパフォーマンス

4. 楽曲

自作楽器の声を十分に活かすため、「うおかりーず(未)」と名付けた楽曲を制作した。これは合成音声のみで構成される電子音響作品である。VOCALOID 鏡音リン・レンの様々な声色で紡がれた「未完成」の土

³ ユアトーン <http://yourtone.jp/> 2019 年 3 月 1 日アクセス

台に、「声」の楽器であるボーカロイドやポケット・ミク、オタマトーン、人工声帯、自作楽器を用いた即興的な演奏が重なることで、その場限りの「完成」された音楽が生まれる。「ヴォカリーズ」とは、もともと、歌い手の技量が試されるクラシック音楽における歌詞のない歌唱法を意味する。この作品では、すべての楽器の声を母音に限定することで、その様子を踏襲している。生きている人の声とは異なる楽器だからこそ、表現と体験の差異が認識され、聞き手が声そのものの知覚に敏感になることを狙った。

土台部分については、Logic Pro X 上の Piapro Studio で組んでおり、鏡音リン・レン混声の最大六声となっている。パフォーマンス時の即興演奏を際立たせるため、その重なりが複雑になりすぎないように、あえて数小節の短い旋律をループさせている。その短い旋律は楽曲の主題でもあり、幅広い層の聞き手にも展開をわかりやすくするため、A-B-A で展開する単純な三部形式で構成した。

5. パフォーマンス

以下に紹介する三公演では、それぞれ別の「うおかりーず(未)」の土台部分を書き下ろした。演奏は主に、楽曲とあわせた三部構成となっている。

5.1. FREQ2018 - 21 世紀初頭の音と音楽⁴

2018 年 3 月 30 日に九州大学大橋キャンパス音響特殊棟でパフォーマンスを発表した。このパフォーマンスは楽器製作の前に行われたため、即興要素にはポケット・ミクとオタマトーンテクノを用いた。一部と二部にポケット・ミクとオタマトーンテクノをそれぞれソロで使用した後、三部では両方を駆使して即興演奏を重ねた。

5.2. つくると！ vol.5⁵

2018 年 7 月 15 日に九州大学大橋キャンパス多次元ホールでパフォーマンスを発表した。即興要素にはオタマトーンテクノと自作楽器であるフレキシブル素材の声道模型を用いた。一部では主にオタマトーンテクノを、二部では声道模型を即興的に演奏し、三部では両方を使用した。

⁴ FREQ2018 - 21 世紀初頭の音と音楽

<http://www.design.kyushu-u.ac.jp/kyushu-u/temp/ce68afa5557217de2b7ca78cc739e7b4.pdf> 2019 年 3 月 1 日アクセス

⁵ つくると！ vol.5 <http://vol5.tsukuruto.net> 2019 年 3 月 1 日アクセス

5.3. COPY CONT ROLL 03⁶

最新のパフォーマンスは、2018年11月29日に薬院IAF SHOPで発表した。即興要素には前回同様、オタマトーンテクノと自作楽器であるフレキシブル素材の声道模型を用いた。一部では主にオタマトーンテクノを、二部と三部では声道模型を即興的に演奏した。それらの楽器に加え、終盤では人工声帯と自身の声を用いた二声によるパフォーマンスも行った。

6. 考察

本章では、公演での発表を通して得られた反応から、楽器、楽曲、パフォーマンスについてそれぞれ考察する。

6.1. 楽器

楽器については、2回目と3回目の公演で、聴衆の声道模型への興味関心の高さが伺えた。後者のパフォーマンス後には、ある観客がTwitterで以下のように反応している。⁷

吉村帆生さんは、最初はオタマトーン、そこから人口咽頭を使った演奏。これは私の語彙では簡単に説明できないぞ…。試み自体がかなり面白かった。母音が出る形になってるやつ（何て言ってたっけ…）もめちゃくちゃ興味深かった。「あ」と「か」の違いを出すには、形以外の要素が必要なんかしら。

最後の文では、前述した全ての「声」の楽器に当てはまる、子音をどう扱うべきかという問題が提起されている。その一方で、母音に特化した声道模型を楽器に活用することは、「ケンペレンの音声合成装置」のように、人間の器官を模して声色の変化を原始的に表現することとなる。それは結果として、聞き手を「機械が声を発する」という事象に客観的に振り返らせることになった。

6.2. 楽曲

楽曲については、パフォーマンス後に観客が土台部分の旋律を鼻歌で歌っていたことも確認されている。このことから、楽曲自体にもある程度のキャッチーさが保持されていることがわかる。それと同時に、実験的な試みであっても、声の「違和感」のみに意識を集

中させず、自然とその声を受け入れる要素として働くキャッチーさが重要であることも実感した。また、それは、三部形式というわかりやすい展開をもつ耳馴染みの良い楽曲であるからこそ、楽器の声の不自然さ、「違和感」を際立たせているとも考えられる。

6.3. パフォーマンス

パフォーマンスについては、3回目の公演で行った人工声帯と自身の声による二声の表現に注目する。これは一人で二声の演奏をするホーミーのような特徴を持っているが、それぞれのパートのピッチが独立して動ける点で、より自由度の高い表現の可能性がある。結果として、一人の演奏者の口から発せられる二声の奇妙な重なりが、改めて、聞き手に声の「違和感」を意識させることとなった。

7. まとめ

本研究を通して、人々が慣れてしまう声の違和感、特に「機械が声を発する」という事象における違和感に注目してきた。

声に対して客観的かつ原始的な表現を可能にする声道模型というツールは、その反応をより良く見極められるものとなった。現状では各母音の模型をそのまま活用しているため、今後は、より滑らかで自由な音声を生み出すような楽器としての形状を改良していく。声道模型を鳴らす際に電気式人工咽頭をそのまま使用している音源部分についても再考する。

聞き手が感じる違和感を解明するにはさらなる追求が必要だが、合成音声特別なものではなくなった現状において、これらの取り組みで新たな表現に挑戦していきたい。

8. 参考文献

- B. Werde, "Could I Get That Song in Elvis, Please?" in *New York Times Online*, 2003. <https://www.nytimes.com/2003/11/23/arts/music-could-i-get-that-song-in-elvis-please.html> 2018年3月1日アクセス
- 秋吉康晴. 2008. 「音響装置論: 19世紀末のフォノグラフの声」『美学芸術学論集』4, 48-64.
- 吉村帆生. 2017. 「声における年齢差と性差の知覚について」学士論文, 京都市立芸術大学.
- H. Dudley & T. H. Tarnoczy, 1950 "The speaking machine of Wolfgang von Kempelen" in *The Journal of the Acoustical Society of America*, 22(2), 151-166.

⁶ COPY CONT ROLL 03 <http://iafshop.tumblr.com/post/179513134603/copy-cont-roll-03> 2019年3月1日アクセス

⁷ emari (emarikko). 2018年11月29日. [Twitter post]. <https://twitter.com/emarikko/status/1068170975231991809> 2019年3月1日アクセス

- John P. Cater 1983. *Electronically Speaking: Computer Speech Generation*. 72-74. Howard M. Sams & Co.
- Heudin, J. C. 2008. *Les Créatures artificielles: Des automates aux mondes virtuels*. Paris: Odile Jacob.
- H. Dudley, 1940 “The vocoder—Electrical recreation of speech” in *Journal of the Society of Motion Picture Engineers*, 34(3), 272-278.
- 越智朝芳. 2017. 「ヴォコーダー論—主体なき声の歌—」 博士論文, 立命館大学.
- H. Dudley, R. R., Riesz, & S. S. A. Watkins, 1939. “A synthetic speaker” in *Journal of the Franklin Institute*, 227(6), 739-764.
- D. Tompkins 2011. *How To Wreck A Nice Beach: The Vocoder from World War II to Hip-Hop*.
- 三輪眞弘, 佐近田展康. 2013. 『兄弟式国際ボタン音素変換標準規格』に至る長くまっすぐな道』『先端芸術音楽創作学会会報』5(2), 13-16.
- 荒井隆行. 2007. 「声道模型」『日本音響学会誌』63(8), 470.
- H. K. Dunn, 1950. “The calculation of vowel resonances, and an electrical vocal tract” in *The Journal of the Acoustical Society of America*, 22(6), 740-753.
- T. Chiba & M. Kajiyama, 1941. *The Vowel ' Its Nature and Structure*, Tokyo-Kaiseikan Pub. Co., Ltd., Tokyo.
- Arai, T. 2001. “The Replication of Chiba and Kajiyama’s Mechanical Models of the Human Vocal Cavity (<Feature Articles> Sixtieth Anniversary of the Publication of The Vowel, Its Nature and Structure by Chiba and Kajiyama).” *Journal of the Phonetic Society of Japan* 5(2), 31-38.
- Maeda, E., Usuki, N., Arai, T., Saika, N., & Murahara, Y. 2004. “Comparing the characteristics of the plate and cylinder type vocal tract models.” *Acoustical Science and Technology* 25(1), 64-65.
- 上見憲弘, 伊福部達, 高橋誠, 松島純一. 1995. 「呼吸圧によるピッチ周波数制御機能のついた人工喉頭の開発」『医用電子と生体工学』33(1), 7-14.
- Nakamura, K., Toda, T., Saruwatari, H., & Shikano, K. 2010. “Evaluation of extremely small sound source signals used in speaking-aid system with statistical voice conversion.” *IEICE TRANSACTIONS on Information and Systems* 93(7), 1909-17.

9. 参考作品

- クリプトン・フューチャー・メディア. 2007. 「初音ミク」 <https://ec.crypton.co.jp/pages/prod/vocaloid/cv01> 2018 年 3 月 1 日アクセス
- Wolfgang von Kempelen. 1791. *Von Kempelen’s talking machine*.
- Homer Dudley. 1928. *Vocoder*.
- Homer Dudley. 1939. *Voder*.
- フォルマント兄弟. 2003. 「兄弟 de ピザ注文」 <https://youtu.be/FFvFlpVjEjM> 2018 年 3 月 1 日アクセス
- YAMAHA. 2004. *VOCALOID* <https://www.vocaloid.com> 2018 年 3 月 1 日アクセス
- 学研. 2014. 「ポケット・ミク」 <http://otonanokagaku.net/nsx39/> 2018 年 3 月 1 日アクセス
- YAMAHA. 2017. *VOCALOID Keyboard* <https://www.vocaloid.com/vocaloidkeyboard/> 2018 年 3 月 1 日アクセス
- 明和電機. 2009. 「オタマトーン」 <https://www.maywadenki.com/otamatone/> 2018 年 3 月 1 日アクセス
- 姜英里. 2018. 「人間に近い声色を持つ声道楽器 -ヴォカリーズ-」『卒業制作優秀作品集2018』<http://www.tamabi.ac.jp/pro/g-works/2018/pd/s5/> 2019 年 2 月 26 日アクセス

10. 著者プロフィール

吉村帆生 (Fou YOSHIMURA)

1995 年福井市生まれ。九州大学芸術工学府修士課程 1 年。「ムラサメ」名義のボカロ P として活動しながら、「声」を楽器としてとらえ、新たな表現へと具現化する様々な試みに挑戦中。

城一裕 (Kazuhiro JO)

1977 年福島県生まれ。東京藝術大学芸術情報センター [AMC] 助教、情報科学芸術大学院大学 [IAMAS] 講師を経て、2016 年 3 月より九州大学芸術工学研究院准教授。山口情報芸術センター [YCAM] 専門委員 (非常勤)。博士 (芸術工学)。専門はメディア・アート。音響学とインタラクションデザインを背景とした現在

の主なプロジェクトには、参加型の音楽の実践である「The SINE WAVE ORCHESTRA」、ありえたかもしれない今をつくり出す「車輪の再発明」、音・文字・グラフィックの関係性を考える「phono/graph」などがある。



この作品は、クリエイティブ・コモンズの表示 - 非営利 - 改変禁止 4.0 国際 ライセンスで提供されています。ライセンスの写しをご覧になるには、<http://creativecommons.org/licenses/by-nc-nd/4.0/> をご覧頂るか、Creative Commons, PO Box 1866, Mountain View, CA 94042, USA までお手紙をお送りください。