

音響変換に基づく音痴の自覚支援環境

Tone-deafness Awareness Support Based on Voice Transformation

橋本 達也, 林 佑樹, 瀬田和久

Tatsuya HASHIMOTO, Yuki HAYASHI, Kazuhisa SETA

大阪府立大学 現代システム科学域 知識情報システム学類

College of Sustainable System Sciences, Osaka Prefecture University

Email: t_hashimoto@ksm.kis.osakafu-u.ac.jp

あらまし：人間の能力を訓練するためには客観視が有効である。しかし、歌唱能力においては歌いながら自分の歌を聴いて評価するのは難しく、また後から自分の歌を聴くことは恥じらいからためらわれる。よって、自分の音痴に自分で気づくことは、きっかけが少なくなり困難である。そこで本研究では、音痴の自覚を促すことを目的として、学習者の歌声を編集することで歌声の他人化を行い、他人化された歌声を学習者自身が聴いて評価することができる環境を開発した。

キーワード：音痴，客観視，振り返り

1. はじめに

人間は運動能力や思考能力，判断能力など多様な能力を持っており，それらの能力を訓練するためには自らを客観視することが有効であることが知られている⁽¹⁾。例えば運動能力では，体の動きを第三者の視点から確認することで，気づいていなかった悪い癖を認識し，パフォーマンスを向上することができる。しかし，人間は能力を使うこととその能力を評価することを同時に行うことは難しい。とりわけ歌唱能力においては，骨導音を聞くことによる自己評価と他者評価が異なることも少なくなく，歌いながら自身の歌声を評価することは難しい。また，録音するなどして歌い終わった後に改めて自分の歌声を聴こうとすることは恥じらいから躊躇してしまうこともある。これらの要因から，自分の音痴に自分で気づくことは難しく，きっかけも少ないと言える。

そこで本研究では，音響変換を用いて学習者の音痴自覚を支援する環境を開発した。

2. アプローチ

歌唱能力を客観視する方法として録音した音声聴く他に，他人に聞かせて指摘してもらうという手法がある。しかし，他人から自分が音痴であることを指摘されたとしても，骨導音と気導音による音の聞こえ方が違うことから得られる実感は薄い。そこで本研究では，歌唱能力を客観視する機会を与える方法として，自分の歌声を音声変換することを考える。この方法では，学習者の歌声を他人の歌声のように聞こえるよう編集することにより，自分の歌を客観視することを促す。

この時，歌声の音程と声質を独立して編集する。音程を変えることで声の高低を変化させる。また，声質を変えることで高低とは別に男性の声や女性の声のように聞こえるように変化させる。これによって，編集後の歌声を別人の歌声のように聞こえるよ

うにする。さらにアニメーションエージェントに歌わせるというアプローチをとることで，他者が歌っているという印象を強くし，歌声の客観視を補助することを狙う。

学習者は，編集した音声データを入力としてエージェントが歌う映像を視聴し，評価する。この一連の流れによって，自分の歌声を客観視しながら評価することができる期待できる。

3. 音痴の自覚支援システム

本研究では，学習者の歌声を他人の歌声のように編集し，編集した音声データを学習者自身に評価させることで，音痴の自覚を支援する環境を構築した。システムは Unity⁽²⁾のデスクトップアプリケーションとして実装している。

図1にシステム画面を，図2にシステム構成図を示す。図1中央部に写っているエージェントが編集された歌声に合わせてリップシンクする。



図1 音痴自覚支援システムの画面

図2にあるように，このシステムは利用前に前準備が必要である。以下，順を追って説明する。

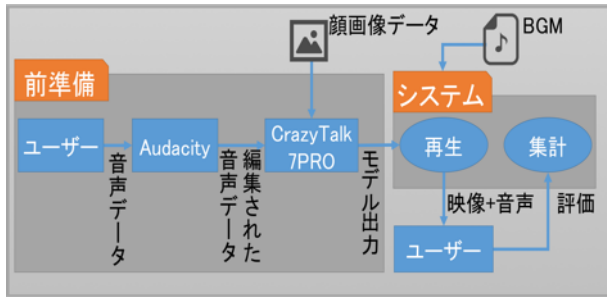


図2 システム構成図

3.1 前処理

システム利用前に、学習者には指定した課題曲を歌ってもらい、Audacityを用いて録音する。音声のみを録音するために、学習者はヘッドホンやイヤホンを用いて曲を聴きながら歌う。さらに録音には、物音や雑音が入らないようにする。これらは歌声以外の要素で自分の歌声であることを認識してしまうことを避けるためである。

次に、録音した歌声の音程(ピッチ)と声質(フォルマント)をAudacity⁽³⁾とRoVee⁽⁴⁾を用いて、独立して編集する。このとき、BGMと歌声に対する音程処理の内容を同じものにするによって、BGMと歌声の間に生じうる違和感を極力少なくする。

編集後、編集した歌声と顔画像データそれぞれに対応したリップシンクデータと、リップシンクを行うモデルを、Crazytalk7 PRO⁽⁵⁾を用いて出力する。

3.2 システム運用

システム内で編集した学習者の歌声と、これをリップシンクしたエージェントの動画を同時に再生し、学習者に視聴させる。そして、学習者は視聴を終えた後に歌声を評価する。評価はマウスクリックを用いて行う。歌声に対して音痴と思ったタイミングでクリックし、学習者がどの部分で違和感を覚えたのか記録する。記録したものは学習者が確認できるようにする。

4. 初期実験

4.1 実験設定

本研究で開発したシステムを運用する前に、音声編集が有効であるかを確認するために、編集した歌声と未編集の歌声を用いて初期実験を行った。

手順は以下の通りである。

1. 「編集後の歌声」を誰の歌かを明かさずに視聴、評価させる。
2. 「編集前の歌声」を誰の歌かを明かさずに視聴、評価させる。
3. 歌い手を学習者に明かす。

本研究では、1,2の評価を比較した。今回の初期実験ではA, B, Cの3名の学習者に音声を録音してもらい、そのうちAとBの2名の学習者に3名分の

音声を評価してもらった。質問内容は「歌声全体への五段階評価」と「自分の歌声がどれか」、「どの歌が同じ人物のものか」の3つである。

歌を5段階評価	学習者A	学習者B
音声1 (編集済A)	3	3
音声2 (編集済C)	2	4
音声3 (編集済B)	4	3
音声4 (未編集C)	3	2
音声5 (未編集B)	4	3
音声6 (未編集A)	3	4

* 矢印は評価の遷移

* 赤字は自分の音声だと気付いたことを表す

図3 実験結果

4.2 実験結果

図3に実験結果を示す。「どの歌が同じ人物のものか」という質問では、音声1~6のうち、同じと考えた番号のペアを作ってもらった。その結果学習者A, Bの両方が全問正解であった。「自分の歌声がどれか」という質問には、自分の歌声だと思った歌の番号を全て書いてもらった。結果、図3にある通り、学習者A, Bの両方が編集前と編集後の自分の歌声をほぼ認識していた。歌声の評価については、自分の歌声に対しては中間の3にしていた。しかし、自分以外の歌声に関しては編集前後で評価に違いがあった。

実験後の聞き取り調査では、環境音によって自分の歌だと認識できたという意見が挙げられた。これらのことから、歌以外の要素で自分の歌かどうか判別できないように音声データの録音・編集を工夫していくことで、歌唱能力に対する評価を変えることができる可能性があると考えた。

5. 今後の課題

音痴自覚の支援を目指す本システムにおいて、その後に音痴を改善する方向に意識を向けるかは、学習者に委ねられている。今後の課題として、学習者が音痴を自覚した後、自身の音痴を改善する意欲が湧くような仕組みを実装したいと考えている。

また、このシステムは手作業を必要とする部分が多く、負担が大きい作業負担を低減する仕組みを考えていきたい。

さらに、歌以外の要素による判別を防ぐための音声データの録音・編集に注力したい。

参考文献

- (1) 古川ら. “身体知研究の潮流—身体知の解明に向けて—” 人工知能学会論文誌, 20(2), 117-128, 2005.
- (2) Unity: <https://unity3d.com/jp>
- (3) Audacity: <https://www.audacityteam.org>
- (4) RoVee1.21: <https://www.g200kg.com/jp/software/rovee.html>
- (5) Crazytalk7PRO: <https://www.ah-soft.com/ct/>