

うたごえ うつ え こうしゅうみさかすいめん しら  
歌声の映し絵: —《甲州三坂水面》の調べ —

(独) 大学入試センター 研究開発部 内田照久  
e-mail: [uchida@rd.dnc.ac.jp](mailto:uchida@rd.dnc.ac.jp)

### 錯覚の説明:

本作品のはじめに提示される歌声は、原曲の歌の節回しの高低を反転させて生成した人工的な歌声です。しかし、

- (1) その旋律が、有名な原曲の歌声を高低反転させたものであることには、まず気がつきません。
- (2) その上で、これまでに聞いたことのない「未知の旋律」が知覚されます。それは、人工的な節回しのはずなのに、音楽的な旋律性を感じます。
- (3) さらに、その反転歌声にビブラートなどの音楽的な技巧さえ感じます。

認知上の過誤という意味では、上記の(1)～(2)が錯覚となります。(3)については、派生的な効果かも知れません。

### 現象の背景:

本作品は、ヒトの声が担っている役割の中から、“歌声”にフォーカスを絞って、“音楽の文脈での錯聴”とも呼べる現象を扱ったものです。

ここでは、歌のメロディの骨格を形作っている声の高さの変化パターン、基本周波数 (fundamental frequency;  $f_0$ ) の軌跡パターンを扱っています。この  $f_0$  パターンは、楽譜に記されているメロディの進行をしっかりと色濃く反映しています。さらに、それと同時に歌声ならではの特徴、音程の跳躍時のターゲット音高への渡り部分での微細な変化をはじめ、ビブラートやシャクリといった揺らぎまで重畳しています。

この作品では、このような纖細な情報を内包する原歌声の  $f_0$  パターンを、その形状を保ったまま、対数周波数軸上で、問答無用で高低反転させています。かなり大胆な変形なので、“音楽的に壊滅的に破綻”してしまって、“世にも恐ろしい音痴”を、この世に生誕させてしまうのかと思いきや、思ひがけず、新たな旋律として音楽的に成立し、歌声としても許容できるものになりました。

このこと自体、たいへん興味深い現象です。しかし、ヒトの知覚の面から考え直してみると、さらに謎が深まります。本作品での  $f_0$  高低反転歌声は、その歌のメロディの高低変化を、文字通り、単純に逆さまにしただけです。また、元の楽曲は誰しも馴染みがあって、よく知っている唱歌です。しかし、この反転歌声を聞いたとき、元の歌の旋律を逆さまにしたメロディだということに気づくことは、相當に難しいようです。図形であれば、高低（上下）反転の図形を想起することは容易です。しかし、変換方法の内容を理解した上で反転歌声を作っている本人でさえ、事前に出来上がりの旋律を想像するのは容易ではありませんでした。楽音のように、規則的な構造をもっていて、その要素が明確であっても、その想起は難しいようです。

また、歌ならではの技巧的な音高遷移を持つ  $f_0$  パターンを反転させた場合、“歌らしさ”が失われてしまって、不自然さが極まる懸念もありました。しかし、それも杞憂に終わりました。なお、日本語音声の時間方向の逆転再生音を聞くと、未知の外国語か、宇宙人の会話か、と感じてしまうほどの違和感が生じます。声や歌声を、時間方向で反転させると、言葉や詞の意味を追うことは、全くできなくなります。しかし、 $f_0$  反転歌声では、多少の不自然さはありますが、概ね、歌声の範疇に留まっています。

これら的心象は、 $f_0$  反転歌声を聴いたヒトの知覚の仕組みが生み出していると言えるでしょう。音楽を聴いている時、私たちは、メロディ、音高の時間的な変化そのものを深く味わっているはずです。それにも関わらず、その旋律が逆さまになっていても、全く気づかないというのは、やっぱり、いかにも、とんでもなく不思議です。

## 作品の解説：

作品のフォーマットについては、前作『お国なまりのエッセンス（2024）』の出だしの展開を踏襲しています。その上で、本作品のオーディエンスは、作品で使用する唱歌については、従前から馴染みがあって既知であるという前提で進めています。ですから、作品の冒頭、歌詞を示した時点で、童謡の“桃太郎”を想起すると仮定しています。そこに、いきなり  $f_0$  反転歌声を提示することになります。オーディエンスには、想定外のメロディが聞こえてくる訳ですが、その時点では、その元々が 童謡の“桃太郎”であることに気がつきません（気づかないはず…です）。

その後段に,  $f_0$ 反転歌声が, 元々の正調 “桃太郎”の逆さま歌声であることを説明していきます。そして, それが “桃太郎”の唄だけに起こることではなく, ある程度一般性をもって生じる現象であることが帰納的に理解されるように, 2 つの楽曲による例示を行いました。

さらに, その説明の中で, 歌とは違って, 図の高低反転の想起は容易だとする具体例として, “逆さ富士”的イメージを借用しています。そして, この“逆さ富士”的イメージは, 最終段の展開にも繋がっていきます。

歌声では, 元々の歌声と  $f_0$  反転歌声は違う旋律に聴こえます。一方, “逆さ富士”では, 湖面に映る姿は相似形の鏡映像を想像するところです。ところが, “富嶽三十六景《甲州三坂水面》”では, 何と一つの図版に, 夏の富士山と, 冬の富士山を, 同時に存立させて描くというギミック, 巧妙な仕掛けが施されていました。元の歌声と  $f_0$  反転歌声の 心象の関係を絵図で表現するとなれば, まさにこの図版だと深く感じ入ったところです。そこで, この図版を提示すると共に, 作品の副題にも織り込みました。

ただ, ここで作品を終えてしまうと, 図版の力に圧倒されたまま終わってしまいます。そこで, 最後の一ひねりとして, “げんこつ山のたぬきさん”的  $f_0$  反転歌声を提示することで, 「本作品は歌声を扱ったものですよ, 忘れないで下さいね。」の注意喚起をさせていただきました。

### 学術的意義:

本作品の  $f_0$  高低反転歌声の生成にあたっては, 見境なく, ただの当てずっぽうで作っていった訳ではありません。多少は勝機を伺いつつ, 一応, それなりの戦略をもって取り組みました。

本作品の  $f_0$  高低反転にあたっては, 対数周波数軸上で反転を行っています。すなわち, *log scale* 上で操作をしています。これまでの応募作品では, 主に言語音声を使用してきました。その中の『声色の罠 (2016)』, 『桃源郷のお国なまり (2019)』, 及び, 『お国なまりのエッセンス (2024)』では, 声の高さやイントネーションを操作しています。そこでは, ヒトが同じと感じる「音の高さの差」が, どの音域でも同じ値になるようにして作られた *mel scale* (1 式) を使って, 聴感上の品質向上に努めてきました。

$$m_0 = 2595 \log_{10} (1 + f_0 / 700) \quad (1)$$

しかし、今回の作品では、音楽として歌声を扱うので、敢えて対数軸 *log scale* を適用しています。それによって、楽音の基本であるオクターブ等価性を担保しています。すなわち、オクターブごとの音高、55 Hz (A1), 110 Hz (A2), 220 Hz (A3), 440 Hz (A4), 880 Hz (A5), 1760 Hz (A6) は、対数軸上では、いずれも等間隔に布置します。さらに、オクターブ内での小目盛にあたる 12 平均律 ( $f = a \cdot 2^{k/12}$   $k = 1 \cdots 12$ ,  $a$  は任意の初期値) の各音も等間隔に並びます。

この *log scale* 上で、楽音のメロディを、適切な中心軸を用いて高低反転させると、反転後の音は 必ず平均律上の音に着地することになります。つまり、反転後も楽譜に記載できる音になります。元になる曲は、調性の制約で構成音が限られますから、反転した後の音列も音楽的な曲になる公算が高いと想定しました。

なお、作曲や編曲の技法として、楽譜上でメロディの推移を鏡映像のように高低反転させる逆行進行が知られています。発想としては、本作品のアイデアと共通する点があります。しかし、その場合、元の主題の調性での音階音の使用に限るなど、楽譜上の見た目は鏡映的ですが、必ずしも同じような旋律にはならないようです。

さて、歌声の特徴としては、音程跳躍時のターゲット音高への渡り部分には、滑らかな音高の変化の直後に、目的の音高を越える瞬時的な変動成分が見られます。これは、オーバーシュート (Overshoot) と呼ばれており、話声と歌声の聴感的な違いを規定するだけでなく、人間らしい歌声として知覚する上で、重要な役割を果たすとされています(齋藤, 2015)。

これは旋律が上行する場合に、歌声が滑らかに高くなっていく時、一瞬、ターゲットの音高よりも高く上がってから、ターゲットの音高に下がって戻る現象をさします。逆に、下行の場合は、低く下がる時に、ターゲットの音高よりも低く下がってから、上がって戻すという振る舞いになります。

これは、歌声らしさを知覚させる特徴でもあるので、反転させた場合に、どうなってしまうかと危惧しました。しかし、よく考えてみると、上行時は上に行き過ぎて戻る、下行時は下に行き過ぎて戻る、という訳なのですから、上行・下行が反転した場合には、ちょうどいい具合に反対方向に行き過ぎて戻る、という振る舞いになります。むしろ、渡りに舟ではないかと考えて、特段の対処を講ずることはせず、当初の単純反転で進めることにしました。

このようにして、ある程度のアタリをつけて、 $f_0$  反転歌声を生成しました。このような手続きで生成された歌声について、内田(2025)は聴覚実験で評価を行っています。その結果、 $f_0$  反転歌声は、“歌声らしい”，“音楽的で旋律的である”，その上で，“知らない旋律”だ、と評価されていました。

この  $f_0$  反転歌声が、音楽的であり得る限界点については、まだ詳細は吟味していません。元の旋律の中に、音階の構成音以外の経過音などがある場合、反転歌声は音楽的に破綻しないか？ 転調するメロディの場合、反転歌声でも自然な転調になるのか？ ハーモニーや和声が付与されている場合、そのまま反転させたらどうなるのか？ これらについては、楽譜上の音高の並びを、対数軸上で計算して反転させてから楽譜に戻せば、音楽学的にも評価できるはずだと思われます。

この現象を追うことは、単に歌声の  $f_0$  パターンが伝える情報に留まらず、私たちが“音楽”を聞く際の、内面に備えている情報の再構築のありようを知る手がかりになるかも知れません。これはまた、今後とも楽しみです。

## 関連文献：

- Morise, M. (2016). D4C, a band-aperiodicity estimator for high-quality speech synthesis. *Speech Communication*, **84**, 57-65.
- 森勢将雅 (2018). 音声分析合成 日本音響学会(編) 音響テクノロジーシリーズ 22 コロナ社
- Morise, M., Yokomori, F., & Ozawa, K. (2016) WORLD: a vocoder-based high-quality speech synthesis system for real-time applications. *IEICE transactions on information and systems*, **E99-D** (7), 1877-1884.
- O'Shaughnessy, (1999). *Speech Communication: Human and Machine*. Wiley-IEEE Press, New York.
- Saitou, T., Unoki, M. & Akagi, M. (2005). Development of an F0 control model based on F0 dynamic characteristics for singing-voice synthesis, *Speech Communication*, **46**, 405-417.
- 齋藤 育 (2015). 人間らしい歌声の理解と合成 日本音響学会 2015 年春季 研究発表会 講演論文集, 1387-1389.
- Tamaru, H., Takamichi. S., Tanji, N. & Saruwatari, H. (2020). JVS-MuSiC: free Japanese multispeaker singing-voice corpus, *arXiv preprint*, arXiv:2001.07044.
- 内田照久 (2025). 「歌う雑音」: 多層動的バンドノイズ音声と音楽コミュニケーション 日本音響学会 第 154 回(2025 年秋季) 研究発表会 講演論文集, 1109-1110.