



選択的電気刺激によるビブラートの速さ制御

内田大悟^{1),2)}, 松原晟都²⁾, 村井昭彦^{1),2)}

- 1) 東京大学 新領域創成科学研究科 (〒277-0882 千葉県柏市柏の葉 5-1-5, uchida-daigo909@g.ecc.u-tokyo.ac.jp)
2) 産業技術総合研究所 人間拡張研究センター (〒277-0882 千葉県柏市柏の葉 6-2-3)

概要: ビブラートは歌唱や楽器演奏の印象評価において重要な役割を果たしている。歌手本人の歌唱に似せたり、好印象なビブラートにしたりするためにビブラートの速さを変えることが有効だと考えられるが、そもそもビブラートの実現自体に高いスキルが求められる。そこで本研究では、喉頭部に対して選択的に電気刺激を行うことでビブラートをかけ、その速さの制御を行った。結果、自然なビブラートとして認識される 4-7 Hz の範囲を含む、4-14 Hz の帯域において刺激周波数に対応した速さのビブラートの制御を実現した。

キーワード: 筋電気刺激, ビブラート, ビブラートの速さの制御, 発声機能介入

1. はじめに

ビブラートとは基本周波数 (F_0) の周期的な変動である。ビブラートは速さと変調幅の二つのパラメータで表される。速さは F_0 の単位時間あたりに生じる揺らぎの回数で、変調幅はビブラート区間の平均音高を中心とした変動の幅のことである。一般に、ビブラートの速さは 4-7 Hz 程度、変調幅は 0.5-2.0 半音 (50-200 cents) の範囲が自然だと感じられる [1]。

ビブラートをかけることで歌唱や楽器演奏の表現の幅が広がる。適切な速さのビブラートをかけることで高い歌唱評価を得られる [2]。また、速すぎたり遅すぎたりするビブラートにより、緊張感や活力の低い印象など、多様な印象を与えられる [3]。しかし、ビブラートの実現自体に高いスキルが求められ、速さの制御は非常に困難である。

表面電極による頸部電気刺激によってビブラートがかかることは示唆されているが、定性評価にとどまっており、電極の大きさや位置に課題がある。Fushimi らは、腹部や頸部を表面電極により電気刺激することで、ビブラートがかかることを報告した [4]。しかし、ユーザインタフェースの構築や被験者の主観評価に重きを置いており、音声波形の評価はなされなかった。また、頸部の刺激では数十mm各の電極が輪状甲状筋 (CT 筋) を覆うよう、頸部前面の上部と下部に斜め方向に配置された。この位置に配置したのは、CT 筋を収縮させることで声帯を伸ばし、音声の F_0 を周期的に上げることを期待したためである。しかし、CT 筋が軟骨から露出している隙間は最大でも 7mm であるため [5]、数十mm各の電極で CT 筋を覆うように配置する場合、精確な刺激が難しい。



図 1 多点筋電気刺激装置。10 mm 間隔で 8×8 個並ぶ電極を用いた。

そこで本研究では、CT 筋周辺の狭い範囲に選択的に電気刺激することでビブラートをかけ、その速さを制御できることを定量的に示した。具体的には本実験では、多点筋電気刺激システムによってビブラートをかけるのに適切な刺激位置を特定し、刺激周波数を変えることでビブラートの速さを制御した。そして、録音した音声データに対してビブラートの速さと変調幅を求め、その制御ができることを評価した。

2. 実験: 周波数変更によるビブラートの速さ制御

2.1 実験システム

開発した多点筋電気刺激システムと刺激位置について説明する。図 1 のように多点筋電気刺激システムとして、定電流回路とアナログスイッチを用いた電極選択回路からなる装置を開発した [6]。電極は 2 mm×3 mm の楕円形電極が中心間距離 10 mm で 8×8 個、計 64 個並んだものを用いた。そのうち最上行および最下行を除く 6×8 の範囲を左右に二分し、縦 2 個×横 2 個を 1 つのマス目とした 3×2 のグリッドを定義して、左右のグリッドがそれぞれ

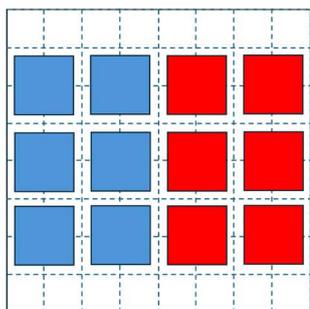


図2 探索する位置のグリッド．赤が陽極，青が陰極を表す．列が隣接しない組を探索した．

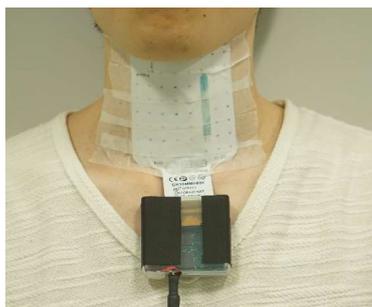


図3 刺激装置を装着した様子．電極は上からテープで固定した．



図4 実験環境．被験者は壁に頭と背をつけた．

陽極，陰極になり，かつ電極の列が隣接しない計 27 通りの組を刺激位置とした（図2）．

2.2 実験手順

本実験は，精確な位置に対し周波数を変更して刺激することで，ビブラートの速さを制御することを目的とする．具体的な実験方法としては，様々な刺激位置に対し電流値を徐々に上げていくことで，刺激に慣れさせつつビブラートがかかりやすい刺激位置を同定し，2 Hz 刻みの 2-20 Hz の周波数で発声中に刺激を行った．

実験は女性 2 名に対して行った．電極は図3のように装着した．刺激は二相性の矩形波でパルス幅 1,000 μs とし，電流値や刺激周波数，刺激位置は実験中に定めた．図4のような実験環境で，無指向性マイク（ECM8000）によって 48 kHz でサンプリングした．発声する音声は/a/とし，マイクと口の距離，方向が試行ごとに変わらないよう，背もたれに背と頭をつけるようにした．また，リラックスして，普段話すような声の大きさ，高さで発声するよう指示した．

詳細な実験手順を示す．まず，実験概要の説明と電極の装着を行った．実験の説明として，大まかな実験の流れに触れ，実際に電気刺激によって喉が震えている動画および

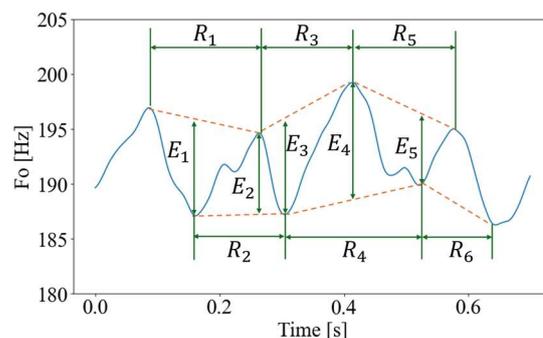


図5 ビブラートパラメータの求め方． R_n の平均が速さ， E_n の平均が変調幅に対応する．

そのときの音声を提示した．電極の装着段階では，装着の手順を示した動画を見せ，ウェットティッシュで頸部を拭いた後，実験協力者自身が鏡を見ながら電極を頸部に装着した．電極が首の中央に，電極の下端が鎖骨のあたりに来るよう指示した．

次に，ビブラートが効果的にかかる刺激位置探索の準備として，電流閾値の設定と発声練習を行った．電流閾値の設定では，電極間距離が最も離れた組に対し，段階的に電流値を上げることで，不快感を生じる電流値を求めた．痛みがないのであれば喉の震えを起こしても更に電流値を上げられること，6 mA 程度まで電流値を上げる人もいることを伝えた．発声練習では 4 秒間/a/を 2 回発声した．

次に，ビブラートが効果的にかかる刺激位置の探索を行った．発声中 1 秒ごとに刺激位置を変えて刺激した．電流値は電流閾値の 50 %から始め，全刺激位置を探索するごとに 20 ポイントずつ上げた．各刺激位置に対する音声データを後述する方法で解析し，ビブラートのかかりやすい刺激位置を求め，以降の実験で用いた．

次に，各刺激周波数における電流閾値の設定および発声練習を行った．電流閾値の設定として，2-20 Hz の 2 Hz 刻みの各刺激周波数において不快感を生じる電流値を求めた．発声練習では，発声してから 1 秒後に 6 Hz の刺激を 3 秒印加することを 2 回行った．

その後，刺激周波数を変更することでビブラートの速さを制御する実験を行った．具体的には，2-20 Hz の範囲の 2 Hz 刻みの周波数で，発声から 1 秒後に 3 秒間刺激した．電流値は各刺激周波数に対して求めた電流閾値とした． F_0 が試行ごとに変わらないようにするため，1 回目と 6 回目の試行の前に非刺激時の発声練習で録音した音声を提示し，その音声と同じ高さで発声するよう指示した．

2.3 解析方法

刺激位置の探索では，各刺激位置に対する 1 秒間の音声から Harvest [7]によって 1 ms ごとに F_0 を抽出し，ハニング窓を適用して FFT を行った．全周波数帯の振幅スペクトルに対する刺激周波数 (6 Hz) の振幅スペクトルの割合が 0.06 以上となる刺激位置をビブラートのかかりやすい位置として以降の実験で用いた．複数の刺激位置で 0.06 以上となった場合は，その中で最大のものを用いた．

周波数を変更してビブラートの速さを制御する実験で

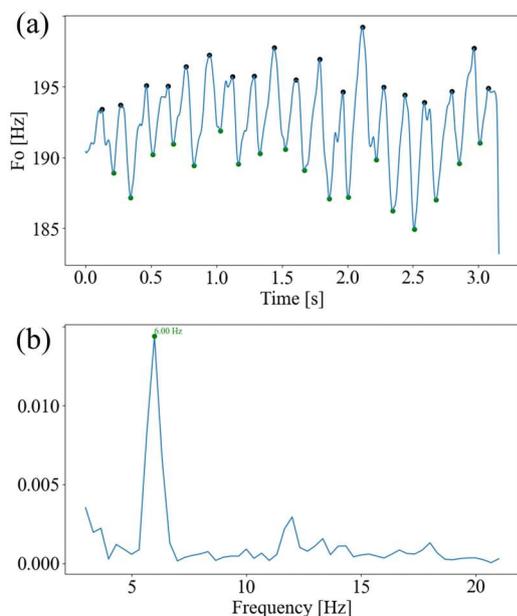


図6 被験者2の6 Hz 刺激時の(a) F_0 波形と(b) 音声スペクトル。 F_0 波形の黒点, 緑点がそれぞれ検出された山, 谷に対応する。

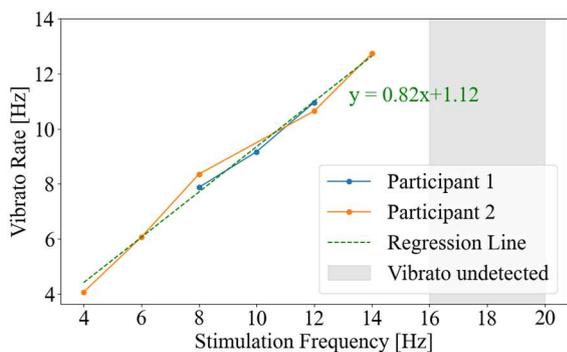


図7 刺激周波数に対するビブラートの速さ。青線が被験者1, 橙線が被験者2の解析結果, 緑線がそれらの線形回帰直線。

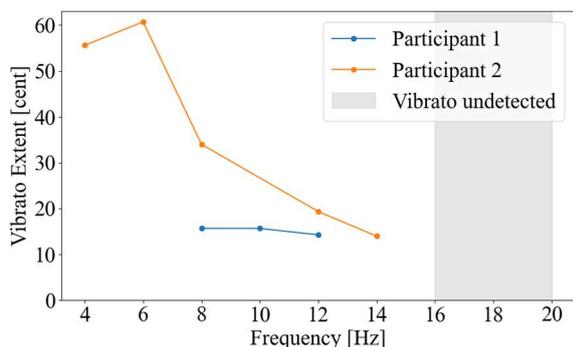


図8 刺激周波数に対するビブラートの変調幅。被験者2で4, 6 Hz 刺激時に50 centsを超えた。

は, まず録音した3秒間の刺激中の音声に対して Harvest [7]を用いて1msごとに F_0 を求め, その振幅スペクトルの割合を算出した。次に振幅スペクトルに対し, 3-21 Hzの範囲でピークとなる周波数を求めた。求めたピーク周波数が刺激周波数と1 Hz以内であれば, 刺激によるビブラートがかかったと見なした。ビブラートが検出された音声の

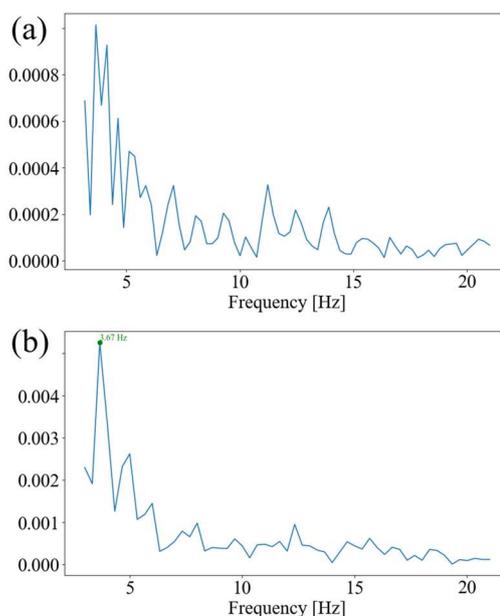


図9 被験者1の(a) 非刺激時および(b) 6 Hz 刺激時の音声スペクトル

F_0 波形に対し, Nakano らの方法 [2]を用いて図5のように R_n , E_n を求め, その平均からビブラートの速さ, 変調幅をそれぞれ算出した。3-21 Hzの範囲でピークを求めたのは, 元の音声に低周波ドリフト成分 (~3 Hz)があったため, 刺激によって低周波のビブラートがかかったとしても, 元の音声のスペクトルに埋もれてしまい, 刺激によってビブラートがかかったとは言い難いためである。ピーク検出の条件は, 一次差分の符号が正から負に変わる点で, ピークの左右2 Hz以内にある最小値のうち大きい方との差が0.001以上であることとした。

2.4 実験結果

それぞれの実験協力者が設定した電流閾値について述べる。実験協力者1は, 刺激周波数が4 Hzのときに最大電流閾値である3.56 mAを設定した。全刺激周波数帯における電流閾値の平均は3.32 mAであった。実験協力者2は, 刺激周波数が4 Hzのときに最大電流閾値である3.80 mAを設定した。全刺激周波数帯における電流閾値の平均は3.14 mAであった。

実験協力者1で10 Hz刺激時, マイクノイズが入ったため, その部分を除く最長区間の1.43秒の音声データに対して解析を行った。2 Hz刻みの4-20 Hzで刺激した音声において, 実験協力者1では8, 10, 12 Hz, 実験協力者2では4, 6, 8, 12, 14 Hzでビブラートパラメータが算出された。例として, 実験協力者2において6 Hzで刺激したときの音声データを解析した結果を図6に示す。図6(a)は F_0 波形で, 黒点と緑点がそれぞれ検出された山, 谷の位置を表す。図6(b)は F_0 のスペクトルであり, ピーク周波数は6.00 Hzと, 刺激周波数である6 Hzと等しい速さでビブラートがかかっていた。

図7にビブラートの速さを示す。刺激周波数に対するビブラートの速さの線形回帰係数は0.82, 切片は1.12となった。係数が1, 切片が0に近いことから, 4-14 Hzの

範囲で刺激周波数とビブラートの速さが概ね対応していた。両実験協力者とも 16 Hz 以上の刺激周波数で、対応する周波数の F_0 スペクトルピークは検出されなかった。つまり、ビブラートは観測されなかった。

図 8 にビブラートの変調幅を示す。実験協力者 1 では 8-12 Hz で約 15 cents の変調幅となった。実験協力者 2 では 4, 6 Hz で 50 cents 以上の値をとり、8 Hz 以上の刺激周波数で変調幅は減少傾向にあった。

2.5 考察

4-14 Hz の範囲で、ビブラートの速さはおおむね刺激周波数と対応していた。そのため、4-14 Hz の範囲で電気刺激することで、ビブラートを刺激周波数と対応する速さで制御できることが示唆された。これは刺激周波数が比較的小さいため、喉頭筋 (CT 筋, TA 筋など) が刺激の間に弛緩の見られる単収縮を繰り返したためであると考えられる。例えば、喉頭筋の活動によって声帯が伸長すると、刺激による収縮期には F_0 が大きく、非刺激時の弛緩期には F_0 が小さくなり、 F_0 の揺らぎが観測される。一方、両実験協力者とも 16 Hz 以上の刺激周波数ではビブラートが検出されなかった。これは刺激周波数が比較的大きいため、喉頭筋が刺激の間に弛緩の見られない強縮をしたためであると考えられる。例えば、喉頭筋の強縮によって声帯が伸長し続けると、非刺激時よりも大きく一定な F_0 の音声となる。

実験協力者 1 で 4, 6 Hz の低周波帯のビブラートが検出されなかったのは、元の音声で 4 Hz 程度の揺らぎを持っており、刺激による周波数ピークが埋もれたためであると考えられる。実際、図 9(a) に示すように実験協力者 1 の声は非刺激時にも 3-6 Hz 付近のスペクトルの割合が大きく、図 9(b) に示す 6 Hz 刺激時のスペクトルでは 6 Hz 付近に顕著なピークが見られない。

ビブラートの速さは 4-7 Hz、変調幅は 50-200 cents の範囲で適切だと感じられやすい [1]。実験協力者 1 では 8-12 Hz の範囲で刺激周波数に対応した速さのビブラートが検出され、変調幅はおおよそ 15 cents だった。実験協力者 2 では 4, 6 Hz で検出されたビブラートの変調幅は 50 cents 以上だった。そのため、実験協力者 1 の 8-12 Hz 刺激時の音声は弱く速いビブラート、実験協力者 2 の 4, 6 Hz 刺激時の音声は適切なビブラートがかかっていると認識されることが考えられる。

2.6 本研究の限界

本実験は 2 名の女性実験協力者に対し、各刺激周波数で 1 回ずつ音声を録音した。実験協力者数が少なく、実験回数も限られていることから、データの再現性やばらつきを評価することが難しい。また、女性に対してのみの実験のため、性別による影響を評価できない。今後の研究では、実験協力者数、刺激周波数ごとの実験回数を増やし、実験協力者に性別の偏りがないようにする。なお、予備実験では男性 3 名に対し、4-14 Hz の帯域でビブラートの速さを

制御できた。

実験協力者 1 ではビブラートの変調幅が 50 cents 以下と適切なビブラート変調幅の範囲外だった [1]。また、予備実験では実験協力者 1 名において、喉が震えるのに十分でない電流レベルでも不快感を生じ、ビブラートが全く検出されなかった。実験協力者によって、喉が十分に震える電流値と不快感を生じ始める電流値が異なるため、電気刺激によるビブラートがかからなかったり、変調幅が小さく弱いビブラートになったりする。

3. おわりに

本研究では、頸部を多点筋電気刺激システムによって選択的に電気刺激することによって喉頭筋を選択的に駆動し、その刺激周波数を変えることでビブラートの速さの制御を行った。結果、認知的に適切なビブラートの速さの範囲を含む周波数範囲で、ビブラートの速さの制御を実現した。将来的には、本研究で提案する手法を用いることで、誰でも歌手や歌の種類によって異なる速さのビブラートをかけ、表現力を高められることが期待される。

謝辞 本研究は、JST, ACT-X, JPMJAX210N 及び、JSPS, 科研費 (研究活動スタート支援), JK23K20018 の助成の結果得られたものである。また、本研究は東京大学倫理審査専門委員会の承認 (審査番号 23-350) を受けて実施された。

参考文献

- [1] Y. Horii, "Acoustic analysis of vocal vibrato: A theoretical interpretation of data," *Journal of Voice*, vol. 3, no. 1, pp. 36-43, 1989.
- [2] T. Nakano, M. Goto, and Y. Hiraga, "An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features," *INTERSPEECH 2006*, pp. 1706-1709, 2006.
- [3] J. Sundberg, "Acoustic and psychoacoustic aspects of vocal vibrato," *STL-QPSR*, vol. 35, no. 2-3, pp. 45-68, 1994.
- [4] R. Fushimi, E. Fujinawa, T. Narumi, et al., "Vibrato-matic: producing vocal vibrato using EMS," 8th Augmented Human International Conference, 2017, pp. 1-5.
- [5] D. Cattano and L. F. Cavallone, "Chapter 30 - Percutaneous Dilational Cricothyrotomy and Tracheostomy," *Benumof and Hagberg's Airway Management (Third Edition)*, C. A. Hagberg, Ed., W.B. Saunders, 2013, pp. 613-639.e3.
- [6] S. Matsubara, T. Watanabe, T. Suzuki, et al., "Optimal Motor Point Search Using Mm-Order Electrode Arrays," *IEEE Access*, vol. 11, pp. 58970-58981, 2023.
- [7] M. Morise, "Harvest: A high-performance fundamental frequency estimator from speech signals," in *INTERSPEECH 2017*, pp. 2321-2325, 2017.