

音響特徴量と抑揚の操作が発話音声の好感度に与える影響の分析*

☆堀池梓哉 (山梨大), 森勢将雅 (明治大)

1 はじめに

発話は他者とコミュニケーションをとる手段の一種であり、発話により、情報を伝達することや、他者の印象に影響を与える。適切な情報を伝達するための発話訓練は、ボイストレーニングや、言語障害のリハビリなどにおいて学習者の発話を洗礼させるためや、語学を学習するトレーニングとしても重要な役割を担う。発話訓練システムによる発話訓練を行う際は、言語的情報を扱う音声認識に基づき、適切な音素系列を発話することが可能となるように訓練することとなる。

発話から得られる印象についての研究が行われており [1], 発話の印象は他者に良い印象を与える重要な要素であると言える。発話の印象を改善するための発話訓練は、ボイストレーナの主観に基づいて行われている。一方で、好みの選択に影響を及ぼした音響特徴量の調査 [2] も行われている。これらのことから、発話の音響特徴量を操作することで印象を操作することが可能と推測できる。

本研究では、発話音声の好感度を改善することを目的とした音声変換手法を提案する。本手法により、発話者個人に合わせた客観的な好感度の指標である好感度測定モデルの生成が期待される。ここで、好感度の定義は、前報 [3] に従い、相手に与える主観的な聴覚的印象と定義する。

2 本研究の位置づけと音声加工法

2.1 前報の評価

前報 [3] において、F0 とスペクトル包絡の操作による好感度改善法を提案した。F0 は、発話者の性別によって傾向が異なることを示した一方で、スペクトル包絡は性別にかかわらず共通の傾向が見られ、その変化量は F0 の変化量によって増減する可能性を示した。また、本変換手法により音声の好感度を改善できることを示唆した。本研究では、F0 の操作に関して、声の高さに加えて声の抑揚の操作を行い、発話好感度の改善を図る。

2.2 抑揚の大きさの制御法

本実験では、前報 [3] の基本周波数のメル尺度による操作とスペクトル包絡の線形シフト操作に加え、抑揚の大きさの制御を行う。この制御は文献 [4] に用いられる式 1 を使用する。ここで、 N は有声部分の分

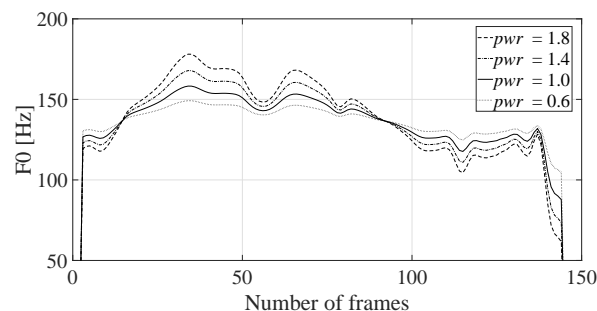


Fig. 1 抑揚の大きさの制御による F0 軌跡

析フレーム総数, $F0(n)$ は n 番目の有声分析フレームでの F0, pwr は抑揚制御のためのべき乗係数, $F0gm$ は有声部分の F0 の幾何平均を表している。また、図 1 に、この制御式を用いた制御による F0 軌跡を示す。縦軸は F0, 横軸はフレームを示す。

$$F0'(n) = F0gm \left(\frac{F0(n)}{F0gm} \right)^{pwr} \quad n = 1, \dots, N \quad (1)$$

3 主観評価実験

F0 操作とスペクトル包絡操作による音声の好感度改善を調査するため、主観評価実験を行った。実験条件を Table 1 に示す。本実験は、音響特徴量を被験者に操作させるため、専用の GUI (Fig. 2) を使用した。まず、操作前に変換前の発話音声を聴取し、聴取した音声を 7 段階で評価させた。評価結果は GUI 上で入力させた。次に、スライダーで F0 とスペクトル包絡を操作し、Play ボタンで操作後の音声を聴取する。最も良い印象のものになるまで操作と聴取を繰り返す。変化量の最大値と最小値は、前報 [3] の実験を基に決定した。良い印象と感じた変化量を Hold ボタンで一時保存し、保存した変化量での変換音声を Replay ボタンで聴取することができる。最良の印象と感じた音声を同様に 7 段階で評価させ、GUI 上に入力させた。その後、Write ボタンで変化量を書き込み、Close ボタンで次の音声へ進む。この試行を 18 音声分繰り返す。

4 実験結果と考察

主観評価実験の結果を Figs. 3, 4 に示す。Fig. 3 は、発話者毎の変換前と変換後の評価値との差の平均を示しており、縦軸は発話者、横軸は変換後の評価値と

* Analysis of relationship between likability and acoustic features and intonation, HORIIKE, Shinya (University of Yamanashi), MORISE, Masanori (Meiji University).

Table 1 音声評価実験の実験条件	
評価音声	好感度音声データベース [6]
評価音声数	6 名 × 3 文章 = 18 音声
発話内容	「桜がきれいです、あなたを愛しています、世界は すばらしいですね、会えてよかったですね。」 「精神は創造的高揚によって一種のメタモルフォー ゼを敢行する。」 「見やがれ、そんなだから、お前は、大学にも、行 かれないんだ、行けるもんか。」
評価者数	6 名 (男性 5 名, 女性 1 名)
評価環境	一般的な教室 (A-weighted SPL : 35 dB)
再生機材	Roland QUAD-CAPTURE SENNHEISER HD 650

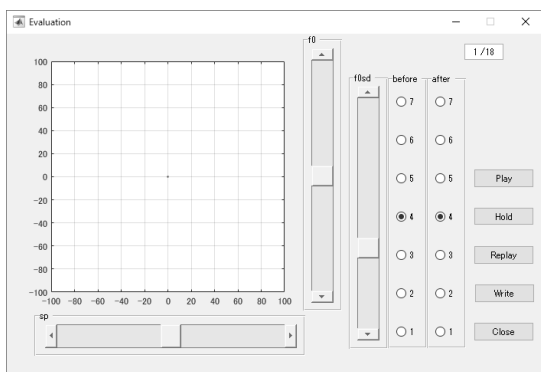


Fig. 2 実験に使用した GUI

の差, 誤差棒は 95%信頼区間を示す. 実験結果より, どの評発話者の音声でも好感度が有意に上がったことを示した. このことから, 発話音声に本変換手法を施すことにより, 好感度が改善される可能性を示した.

Fig. 4 に発話者毎の最大となる好感度を得られた音声における F0 の変化量とスペクトル包絡の変化量, 抑揚の変化率の平均を発話者毎に示す. 縦軸は発話者, 横軸は左のグラフから F0 の変化量, スペクトル包絡の変化量, 抑揚の変化率を示す. また, 誤差棒は 95%信頼区間を示す. F0 の変化に関して, 女性発話者の音声は全て有意に高く変化させる傾向が見られたが, 男性 3 の音声は変化量に有意差は見られなかった. これは, 以前の実験結果 [3] と同様に, 女性発話者の音声と男性発話者の音声において, F0 の変化量の傾向が異なることを示したといえる. また, スペクトル包絡の変化に関して, 全ての音声で有意に周波数軸の正方向に変化させていることが示された. 一方で, スペクトル包絡の変化量の性差に関して, 女性発話者の音声は約 40 Hz から 55 Hz の変化量に対し, 男性発話者の音声は約 20 Hz から 30 Hz という結果となった. これらの結果は, 前回の実験を再現できたことを示す. 抑揚の変化率に関して, 全ての発話者において 1.0 よりも有意に高い変化率にする傾向が見られた. これは, 性別にかかわらず, 抑揚を大きくすることで好感度が改善する可能性を示した. また, 文献 [4] において Big Five 特性印象 [5] のうち外向性の評価は, 抑揚が大きくなるにれて上昇していくことが示され

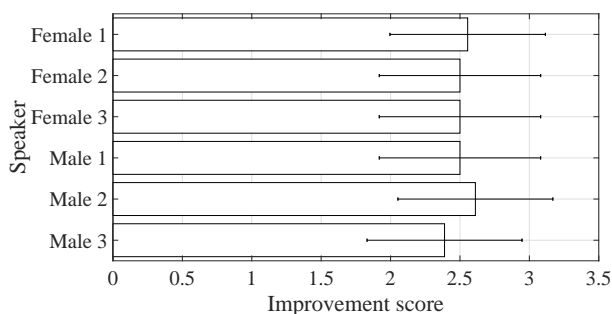


Fig. 3 変換前の評価値毎の変換後の評価値との差

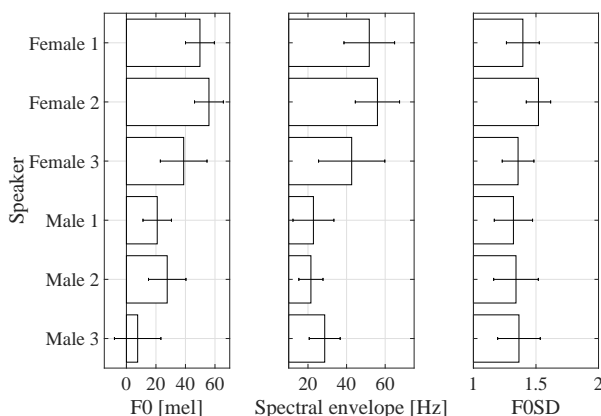


Fig. 4 発話者毎の最大となる好感度を得られた音声における各パラメータの平均値

ている. 外向性が高い性格印象を与える音声は客観的な好感度も高い可能性を示唆する.

5 おわりに

本研究では, 音響特徴量と抑揚の操作が音声の好感度に与える影響を調査した. 本変換手法により音声の好感度が改善されることを示し, F0 とスペクトル包絡は前報 [3] と同様の結果を示した. また, 発話者の性別によらず抑揚を有意に大きくする傾向が見られた.

今後の課題として, 音声波形から好感度測定モデルの開発を行うことが挙げられる. このモデルにより, 効果的な音声トレーニングが行えると考えられる.

謝辞 本研究は, 科研費 JP16H05899, JP16H01734 の支援を受けた.

参考文献

- [1] S. A. Collins, *Animal Behaviour*, 60(6), pp. 773–780, (2000).
- [2] K. Pisanski *et al*, in *Proc. RSB2018*, 285(1893), pp. 1–8, (2018).
- [3] 堀池ら, 音講論 (春), pp. 1141–1142, (2019).
- [4] 内田, 心理学研究, 76(4), pp. 382–390, (2005).
- [5] L. R. Goldberg, *Jurnal of Personality and Social Psychology*, 59(6), pp. 1216–1229, (1990).
- [6] 森勢ら, 音講論 (秋), pp. 395–396, (2017).