

発話音声の変換が聞き手に及ぼす感情変化に 現感情が与える影響の調査

井上修太[†] 田谷昭仁[‡] 戸辺義人[♯]

青山学院大学大学院理工学研究科理工学専攻[†]

東京大学生産技術研究所[‡]

青山学院大学理工学部情報テクノロジー学科[♯]

1. はじめに

近年, COVID-19 によりオンラインコミュニケーションが増加している. オンラインコミュニケーションの問題点として, 実空間と異なり, 相手に感情が伝わりづらいことがある. こうした中, 話し手と聞き手において感情状態に差異が生じ, お互いの意欲や士気低下などを招くことがある. この差異を埋めるために, 音声聞き手の感情に与える影響を考慮した音声変換を行うことで, より円滑なオンラインコミュニケーションが期待される. しかし, 同じ特徴の音声であっても聞き手の感情変化が, その時点での感情状態によって異なる可能性があるため, 本研究ではその影響を調査する.

2. 関連研究

片山¹⁾らは被験者の感情に応じた会話エージェントの最適な音声を探るためにユーザアンケートを行った. しかし, アンケートは想像上の回答のため, 実際にその感情のときの回答とは異なる可能性がある. また, Daniel²⁾らは音楽が人々の気分や感情を変化させることができることを発見した. 我々はこの音楽による感情誘導を用いる. すなわち, 被験者に対してアンケート直前に音楽による気分誘導を行い, 従来とは異なる本当の感情状態, またはより聞き手が感情を想像しやすい状態で回答を行ってもらおう.

3. 本実験で用いる感情分類

本章では実験で使用した感情分類について述べる.

3.1. 話し手の音声

話し手の感情分類は Happy, Relax, Angry, Sad, Neutral である. これは, ラッセルの感情円環モデルにおける第 1 象限から第 4 象限とそ

Investigation of the Influence of the Present Emotion on the Emotional Change that the Transformation of Speech Utterances Affects the Listener

[†] Shuta Inoue, Yoshito Tobe / Aoyama Gakuin University

[‡] Akihito Taya / The University of Tokyo

の中央に位置する neutral に由来する. さらに, 各々の感情で収録した音声に対して音声加工を行. 音声加工については, 早く, 遅く, 高く, 低く, 棒読み, 加工なしの 6 種類である. すなわち, 感情分類 5 種類と音声加工 6 種類で合計 30 種類の音声を作成される. なお, 今回用いた音声は一般人 1 名によって収録された. これは, 本システムの目的である音声変換の利用者を普段から音声に感情を込めることを意識している声優などではなく一般人と想定しているためである.

3.2. 聞き手の感情誘導

聞き手の感情分類は話し手と同様, Happy, Relax, Angry, Sad, Neutral である. 前述の通り, この感情誘導に音楽を用いる. Happy, Relax, Sad については, インターネット上に公開されている知的財産権の生じない音源を使用した. なお, 音源がその感情の誘導に適切かはパイロットテストを行い決定された. Neutral についてはとくに音源を使用しなかった. Angry については, 感情誘導をする音源が見つからなかった. そのため, ホワイトノイズを定期的に与えることでイライラを誘発することとする.

4. 評価実験および結果

同じ特徴の音声を聞いたとき, 聞き手の感情状態によって知覚が異なるのか評価実験を行った.

4.1. 評価実験

本実験は, 被験者 10 人によって行われた. 流れとして, 開始と同時に 90 秒間音楽を用いた感情誘導を行う. 次に, 30 秒間を一つのサイクルとし, 初めの 10 秒で短い音声 2 回流れ, 残りの 20 秒でその音声に対して以下の 3 項目についてアンケートに回答する.

Q1. 音声の快不快度 (1:不快 - 10:快)

Q2. 音声の違和感 (1:違和感 - 10:自然)

Q3. 音声の推定感情 (A:happy, B:relax)

表1 快不快と違和感の結果

	happy	relax	angry	sad	neutral
Q1	5.354	5.261	4.504	4.818	5.004
Q2	4.839	4.661	4.671	4.804	4.582

表2 音声加工別に比較した結果

	無	早	遅	高	低	棒
Q1	5.988	5.260	5.316	4.148	4.260	4.933
Q2	7.416	7.020	5.196	2.532	2.864	2.260

C:angry 4:sad 5:neutral)

Q1 の目的は、聞き手の感情によって同じ音声聞いたときに好感度の差があるかを調査するためである。Q2 の目的は、加工を行った音声を流した際に聞き手の感情によって違和感の度合いに差が出るかを調査するためである。Q3 の目的は、同じ音声を聞いた際に聞き手の心理状態によっては他の感情表現に聞こえるか調査するためである。これらを上記 3.2 に基づき、聞き手の5つの感情状態において評価を行い比較する。

4.2. 評価結果および考察

Q1 および Q2 の各感情におけるアンケート結果の平均スコアを表1に示す。Q1に関する傾向としては、聞き手の感情がラッセルの感情円環モデルに基づく快のとき、話し手の音声の好感度の高くなる結果となった。この結果に対し、複数群のデータから母平均が等しいかを解析する一元配置分散分析を行った。有意水準を 0.05、帰無仮説を「聞き手の感情状態によって話し手の声に対する印象は変わらない」とした結果、5つの感情状態での p 値は 0.083 となり、帰無仮説を棄却することはできなかった。しかし、neutral を除いた4つの感情状態で同様の分析を行った結果、p 値は 0.037 となり、帰無仮説を棄却する、すなわち同じ音声を聞いた際、「聞き手の感情状態によって話し手の印象は異なる」という結果となった。Q2 に関する傾向としては、neutral の状態が最も声の違和感が大きく、ほかの感情状態によって差は大きく見られなかった。この理由としては、neutral の状態のみ音楽を流さなため、純粋な音声に集中した結果だと思われる。Q1 と同様、聞き手の感情状態によって話し手の声に対する違和感が変わるのか一元配置分散分析を行った。その結果違和感には影響しなかった。

次に、加工別に見た Q1 および Q2 の結果を表2に示す。今回の音声加工ではピッチの変化を極端に行ったため、ピッチに関する項目の魅力度

表3 聞き手が判定した話し手の感情状態の割合

Q	happy	relax	angry	sad	neutral
happy	4.8%	0.7%	0.0%	0.0%	2.7%
relax	0.6%	10.0%	0.1%	0.7%	3.6%
angry	0.5%	0.4%	16.6%	1.0%	4.0%
sad	0.0%	1.2%	0.5%	10.8%	2.1%
neutral	2.4%	2.9%	3.8%	3.3%	27.3%

と違和感のスコアが両者とも低くなる結果となった。今後は、加工を少しずつ変化させ、どの加工度合いから違和感があるか調査する必要がある。そのほかの特徴として、棒読みの加工は非常に違和感がある一方でピッチの変化よりも好感度が高い結果となった。

次に聞き手の感情状態が neutral のときに判定した話し手の感情(列)と neutral 以外の感情状態のときに判定した話し手の感情(列)の割合を表3に示す。全体的な傾向として、話し手の感情を neutral と判定する割合が4割、happy と判定する割合が1割弱と差のある結果となった。また、聞き手の感情が neutral のときと他の感情状態のときに認識した話し手の感情が同じ感情に聞こえた割合は約7割となった。反対に残りの3割は異なる結果となった。以上のことから、同じ特徴の音声であっても聞き手の感情変化が、その時点での感情状態によって異なる可能性があると言える結果となった。

5. むすび

本研究では、聞き手の感情に与える影響を考慮した音声変換を行うシステムの作成を目的としている。そこで本稿では同じ特徴の音声であっても聞き手の感情変化が、その時点での感情状態によって異なる可能性があるため、影響を調査した。その結果、聞き手の感情状態によって音声の魅力度や違和感など複数の項目で影響がある可能性が見られた。

今後の展望としては、音声加工のパラメータ追加、またパラメータ度合いの変化による聞き手への影響調査を検討している。その上で聞き手の感情に与える影響を考慮した音声変換システムの作成を行う。

参考文献

- 1) Katayama, S., Mathur, A., Nakazawa, J., et al.: Situation-Aware Emotion Regulation of Conversational Agents with Kinetic Earables, In 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, pp.725–731.
- 2) Västfjäll, D.: Emotion induction through music: a review of the musical mood induction procedure, *Musicae Scientiae*, vol.5, pp.173–211.