

# 令和 元 年度 学 位 論 文 内 容 要 旨

## 論 文 題 目

ビブラートに着目した歌唱音声の類似度評価  
に関する研究

名城大学大学院 理工学 研究科 電気電子・情報・材料工学 専攻 博士後期 課程

氏 名

鈴木 千文

現在の歌唱力評価システムでは、歌手の歌唱方法を検出するだけであり、歌手の歌唱方法に関する詳細な評価は行われていない。そこで、本研究では歌手の歌唱方法を考慮した評価ができることが望ましいと考え、歌唱音声の類似度評価を行うことを目的とする。先行研究より、歌唱音声においてビブラートは歌手の個人性を反映することが示唆されている。また、物真似歌唱においてもビブラートは重要な歌唱方法であることが示されている。さらに、物真似歌唱によっては、ビブラートの速さより深さを変化させていることも報告されているため、ビブラートの深さの違いが歌手の個人性を判断する際の手がかりの一つになると考えられる。それを踏まえ、本研究では歌唱音声の類似度評価を行う指標として、ビブラートの深さを反映するビブラート特徴量を構築する。

まず、ビブラートの検出を目的として算出されたビブラート特徴量を従来法とし、従来法の算出方法を改良して算出されたビブラート特徴量を提案法とする。本研究で分析対象としている物真似なし歌唱音声と物真似歌唱音声、さらに CD 音源の目標歌手のビブラート音声に対し算出した。比較結果より、従来法では、物真似歌唱音声の方が物真似なし歌唱音声よりビブラートの深さが大きくなっているにも拘らず、ビブラート特徴量の値はほとんど変化しないことが分かった。一方で、提案法では、ビブラートの深さが大きくなるにつれ、提案したビブラートの値も大きくなることが分かった。さらに、今回用いた物真似歌唱音声は、物真似なし歌唱音声よりビブラートの深さが目標歌手のそれに近づいており、それに応じてビブラート特徴量の値も近づいていることも分かった。さらに、提案法によるビブラート特徴量の評価のために、物真似なし歌唱音声と物真似歌唱音声に対し従来法と提案法のビブラート特徴量をそれぞれ算出し、目標歌手のそれらとの歪みを算出した。その結果、従来法では物真似なし歌唱音声と物真似歌唱音声のビブラート特徴量の値の差がごく小さくなるため、目標歌手にビブラートの深さが近くなる場合でも歪みの差が大きくなる可能性がある。一方で、提案法では物真似歌唱音声のビブラート特徴量と目標歌手のそれとの歪みが小さくなり、類似度評価に適していることが明らかとなった。

次に、構築したビブラート特徴量から算出されるビブラート特徴量歪みと、主観評価実験により得られた評価値の関係性について調査した。この調査により、構築したビブラート特徴量とビブラート知覚の関係性を明確にできると考えられる。この調査では、類似度が高くなる場合は目標歌手の再現度が高くなると考え、類似度を主観的再現度と定義することとした。主観評実験は2種類実施しており、一つは目標歌手のビブラートを記憶した状態でビブラート合成音の主観的再現度の評価を行うものであり、もう一つは、目標歌手のビブラート音声とビブラート合成音を聴き比べることで主観的再現度を評価するものである。

まず、記憶による主観評価実験により得られた結果について述べる。この評価では、目標歌手がデフォルメしていた場合でも主観的再現度が高くなるかを検証する。実験に使用したビブラート合成音は、目標歌手の本来のビブラートの深さを持つもの、それより小さい深さを持つものとデフォルメした状態を含むものである。さらに、ビブラート特徴量歪みは、物真似歌手のビブラート音声と、作成したビブラート合成音に対し算出されたビブラート特徴量から得られる。これにより、ビブラートの深さが近い場合にはビブラート特徴量歪みは小さくなること、さらに物真似歌手がデフォルメをしている場合には、ビブラートの深さが大きいビブラート合成音との歪みが小さくなることを確認する。ビブラート特徴量歪みより、ビブラートの深さが近くなる場合ではビブラート特徴量歪みが小さくなることが示された。また、物真似歌手の中にはデフォルメをしている歌手もいることが分かり、さらに主観的再現度より、目標歌手によってはデフォルメしている場合に主観的再

(名城大学大学院)

現度が 1.2 倍程度高くなるケースがあることが分かった。

次に、聴き比べによる主観評価実験により得られた結果について述べる。この実験では記憶による主観評価実験で得られた傾向が、聴き比べによる実験でも観察されるかを確認するために実施した。実験に用いたビブラート合成音は、目標歌手のビブラートの他に、目標歌手とはビブラートの深さや速さが異なる歌手のビブラートも用いて作成されたものである。また、聴き比べに用いるリファレンス音声は、ビブラートの深さが小さい場合と大きい場合の 2 種類を用意した。その結果、記憶による主観評価実験と同様に、目標歌手のビブラートの深さよりやや大きいビブラートの深さを持つビブラート合成音でも、再現度が高くなるかあるいは変わらない傾向が見られた。つまり、被験者はデフォルメしたビブラートでも、目標歌手のビブラートであると許容しているといえる。さらに、ビブラートの深さが大きくなるにつれ、ビブラートの速さの違いが知覚されやすくなり、主観的再現度に影響を与えることも分かった。

そこで、本研究ではビブラートの深さと速さの関係性の調査を行った。この調査は、微細変動および大局的な変動の有無がビブラートの自然性へ与える影響に着目し行った。これらの変動に着目した理由として、目標歌手のビブラートの基本周波数の時系列を観察すると、ビブラートの深さが緩やかに変化する大局的な変動と、基本周波数が細かく変動する微細変動の 2 種類の変動を含むことが確認されたためである。そのため、ビブラートの深さと速さを変化させた場合に、これらの 2 種類の変動がビブラートの自然性に与える影響を調査することとする。この実験で用いたビブラート合成音は、微細変動および大局的な変動を含む場合と含まない場合で作成したものである。主観評価実験の結果より、目標歌手によらず、微細変動がビブラートの自然性に影響を与える顕著な傾向は見られなかった。しかし、この実験ではスペクトルの変動を含むビブラート合成音を用いたため、それがビブラートの自然性に影響を与えた可能性がある。そこで、スペクトルの変動がある場合とない場合において、ビブラート波形の違いがビブラートの自然性に与える影響を調査する実験を行った。

次の実験では、スペクトル変動がある場合とない場合において、大局的な変動のみを含むビブラート波形とそれを含まないビブラート波形を用いてビブラート合成音を作成した。主観評価実験の結果、目標歌手やビブラートの深さや速さによらず、ビブラート波形の違いがビブラートの自然性に与える影響はほとんどないことが分かった。つまり、この 2 つの実験により微細変動および大局的な変動はビブラートの自然性には影響しないことが示された。また、スペクトル変動の有無においては、スペクトル変動を含む場合では、含まない場合に比べてビブラートの自然性が高いことが分かった。さらに、目標歌手のビブラートの深さや速さによってスペクトル変動の有無がビブラートの自然性へ与える影響の大きさが変化する可能性があることが分かった。

このようにして本研究では、ビブラートに着目した歌唱音声の類似度評価を行うためにビブラート特徴量を構築し、ビブラート知覚との関係性を調査した。このビブラート特徴量により、用いた歌声データに対しては、ビブラートの深さにおける類似度を測ることが可能になったと考えられる。さらに、ビブラートの主観的再現度より、ビブラートの深さにおけるデフォルメを許容する範囲が明確になったと考えられる。今後は、構築したビブラート特徴量を、デフォルメを考慮したものに改良する予定であり、ビブラートの速さを反映したビブラート特徴量の構築も行う。また、ビブラートの深さと速さの関係性の調査に関しては、デフォルメの範囲に着目して調査することが挙げられる。

(名城大学大学院)