

研究概要報告書【サウンド技術振興部門】

(1/4)

研究テーマ	歌唱における顔面の動きと皮膚振動感覚および歌声の音響的特徴の関係性解明	報告書作成者	高橋純
研究従事者	高橋純		
研究目的	歌は、私たちにとって自己を表現するためのツールであり、人々に感動や喜びを与えることができる。しかし、歌唱技術は感覚的なものであり、歌声や発声器官は目に見えないことから訓練での上達は難しく、自身の歌声や歌唱にコンプレックスをもつ人々は少なくない。 歌唱訓練において、目の開き方、口角の高さ、口の開け方などの顔面の動きが歌声に影響を与えることは経験的に知られている。しかし、これらに客観的な根拠はなく、なぜ歌声に影響があるのかは明らかになっていない。これまで、申請者はリアルタイム MRI を用いた歌唱時の声道（口腔、咽頭、口頭）の制御に関する研究をおこない、口唇や口腔の形状変化が歌声の周波数特性に影響を与えていることを明らかにしてきた。例えば、優れた歌手に共通する音響的特徴として「歌い手のフォルマント(Singer's formant)」が存在する。歌い手のフォルマントは 3kHz 付近に現れるエネルギーの集約であり、この帯域は人間が最も可聴しやすく、さらに他の楽器にはないエネルギーの高まりであることから、オーケストラなどの楽音にマスクされにくいという特徴を持っている。これまでの研究から、プロの歌手は、歌い手のフォルマントの割合が高い歌声を出す際に、喉頭が下降して咽頭腔を拡大させていることがわかっている。この喉頭腔と咽頭腔の大きな空間は、歌い手のフォルマントの生成に必要な要因とされており、口唇を開くことで、下顎を下降させて喉頭を下方に維持させることから、口唇の開き具合は歌声に音響的な影響を与えていると考えられる。特に、この傾向は高音域の発声で顕著になり、プロの歌手は、高音域の発声になればなるほど口唇を開いていく。このことから、口唇の開き具合と歌い手のフォルマントの生成には大きな関わりがあると考えられる。 また、口唇の開き具合は歌唱表現とも大きく関わっている。これまでの研究から、プロの歌い手が、明るい声と暗い声の表現を歌い分けた際、暗い声では、より喉頭が下降し咽頭が拡大するのに対して、明るい声では、口唇を通常よりも開いて発声していることが明らかになっている。口唇を開くことで、歌声の音響的特徴が変化することが確認されており、明るい声では第2フォルマントと第3フォルマントが上昇することが明らかになっており、簡易的な声道モデルを用いたシミュレーションにおいても口唇が開くことで第2フォルマントが上昇することが確認されている。 これらの知見を踏まえて、本研究では、顔面の動きをリアルタイムにトラッキングできるシステムを用いて歌唱における顔面の動きを定量化し、同時に収録された音声を分析することによって、歌唱中に歌い手の顔面はどのような動きをしているのか、また口唇の開きや口角の動きが、歌声の音響的特徴にどのような影響を及ぼすのかを検討する。		

研究内容	実験方法 歌唱における顔面の動きは、歌声の音色や音質のためだけではなく、歌唱時の音楽的な表現に大きく関わる要素である。そのため単純に顔面の各部位を動かして歌声の音響的特徴がどう変化するかを検討しても、実際の歌唱行動とは乖離した事象である。そこで、音楽における「表現」に注目し、表現をつけて歌唱する中で、顔面の各部位がどのような動きや特徴を持つのかを定量的に示し、歌声の音響的特徴にどのような影響を与えるのかを検討する。 実験参加者 声楽の経験がある女性の歌い手 5 名 (A,B,C,D,E) を選定した。声種はソプラノ、メゾソプラノ等、声域に個人差があるが、後述する歌唱課題はどの声種でも歌える音域であることから、本実験ではそれらを区別することはしていない。 歌唱課題 童謡「チューリップ(へ長調 F dur)」を用いた。童謡チューリップは、女性のどの声種でも比較的に歌いやすい音域であり、歌詞がシンプルで曲自体の難易度が低く、事前の練習は不必要であることから選曲した。 歌唱収録 歌唱者は立位で、イヤホンから流れる伴奏に合わせて、童謡「チューリップ」の楽譜を見ながら2つの表出条件(表情あり・表情なし)で計2回歌唱した。1回目は「子供に歌ってあげるように表現豊かに」、2回目は「表現をせずに歌う」ように指示した。歌唱者の右方と後方にはグリーンバックを配置し、顔の正面から 2.0 m の位置に目線の高さの台に乗せたタブレットを用いて楽譜を表示させ、顔の正面から約 2.4 m 離れた位置にカメラを固定して、4K 60 fps にて顔の領域の動画を撮影した。動画と同時に音声の収録も行った。コンデンサマイクをオーディオインタフェースを経由して PC に接続し、標本化周波数 44.1 kHz、量子化ビット数 16 bit にてオーディオ編集ソフトウェア Audacity で録音した。 分析方法 画像認識 AI である、Media Pipe Face landmark detection を使用して顔ランドマーク座標を取得した。取得した顔ランドマークから平均座標を算出して平均顔とし、2つの条件(表情あり・表情なし)での差を求めた。歌唱中の頭部の動きを除去するため、座標は鼻根点を基準に補正した。また、左右の外眼角(目尻)の距離を相対値 100 として顔ランドマークの座標を正規化した。音声分析は STRAIGHT 分析法を用いて平滑化スペクトログラムを求め、任意のフレーム数の平均スペクトルを求めた。
------	---

研究のポイント	顔の表情や、それを用いた歌唱訓練に関する書籍は多数出版されており、その中で様々な訓練法が紹介されているが、効果が科学的に実証されているものは少なく、その根拠や原理も解明されていない。しかしながら、これまで一定の効果が認められていることも事実であり、単なる歌手やその学習者の経験則として無視できるものではなく、科学的な根拠や原理を明らかにすることによって、客観的資料に基づいた訓練法として活用できる可能性は十分にある。 本研究では、歌唱中の顔面の各部位を画像認識 AI を用いることにより、その形状や動きを定量化することができる。歌唱中の顔面の動きのなかで、声道の端である口唇の開閉度合いは直接的に歌声の音響的な特性に関わりがあると考えられるが、それ以外の部位が歌声にどう関わるのかという知見は申請者の知る限りない。これらを踏まえて、歌唱中の顔面の動きがどのようなになっているのかを定量的に観察し、歌唱表現や歌声との関わりを検討することは本研究の重要なポイントである。
研究結果	図1は、童謡「チューリップ」の冒頭部分(さいた、さいた、チューリップのはなが)の歌唱における顔面の各トラッキングポイントの平均位置を算出して得た平均顔を、さらに歌唱者5人分の値を平均して、表現ありと表現なしでの差の大きさを示したものである。図1から、表現ありと表現なしでは、口角の平均位置に大きな差が見られ、それに次いで頬の領域でも差が見られた。図2は、歌唱者Eの表現ありと表現なしにおける平均顔である。表現ありでは表現なしに比べて口角が上がることで口唇が開き、頬も上方に引き上げられている。他の歌唱者の平均顔でも同様の傾向が見られ、実際に撮影した動画からもその特徴を確認することができた。この結果から、歌唱において主に口角や頬に表現としての顔面の動きが現れるということが示唆された。図3は、STRAIGHT 分析法によって算出した各歌唱者のチューリップ冒頭部分の平均スペクトルである(青線1:表現あり、赤線2:表現なし)。図3に示すとおり、表現ありと表現なしでは、2-4kHz 付近で差が見られ、特に実験参加者 A,B,E はその差が大きく、表現ありの方が高い値であることがわかる。この結果から、表現ありで口角と頬が上がり、口唇が開くことによって、歌声の 2-4kHz 帯域に音響的な変化が起きたことが示唆された。
今後の課題	本実験で、顔面の各部位の位置や動きを定量化し、歌唱中の表情や歌声の音響的特徴との関係を検討することができた。しかしながら、歌唱課題は童謡1曲であったことから、さらに他の曲調や内容が異なる楽曲の表現では、どのような顔面の動きになるのかということは明らかになっておらず、更なる検証が必要である。また、音声分析に関して、本研究では歌唱全体の平均スペクトルを観察したが、例えば隣り合う音同士での変化量や、その変化の合計を指標化することにより、さらに顔面の挙動と歌声の音響的特徴の関係性が明らかになる可能性がある。 また、当初予定していたスキャニング型レーザードップラ振動計による歌唱時の顔面の皮膚振動の計測実験は、予算の都合上予備実験しか行えず、そのデータから成果を得ることはできなかった。

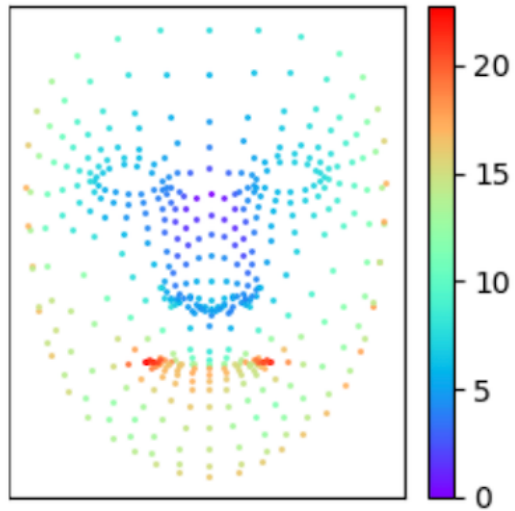


図1 歌唱者間で平均した表情あり・なしにおける平均位置の差

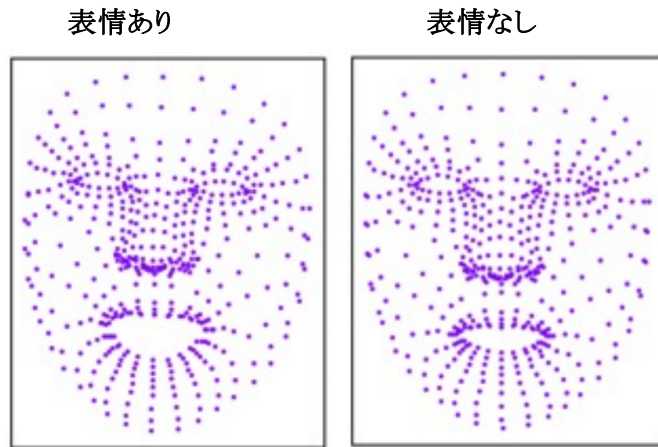


図 2 歌唱者 E における表情あり・なしの平均顔の違い

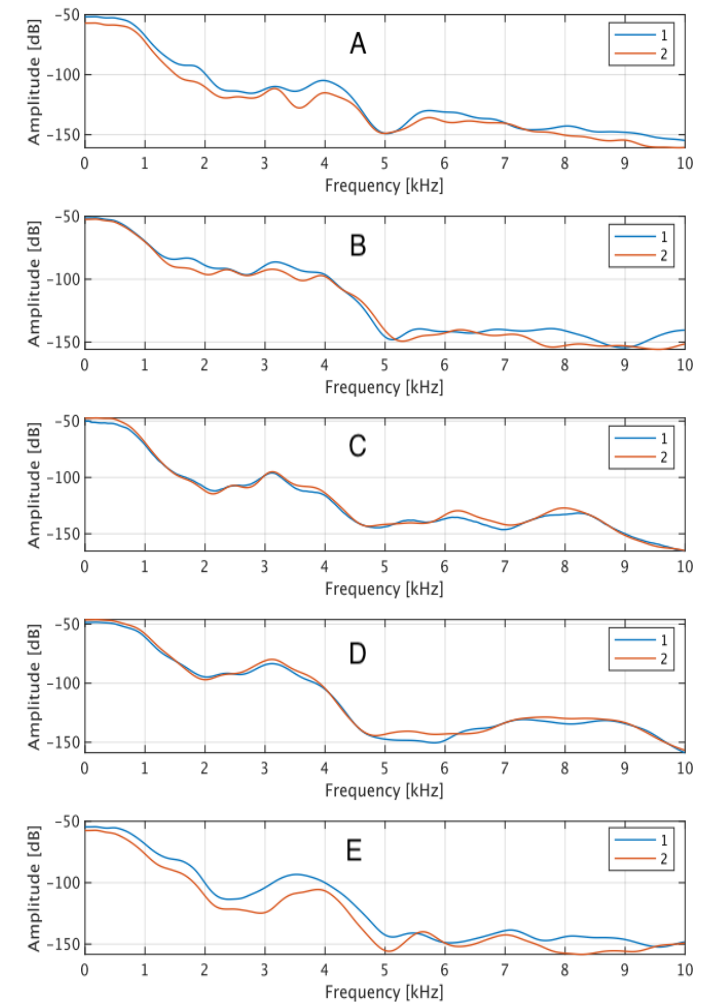


図 3 歌唱における各歌唱者の平均スペクトル