

イタリア歌曲を対象とした歌声の習熟度に関連する 周波数特性に基づく音響特徴量の変化

吉田 祥^{†1} 香山 瑞恵^{†2} 池田 京子^{†3} 山下 泰樹^{†4} 伊東 一典^{†2} 浅沼 和志^{†5}

概要：本研究では、歌声の響きに関連する周波数特性の強度や割合の定量化を検討し、歌声評価指標として提案することを目的としている。これまでの研究で、心理的印象に影響を与える音響特徴量が歌唱指導により変化することが確認されている。また、この音響特徴量を評価するための声楽初学者に特化した歌唱評価指標を検討してきた。本稿では、これらの成果をふまえ、初学者と声楽家によるイタリア歌曲歌唱を対象に、本指標の妥当性を検討した成果に関する考察を試みる。その結果、歌声の習熟度を評価する指標としての可能性を見出した成果について述べる。

キーワード：歌声、周波数特性、歌唱評価指標、音響特徴量、指導前後、定量化、Singer's formant

Changes in acoustic features based on frequency features related to skill-development of singing voices for Italian songs

Sho YOSHIDA^{†1} Mizue KAYAMA^{†2} Kyoko IKEDA^{†3}
Yasuki YAMASHITA^{†4} Kazunori ITOH^{†2} Kazushi ASANUMA^{†5}

Abstract: In this research, we aim to propose quantitative strength and proportion of frequency characteristics related to singing voices and propose it as a singing voice evaluation index. In previous studies, acoustic features that affect psychological impression It has been confirmed that the quantity varies according to singing instruction. We have also studied singing evaluation indexes specialized for singing voice students to evaluate this acoustic feature quantity. In this paper, based on these results, we will try to consider the results of examining the relevance of this index for Italian singing songs by first scholars and vocalists. As a result, we will describe the results of finding the possibility as an indicator to evaluate the proficiency of singing voice.

Keywords: Singing Voice, Features of frequency, Metrics for Singing Voice, Acoustical Features, Personalized, Quantification, Singer's formant

1. はじめに

歌声研究分野において、歌唱力と音響特徴量の関係について長年に渡って検討がなされている。物理的に良い声とされている条件に関する研究や不快な歌声の音響特徴量に関する調査など歌唱の習熟度と音響特徴量の関係した研究が数多くある^[1,2]。しかし、これらの研究はプロ歌手を対象としたものが多く、声楽初学者の歌唱指導と各音響特徴量の関係についてまだ議論が十分にされていない。

また、歌声は特異なスペクトル形状を持つことが知られている。その1つとして、歌声の3kHz付近のスペクトルピークがあり、Singer's formant と称される^[1,3]。これは母音によらず一定で、男女に関わらず「響く声」「艶のある声」

を特徴付けるといわれる。また、典型的な良い歌声の条件としてビブラートやSinger's formantの存在が指摘されており、これらを有することにより「歌声らしい」という聴感的印象となる^[4]。

本研究では、歌声の響きに関連する周波数特性の強度や割合の定量化を検討し、歌声評価指標として提案することを目的としている。これまでの研究で、心理的印象に影響を与える音響特徴量が歌唱指導により変化することが確認されている^[5]。また、この音響特徴量を評価するための声楽初学者に特化した歌唱評価指標としてSinger's formantの強度と割合の定量化を行ってきた。そして、これらが初学者とプロ歌手の歌声分析から、いずれかもしくはその両方の上昇が声楽の上達に必要であることが分かっている^[6]。本稿では、これらの成果をふまえ、イタリア歌曲を対象とし、既提案指標で初学者の指導前後、プロ歌手の声種別の差異を客観的に評価する。そして、イタリア歌曲に対しても既提案指標が適用できるかの考察を行う。

2. 本研究で用いる評価指標

本章では、本研究での歌声分析における歌声評価指標に

^{†1} 信州大学大学院
Graduate School of Shinshu University
^{†2} 信州大学工学部
Faculty of Engineering, Shinshu University
^{†3} 信州大学教育学部
Faculty of Education, Shinshu University
^{†4} 長野県工科短期大学校
Nagano Prefectural Institute of Technology
^{†5} 国立長野工業高等専門学校
National Institute of Technology, Nagano College

ついて述べる．ここでは歌声音声に含まれる周波数成分から習熟度に関係すると思われる響きの有無や程度について検討するため，高次（2.4~4.0 kHz）の倍音に着目する．本研究では，この帯域の周波数成分の割合と強さを定量化した結果を歌声評価指標とする．

2.1 周波数成分の割合

歌声データに対する高速フーリエ変換（Fast Fourier Transform : FFT）の結果に基づき，高次倍音成分の割合を算出する．歌声データの FFT により基本周波数と倍音がピークとして現れる．FFT の分析条件を以下に示す．

窓関数：ハミング窓，フレーム周期：10 ms，窓長：30 ms
FFT 後のデータにおける振幅スペクトル特性の大きさを d とすると，RMS 値は式(1)により求められる．

$$\text{RMS} = 10^{\left(\frac{d+80}{20}\right)} \quad (1)$$

RMS から歌声の Singer's formant が含まれるとされる 4.0 kHz までの帯域成分のうち，2.4~4.0 kHz の帯域の合計値の割合を SFR とする．以下の式(2)で算出する．

$$\text{SFR} = \frac{2.4 \sim 4.0 \text{ kHz の RMS 値の合計}}{4.0 \text{ kHz までの RMS 値の合計}} \times 100 \quad (2)$$

2.2 周波数成分の強さ

2.2.1 LPC 包絡線の算出方法

歌声に対する LPC 包絡線から高次倍音成分の強さを算出する．LPC 包絡線は以下の手順で求めるものとする．

- I. 音声データを読み込む．
- II. 音声波形の中心部分を切り出す．波形 x とする．
- III. プリエンファシスフィルタをかける．
- IV. ハミング窓をかける．
- V. LPC 次数 12 として，FFT のサンプル数 N ，自己相関関数のサイズを τ として，波形 x の自己相関 r を以下の式(3)で求める．

$$r(\tau) = \frac{1}{N-\tau} \sum_{k=0}^{N-1-\tau} x_k x_{k+\tau} \quad (\tau=0,1,\dots,\tau_{\max}) \quad (3)$$

- VI. 自己相関 r を用いて，LevinsonDurbin アルゴリズムから LPC 係数を算出する．

手順Ⅲではプリエンファシスフィルタを用いる．このフィルタは，音声のスペクトルの平坦化しダイナミックレンジを圧縮することで実効的な SNR を高める．また，手順Ⅴの自己相関関数は先行研究^[6]での分析時の計算式とは異なり，サンプル数から自己相関関数のサイズを減じた値で除算することとした．

2.2.2 Q 値

2.2.1 項に示した手順で求められた LPC 包絡線を用いて，倍音成分の強さとしての Singer's formant の鋭さを算出する．LPC 包絡線における Singer's formant の鋭さを Q 値とする．2.4~4.0 kHz の帯域において，最大ピークの周波数を f_a ，

ピークの両側に -3 dB となる周波数を f_b , f_c ($f_b > f_c$) とし，Q 値を以下の式(4)で算出する．

$$Q = \frac{f_a}{f_b - f_c} \quad (4)$$

2.2.3 2 凸

声楽初学者の場合，Singer's formant 帯域の成分が十分でなく，Q 値が算出しづらいことがわかっている^[7]．そこで，3 kHz 付近のピークの高さを Q 値に代わる指標とすることとした．図 1 にこの指標の求め方を示す．縦軸は周波数成分の強度の相対レベル，横軸は周波数を表す．Singer's formant が含まれるとされる帯域のピークとそのピークの直前にある谷との差を 2 凸とする．

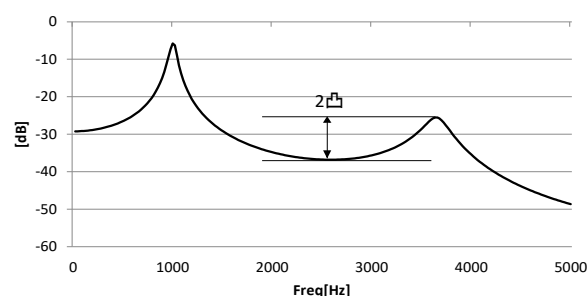


図 1 2 凸の求め方

3. 初学者の長期指導前後における評価実験

本章では，前章で示した評価指標を用いて初学者の指導前後における歌唱習熟度の変化を評価する実験について述べる．

3.1 実験環境

本実験の被験者は，女性の初学者 4 名（被験者 A, B, C, D）である．初学者は，教育学部音楽教育コース所属の大学生であり，大学のカリキュラムに沿ったレッスンを継続的に受けている．このカリキュラムは，最初の 1 年間に基本的な呼吸法や発声法などを学び，歌唱技術の習得のための指導が行われる．次の半年間では，近代イタリア歌曲などを教材に用いて，習得した歌唱技術を発展させる指導が行われる．本実験では，これらの指導を受ける前を指導前，これらの指導を 1 年半受けた後を指導後とする．

本実験で使用する楽曲は，指導者の意見から「Caro mio ben」（作詞：不明，作曲：Tommaso Giordani）とした．被験者は，分析対象とする区間を独唱し，各被験者につき 1 回の収録で 2 回ずつの歌唱をさせた．収録は音の反響しない静かな部屋で行い，レコーダは ICR-PS605 RM (SANYO) を使用した．被験者にはレコーダから 2 m 離れた正面立位で歌唱させた．音声はサンプリング周波数 44.1 kHz，量子化ビット数 16 bit で収録した．分析対象とするデータは，指導者に録音した 2 つの歌声データを聴かせ，被験者の習熟の程度がよく表れていると判断された 1 つの歌声とした．

また、「Caro mio ben」内の歌詞「Cessa」区間の母音/e/と「Tanto」区間の母音/a/のそれぞれの安定区間を評価対象区間とした。

また、本稿では、各指標のデータ偏りを可視化するために、独自に定めた図を用いる。描画方法は、次の通りである。グラフ内の各被験者のプロット点は、各被験者の歌声データの解析結果に対して SFR と 2 凸の各中央値とする。また、全データがどの位置に分布するかを表すために、2 つの指標それぞれの中央値を線で描画する。縦線の長さは、2 凸の下限が第 1 四分位値と上限が第 3 四分位値とする。同様に、横線の長さは SFR の下限が第 1 四分位値と上限が第 3 四分位値とする。そのため、図は 2 本の線の交点に点がプロットされるようになる。

3.2 実験結果

初学者の SFR と 2 凸の Cessa 区間の分析結果を図 2, Tanto 区間の分析結果を図 3 に示す。いずれのグラフも横軸が SFR であり、縦軸は 2 凸である。指導前の声を濃灰、指導後を薄灰で表す。また、各図のアルファベットはその図の被験者を表す。

図 2 に示す初学者の Cessa 区間の結果では、指導前後を比較すると被験者 A 以外の 3 名の被験者については少なくともどちらかの指標が上昇していた。被験者 A は SFR の中央値が 2.7、2 凸の中央値が 7.7 の下降がみられた。図 3 に示す初学者の Tanto 区間の結果では、指導前後を比較すると被験者 C, D においては、いずれの指標にも上昇が確認できた。被験者 A, B の 2 凸の中央値が 0.0 と変化がみられなかった。

3.3 考察

解析結果から、被験者 A 以外の Cessa 区間、Tanto 区間については、SFR、2 凸の数値の維持、または上昇という

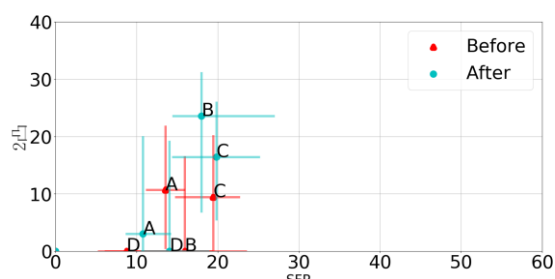


図 2 初学者の指導前後の Cessa

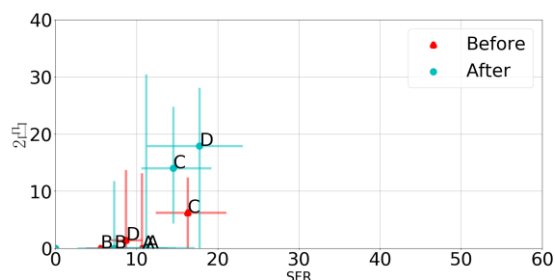


図 3 初学者の指導前後の Tanto

結果となり、指導が進むにつれて習熟度の変化が伺えるとともに本実験で用いた指標の妥当性が示唆される。被験者 A の Cessa 区間のみが SFR、2 凸どちらの数値も低下していた。このデータについて、指導者のヒアリングによる印象評価を行った。その結果、指導前が「まだ響かず、平べったい声」、指導後が「成長は感じられるが、収録日のコンディションが悪い」というコメントを得た。そのため、数値が下がった原因として被験者 A の声のコンディションに問題があり、声が出ていなかったことが考えられる。また、歌唱している母音が異なったが、母音ごとの差異は確認されなかった。特に女性歌手の場合、高いピッチでの歌声のため、プロ歌手であっても母音のフォルマント周波数が母音ごとにほとんど差異がないことが知られている^[8]。これらのことから、本実験で用いた評価指標と歌唱の習熟度との関係が示唆される。

4. プロ歌手の声種別による変化

本章では、2 章で示した指標を用いて、プロ歌手の声種別による歌声の変化を解析する。そして、これらの指標の声種別の有用性や性別による変化を考察する。

4.1 解析条件

本解析における被験者は、演奏活動や音楽大学で歌唱指導も行っているプロ歌手男性 6 名、女性 7 名である。各プロ歌手の詳細な情報を表 1 に示す。収録環境、解析方法、解析対象、解析データの可視化方法等の実験環境は 3 章と同様である。

4.2 解析結果

プロ歌手の SFR と 2 凸の結果を 3 章と同様の手法で図 4 ～7 に示す。いずれのグラフも横軸が SFR であり、縦軸は 2 凸である。図 4 は女性の Cessa 区間、図 5 は女性の Tanto 区間、図 6 は男性の Cessa 区間、図 7 は男性の Tanto 区間を表す。また、グラフ中のラベルは、Mez. がメゾ・ソプラノ、Sop. がソプラノ、Ten. がテノール、BB. がバス・バリトン、Br. がバリトンである。各図のアルファベットは、表 1 に示したプロ歌手と対応している。

解析結果から、3 章の初学者と比較した場合に図 4 の女性 Cessa 区間のプロ歌手 C, I 以外を除いたプロ歌手は SFR、2 凸ともに高い数値となった。また、プロ歌手の男女別に比較すると、男性のほうが SFR が高くなる傾向であった。図 7 の男性の Tanto 区間は、プロ歌手 J の SFR の第 1 四分位と第 3 四分位との差が 17.0 と大きくなり、プロ歌手 E は 2 凸が最高値となった。また、プロ歌手 L の SFR が 39.6、プロ歌手 E の 2 凸が 25.6 で最高値となった。しかし、ソプラノのプロ歌手 I の SFR は 11.2、2 凸が 0.0 と最低値となった。図 5 の女性 Tanto 区間は、プロ歌手 C の SFR は 14.7、2 凸が 15.9 で最低値となり、ソプラノのプロ歌手 I の SFR は 16.9 と低くなったが、2 凸が 25.8 で最高値となった。図 6 の男性の Cessa 区間は、バス・バリトンのプロ歌手 E の

表 1 プロ歌手の情報

被験者	所属・演奏キャリア	指導歴	声種
A	藤原歌劇団・28年	25年	バス・バリトン
B	関西二期会・20年	11年	ソプラノ
C	東京二期会・33年	21年	ソプラノ
D	藤原歌劇団・28年	5年	バリトン
E	東京二期会・37年	28年	バス・バリトン
F	東京二期会・37年	21年	メゾ・ソプラノ
G	藤原歌劇団・17年	7年	メゾ・ソプラノ
H	藤原歌劇団・25年	14年	バリトン
I	東京二期会・35年	29年	ソプラノ
J	藤原歌劇団・16年	16年	バリトン
K	東京二期会・32年	28年	ソプラノ
L	東京二期会・19年	13年	テノール
M	東京二期会・26年	21年	メゾ・ソプラノ

SFR は 23.0, 2 凸が 0.0 と最低値となった。

4.3 考察

ここでは、声種や性別における SFR と 2 凸の差異について性別ごとに考察を行う。

4.3.1 女性プロ歌手

女性プロ歌手は Cessa 区間, Tanto 区間に共通してみられる特徴があった。プロ歌手 C, I の声種はソプラノであり、メゾ・ソプラノのプロ歌手と比較した場合に SFR が低い値となる傾向が確認された。Singer's formant は、母音のフォルマント周波数に依存していると言われている^[8]。そのため、歌声が話しより歌声の方がはるかに大きく聞こえフォルマント周波数の特定が困難であるソプラノ歌手は、他の声種と比較すると数値が低く出る傾向にある。このことから、これらの指標が Singer's formant との関係性が強く、歌声評価指標としての妥当性が示唆された。しかし、声種がソプラノであるがメゾ・ソプラノと同程度の数値を示したプロ歌手が B, K である。これは、本稿で示した歌声評価指標では表せない歌唱習熟度の指標の存在が示唆され、さらに詳細に検討する必要がある。

4.3.2 男性プロ歌手

男性プロ歌手は、Cessa 区間でプロ歌手 E が他のプロ歌手と比べて大きく数値が低くなる結果となった。しかし、同一歌唱である Tanto 区間では他のプロ歌手との差がみられなかったため、収音日のコンディションの影響で本来の声ではなかったことが考えられる。その他のプロ歌手は Cessa 区間, Tanto 区間ともに女性と比べ高い数値を示した。男性歌手には Singer's formant が存在することが知られている^[4]。そのため、先行研究と本解析結果との合致する部分があり、本指標の歌声評価指標としての妥当性が示唆される。また、先行研究から Singer's formant の中心周波数はテ

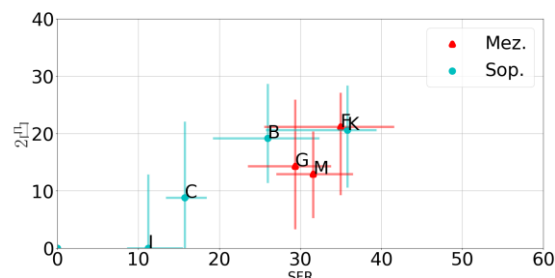


図 4 女性プロ歌手 Cessa 区間

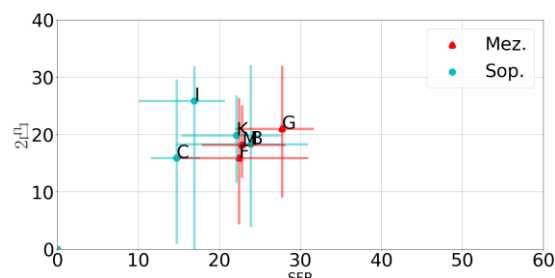


図 5 女性プロ歌手 Tanto 区間

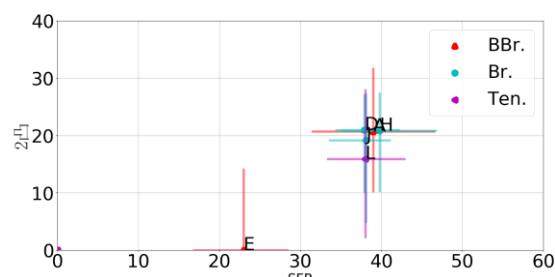


図 6 男性プロ歌手 Cessa 区間

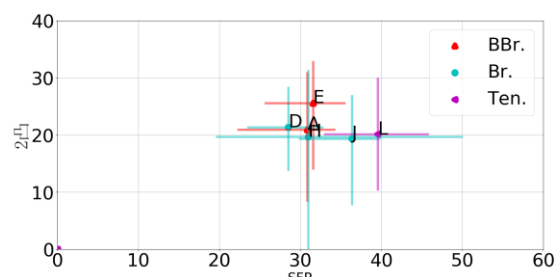


図 7 男性プロ歌手 Tanto 区間

ノール歌手が 3～3.8kHz, その他の声種の男性歌手も 3.0kHz 付近で強い Singer's formant を生成していることが知られている^[8]。Cessa 区間, Tanto 区間ともに男性歌手は他の声種と比べて高い数値を示す傾向であった。このことから本指標の歌声評価指標としての妥当性が示唆される。

5. おわりに

本稿では、初学者とプロ歌手のイタリア歌曲を対象とした歌声に対して、響きに関する音響特徴量による評価結果から歌声の習熟度や声種の区別との関連について考察した。その結果、本研究で用いる評価指標により、初学者の

指導前後の歌声と、プロ歌手の声種を区別することができることが示唆された。これらの成果より、本研究で用いる評価指標と歌唱の習熟度との関係が示唆された。また、これらの評価指標を用いることで、歌声を客観的に評価できる可能性を見出した。しかし、提案指標では説明ができない結果もあった。今後は、主観的評価とそれに対応する客観的評価としての解析結果をさらに増やしつつ、長期的かつ個人に特化した歌声評価の分析を進めていく。また、今回用いた歌声を評価する指標以外にも歌唱の習熟度に影響を与える指標の探究も進めていく。

謝辞 本研究の一部は科研費 15K01022 の支援を受けた。

参考文献

- [1] W.T.Bartholomew. A Physical Definition of "Good Voice- Quality" in the Male Voice, J. Acoust. Soc. Am., 6(1), p.25-33, 1934.
- [2] J.Sundberg. The KTH synthesis singing. Advances in Cognitive Psychology , Special issue on Music Performance, 2(2-3), p.131-143, 2006.
- [3] J. Sundberg. Articulatory interpretation of the 'singing formant'. J. Acoust. Soc. Am., 55(4), p.838-844, 1974.
- [4] 齋藤毅他. 歌声らしさの知覚モデルに基づいた歌声特有の音響特徴量の分析. 日本音響学会誌. 64(5), p.405-417, 2008.
- [5] 山辺大貴他. 歌声の心理的印象と音響特徴量との対応付けによる歌唱の習熟度評価に関する基礎的検討. 信学技報応用音響. 112(266), p.61-66, 2012.
- [6] 野田美春他. 歌唱の習熟度に関連する周波数特性に基づく初学者とプロ歌手の音響特徴量に関する考察. 信学技報教育工学. 115(444), p.35-40, 2016.
- [7] 佐久間雄輝他. 歌声の習熟度に関連する周波数特性からみた音響特徴量. 信学技報教育工学. 114(441), p.45-50, 2015.
- [8] ヨハン・スンドベリ. 歌声の科学. 東京電機大学出版局. P.115-130. 2007.