

歌唱における歌い手のフォルマントの安定性 —母音の移行を含む歌唱のスペクトル分析—

高橋 純[†] 津崎 実[‡]

[†] 京都市立芸術大学大学院音楽研究科 〒610-1197 京都市西京区大枝沓掛町13-6

[‡] 京都市立芸術大学 〒610-1197 京都市西京区大枝沓掛町13-6

E-mail: [†]johnandiamo@gmail.com, [‡]minoru.tsuzaki@kcua.ac.jp

あらまし クラシック歌唱では、ベルカント唱法という理想的な声質を実現する歌唱法があると言われている。優れた歌い手の歌声をスペクトル分析すると、共通する音響特性として歌い手のフォルマント(Singers Formant)が存在する(Sundberg,2007)。しかし、歌詞や旋律に伴い母音や音高が変化の中で、歌い手のフォルマントの生起や特性の時間的なふるまいについては、未だ十分に調査されていない。本研究では音声分析によって「音高・母音の変化」に対する歌い手のフォルマントの安定性を検討し、被験者の熟達度との相関を見た。その結果、熟達度の高いプロの歌手に優位性が見られた。

キーワード 歌い手のフォルマント, 歌唱, 母音, スペクトル分析,

Stability of Singer's Formant in Singing —Spectral Analysis of Singing including the transition of vowel—

Jun TAKAHASHI[†] and Minoru TUZAKI[‡]

[†] Kyoto City University of Arts 13-6 Kutsukake-cho, Oe, Nishikyo-ku, Kyoto 610-1197, Japan

[‡] Kyoto City University of Arts 13-6 Kutsukake-cho, Oe, Nishikyo-ku, Kyoto 610-1197, Japan

E-mail: [†]johnandiamo@gmail.com, [‡]minoru.tsuzaki@kcua.ac.jp

Abstract Abstract

The Bel Canto is one of the singing technique in the classic singing, and it is assumed to have an ideal voice quality. The existence of the “singers’ formants” has been regarded as a common acoustic property on the basis of acoustic analyses of the voices sung by excellent singers. Few studies, however, have focused on the stability of the singers’ formant against the changes in vowels and tone pitch which should be made with the lyrics, melody. In the current study we examine the stability of the singers’ formants against the change in pitches and -vowels by the spectral analysis. Its correlation with the level of experience of the singers was investigated, the proficiency of the subject. The results revealed that singers of high proficiency provided scores showing less variations in singers’ formants than singers of less proficiency..

Keywords Singer's formant, singing, Vowel, Spectral Analysis

1.はじめに

1.1 研究の背景

今日の音楽教育にける演奏指導は、主に指導者自身の経験と主観的に得られた知識に基づいて行われている。その結果、学習者は教師の求める理想的な「音」を出すために、曖昧な感覚に頼らざるをえない現状におかれている。このことが、その演奏技能を習得することの困難さを生む一因となっている可能性は高い。

特に「歌唱」は他の楽器と比べると、音の発生に必要な動作を外部から観察できないという制約がある。歌手はこれらを常に自分の制御のもとにおいておかなければならない(Husler,1987)のである。例えば、クラシック歌唱ではベルカント唱法という理想的な声質を実現する歌唱法がある(Reid,1987)と言われている。しかしながら、実際にベルカントという言葉が指す実体が何であるかは、その用語の仕様のされ方の歴史的な変遷、歌唱法がある地域から別の

地域に伝わる際の変化などの影響もあり、未だに曖昧な部分が残っている。

その一方で自然科学的な解明を試みた研究結果から、その物理的な特徴が歌手のフォルマント(Singer's Formant, 以下SF)の存在であるとの認識に至っている。但し、この研究は歌唱された結果としての音響信号の分析結果であり、SFが存在しているからというだけで、それが理想的な発声であるということは断言できない。歌手が楽曲を演奏する際に、SFをどのぐらい維持しているのかは解明されていないからである。SFが楽曲中のすべての歌声に安定して存在しているのか、あるいは局所的に存在しているのかは現段階では分かっていない。それはこれまでの先行研究が歌唱全体の時間平均を観察した、長時間スペクトルによって得られているからであり、そのいずれであるのかを調べる必要性はまだ残された課題である。

1.2. 歌手のフォルマント

先に述べたように、優れた歌手の歌声をスペクトル分析すると、共通する音響特性として歌手のフォルマントが存在する。歌声にはフォルマント(Formant)と呼ばれる周波数群が存在し、それらを低い方から第1フォルマント、第2フォルマントと呼んでいく。第1フォルマントと第2フォルマントは、主に口腔内の面積と舌の位置の関係で決定され、母音の形成に関与している。SFは主に第3、第4、第5フォルマントから成り、そのエネルギーが3 kHz付近に集約されることで生成されている。3 kHz付近は、人間の最も可聴しやすい周波数域であため、人間の耳にはより聞こえやすい声になる。また、歌手の声がオーケストラと共に歌っても聞こえる理由は、オーケストラの音の長時間スペクトルには3 kHz付近に共振がそれほど顕著に現れないため、その付近ではオーケストラの音によって歌声がマスクされにくいという傾向を生むからである。優れた歌手の声が、楽器にはないフォルマントのピークを持つことにより、オーケストラと共に歌っても歌声が届くのである。

1.3. 歌声のスペクトル

歌声は、声帯で作られる喉頭音源が、声道（喉頭腔と咽頭腔）によるフィルターを通過することによって作られる。声帯の振動によって発生する喉頭音源スペクトルが、声道の断面積から算出された声道断面関数によって、歌声のスペクトルを作り出すのである。それを簡略化した図を(Fig.1)に示す。実際には、この2要素に加えて放射特性が関わる。喉頭音源は、声帯の厚さや張力などの要因によって変化し、音高が変化するという事は、声帯の張力や厚みなどの物理特性が変化することである。声帯は肺から供給される呼気圧を受けて振動する様式が変わり、振動の周期を変えることに繋がる。また声道断面関数は母音の生成にも深く関わっており、声道断面関数から母音が決まるのである。母音が変わるといことは、声道断面関数が変わることであり、声道自体の物理的な変化を意味する。また実際の発声器官において、これらの物理的变化はお互

いに影響を受け、それに伴い喉頭の位置や、声道の形状も随時変化する。

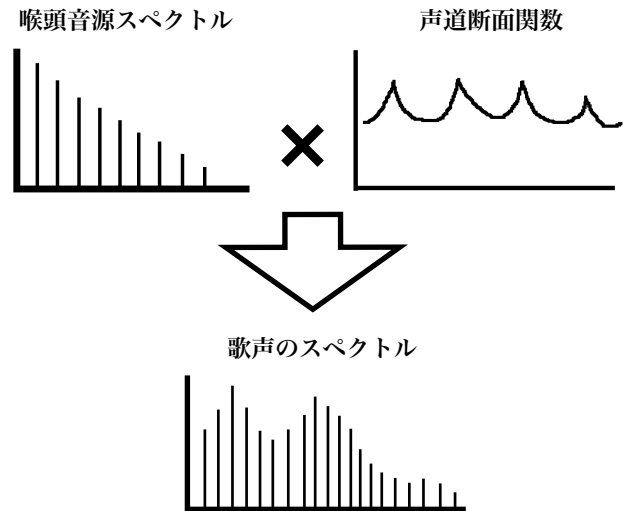


Figure.1 歌声の仕組み

2. 目的

先行研究によって見出されたSFの存在は、歌唱全体の長時間スペクトルの形状から観測されたものである。しかし、楽曲中の歌詞や旋律に伴い、母音や音高が変化する中で、SFの生起や特性の時間的なふるまいについては、未だ十分に調査されていない。歌唱には「旋律・歌詞」という二つの要素があり(Fig.2)、これをほぼ独立に変化させることが歌手には求められる。即ち、旋律における「音高の変化」と歌詞に含まれる「母音の変化」を同時に達成しつつ、歌手のフォルマントに代表される安定した声の質を保つことが要求される。

本研究の目的は「音高・母音の変化」に対する歌手のフォルマントの安定性を音声分析によって調べ、歌唱者の熟達度との相関を見ることである。

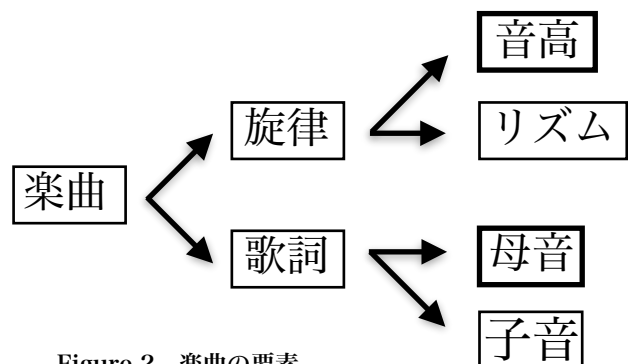


Figure.2 楽曲の要素

3.方法

3.1.実験方法

次の2点を軸として研究を進める.

- 1.「母音・音高」の要素を直交配置した歌唱課題の作成.
- 2.技量レベルの異なる歌唱者による課題曲の歌唱収録.

先にも述べた通り,歌声のスペクトルは声帯で作られる喉頭音源が声道に響くことによって作られる.優れた歌い手ほど,この声帯と声道が変化する状況で,SFを安定させることに成功しているか否かをSTRAIGHT平滑スペクトルを上に見られるSF付近のゲインの大きさを観察し,被験者の熟達度との相関比較することが本研究の手法上の特徴である.

3.2.歌唱課題

任意のある1音を主音として決定し,その完全五度上の音と,完全四度下の音を選び,合計3音を設定する.今回の実験は「母音の変化」と「音高の変化」の歌声を分析することから,2音間における音の移行を観察するために,3つのパターンの音型を作成した.一つ目は,同音程同士の移行(Unison.以下Uni),二つ目は,主音から,完全五度上への上行跳躍進行(Upward 5th.以下U5th),そして,三つ目は,主音から完全四度下への下行跳躍進行(Downward 4th.以下D4th)である.そしてそれぞれに,主要三母音である'a','i','u'を当てはめ,合計27通り(3通りUni,U5th,D4th×9通りaa,ai,au,ii,ia,iu,uu,ua,ui)の歌唱課題を設定した.歌唱被験者には,各自の歌いやすい三音を自己申告で選択してもらった.また,最高音には,声区転換であるパッサッジョ(passaggio)を通過した音を含むように指示した.

3.3.歌唱被験者

今回の実験では,ある程度の声楽経験がある歌唱者が実験対象であり,その被験者の熟達度と歌唱データとを相関比較することが本研究の目的である.そのため,声楽の中度経験者と高度経験者という2群に分けて,その対比を取ることにした.中度経験者には,京都市立芸術大学の声楽専攻の学部生と院生を含む学生(バリトン6名,テノール1名)7人,高度経験者には,同大学卒業生であり,より経験年数の高いプロのバリトン歌手2名とテノール歌手1名の合計10人に,先に述べた歌唱課題を課し録音した.先行研究から,女性の歌声では,SFの観測が難しいことがわかっているため,本実験では男性の歌声のみを扱うこととした.

3.4.録音方法

録音は,京都市立芸術大学内の静寂な部屋で実施した.ワイヤレスマイク(SENNHEISER, evolution wireless G3)をオーディオ・インターフェイス(Apogee Duet)を経由し,PC(Apple, MacBook pro retina)に接続した.また,サンプリング周波数は,44.1kHzで,量子化ビット数は24bitの圧縮なしでデジタル録音した.収録にはヘッドセット・タイプのワイヤレスマイクを使用した.その理由

は,マイクと歌唱者の口元の距離や方向性によって歌唱収録音声が変わるという事実があり,スタンド・マイクにするとスタジオ収録になれていない歌唱者などの場合は,姿勢の安定性を必要以上に録音側の都合だけで強いることとなるためである.

3.4.特徴量の計算方法

各歌唱被験者から得られた,27通の録音データを,STRAIGHT分析法によって平滑化スペクトルを求めた.音声分析変換合成システムSTRAIGHTでは,入力された音声は,基本周波数等の音源情報と,調音器官により形成された声道等の特性を示す滑らかな時間周波数特性に分解される(河原.2002).つまり,声を基本周波数や非周期音(ノイズなど),スペクトルなどに分離して,それぞれを分析したり,また操作や合成を行ったりすることが出来るのである.STRAIGHT分析法より得られたサウンドスペクトログラムから,音程変化部,母音変化部を目視によって決定し,遷移部を除いた1秒間の定常部に対するスペクトルを分析対象とした.そして切り出した1秒間の平均スペクトルを求め,その変化量をRMS(二乗平均平方根)によって求め特徴量とした.

4.結果

4.1.比較周波数帯の算出

SFの変化量を見るために,それぞれの録音データの比較周波数帯を決める必要があった.SFは主に第3,第4,第5フォルマントが3kHz付近に集約されることで生成されている(Sundberg.2007)ため,この3kHz付近を基点として,2つの方法で周波数帯を決定した.まず単純に,3kHzを中心とした2kHzから4kHzまでの周波数域について,各周波数ビンでのRMSを求めた.予備的に実際のデータを目視した結果,1区間と2区間におけるSFの変動が,おおよそ2kHzから4kHzまでの間で起こっていることがわかったことからこの区間を選定した.

そしてもう一つの2区間の変動の指標として,2kHz以上にある極大点を基準として,その前後500Hzを対象とする方法である(以下ピーク調整型RMSとする).2kHz以上にした理由は,Sundberg(2007)の研究から,母音の生成に寄与している.第1,第2フォルマントの平均は,男性においては2000Hzまでにあることがわかっているためである.そのため,必然的に第1,第2フォルマントは除外され,第3フォルマント以降の最大フォルマントであるSFが選ばれることになる.事前の検討により,SFの最大値を中心にする,その前後500Hzの範囲に共振の山がほぼ入ることが確認された.従って,第1の指標よりも狭い1000Hzという範囲で変化量をみることとなる.

4.2.録音データの正規化

録音データを分析していく上で,大きな問題点となったのが,1区間と2区間のSFを含む全体のパワーの違いである.これは,歌唱者によっても様々で,母音の変化や音高の変化に伴いパワーも変化した.しかし,中にはSFに関

してほとんどパワーが変わらないままのものも確認できた。これが歌唱者の技術に含むべきものかどうかということには議論の余地があり、今回の実験では、その両方を観察することにした。

よって、以下に呈示する第1区間と第2区間のSFの変動度合いの指標には周波数帯域の選択法の2種類と、正規化の有無の2種類を組合わせた4種が存在することとなる。

4.3.同母音間と違母音間における音高の変化

27通りの歌唱課題をUni,U5th,D4th×aa,ii,uuという9通りの同母音間における音程変化(以下same)と,Uni,U5th,D4th×ai,au,ia,iu,ua,uiという18通りの違母音間における音程変化(以下diff)という二つに分けて変化量を観察することにした。

4.3.母音組み合わせ・正規化の条件別の歌唱経験による差

10人の歌唱被験者中から、中度経験者として学生の7人と、高度経験者として、より経験年数の高いプロ歌手の3人に分けて、各組合せの第1区間と第2区間の差をRMS値で表した。(図中では中度経験者の学生群をama,高度経験者のプロ群をproと示す)。

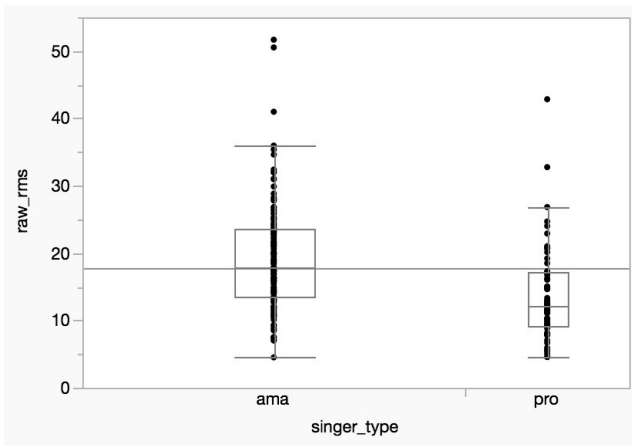


Figure.3
 パワー 正規化無し・2-4 kHz帯域のRMS指標・違母音

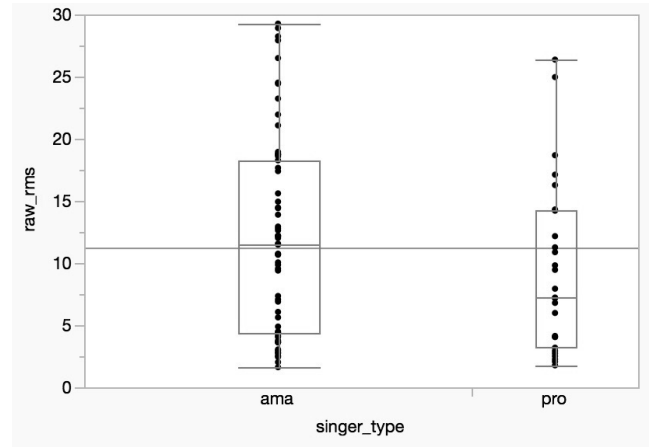


Figure.4
 パワー正規化無し・2-4 kHz帯域のRMS指標・同母音

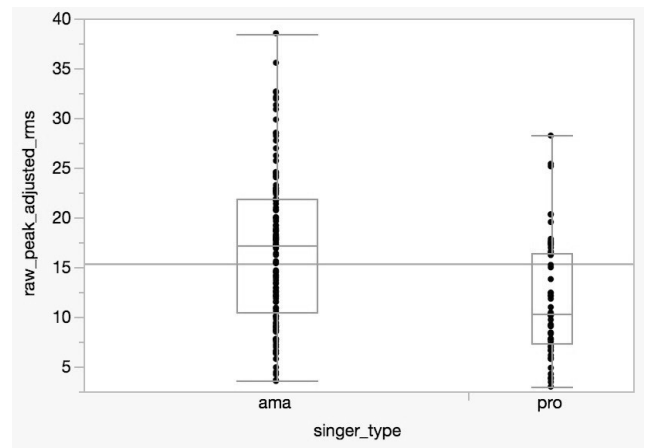


Figure.5
 パワー正規化無し・2kHz以上のフォルマントピークの前後500HzのRMS指標・違母音

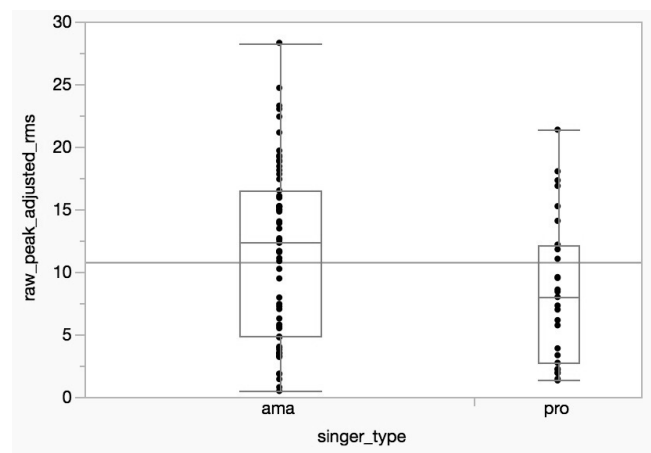


Figure.6
 パワー正規化無し・2kHz以上のフォルマントピークの前後500HzのRMS指標・同母音

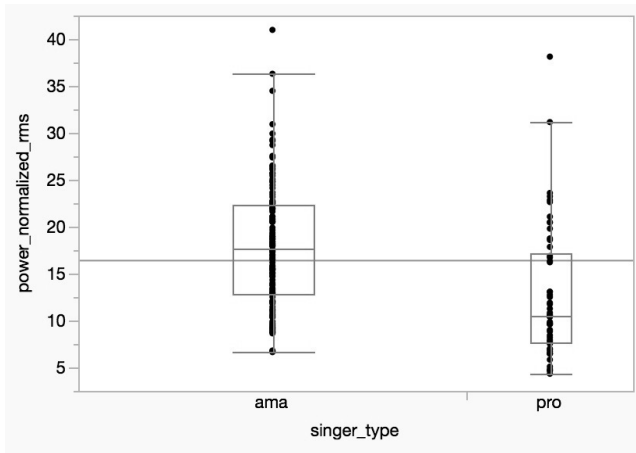


Figure.7

パワー正規化あり・2-4 kHz帯域のRMS指標・違母音

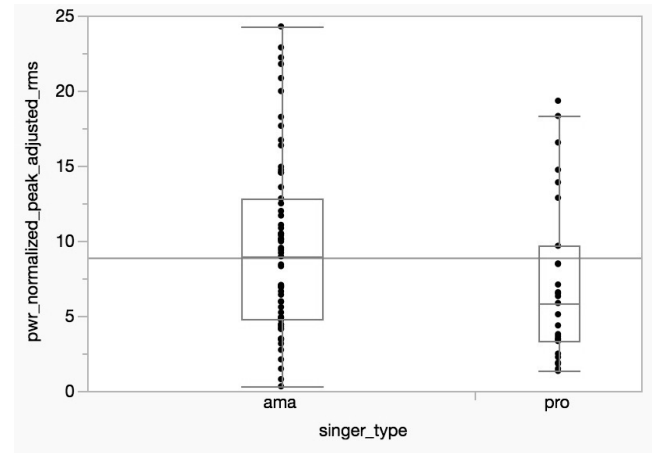


Figure.10

パワー正規化あり・2kHz以上のフォルマントピークの後500HzのRMS指標・同母音

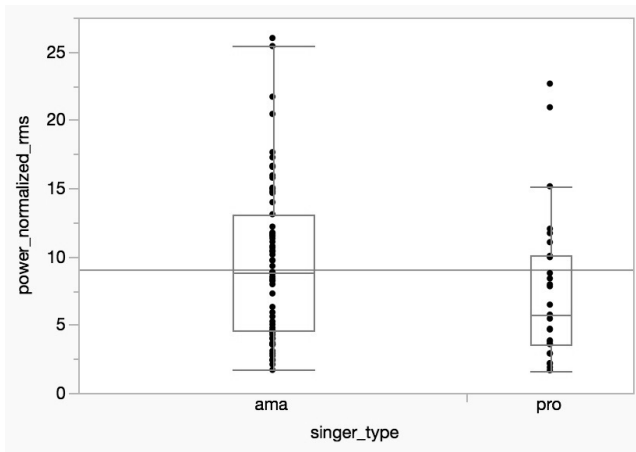


Figure.8

パワー正規化あり・2-4 kHz帯域のRMS指標・同母音

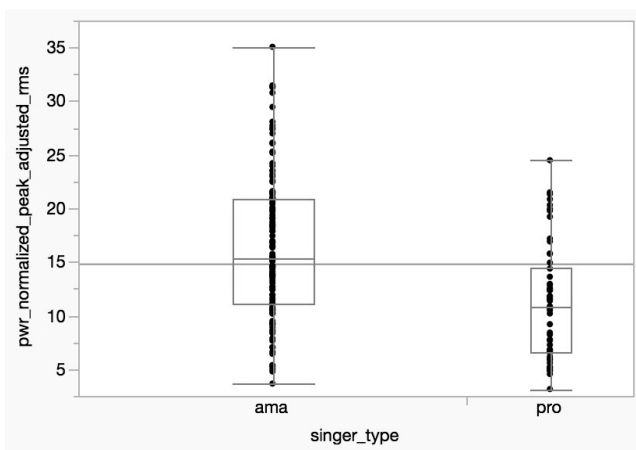


Figure.9

パワー正規化あり・2kHz以上のフォルマントピークの後500HzのRMS指標・違母音

パワー正規化無しの2-4 kHz帯域のRMS指標で、違母音について求めた結果は、学生の中央値が17.86で、プロの中央値が12.07となり、 X^2 検定を行うと $[X^2=15.15 \text{ df}=1 \text{ p}<0.0001]$ となり、有意差が観測できた。(Fig.3)パワー正規化無しの2-4 kHz帯域のRMS指標で、同母音について求めた結果はでは、学生の中央値が11.47で、プロの中央値が7.221となり、 X^2 検定を行うと $[X^2=2.563 \text{ df}=1 \text{ p}<0.109]$ となり、有意差はでなかった。(Fig.4)パワー正規化無しの2kHz以上のフォルマントピークの後500HzのRMS指標で、違母音について求めた結果では、学生の中央値が17.17で、プロの中央値が10.23となり、 X^2 検定を行うと $[X^2=10.52 \text{ df}=1 \text{ p}<0.001]$ となり、有意差が観測できた。(Fig.5)パワー正規化無しの2kHz以上のフォルマントピークの後500HzのRMS指標で、同母音について求めた結果では、学生の中央値が12.31で、プロの中央値が7.97となり、 X^2 検定を行うと $[X^2=4.238 \text{ df}=1 \text{ p}<0.039]$ となり、有意差が観測できた。(Fig.6)パワー正規化ありの2-4 kHz帯域のRMS指標で、違母音について求めた結果では、学生の中央値が12.91で、プロの中央値が7.708となり、 X^2 検定を行うと $[X^2=10.52 \text{ df}=1 \text{ p}<0.001]$ となり、有意差が観測できた。(Fig.7)パワー正規化ありの2-4 kHz帯域のRMS指標で、同母音について求めた結果では、学生の中央値が8.857で、プロの中央値が5.731となり、 X^2 検定を行うと $[X^2=4.238 \text{ df}=1 \text{ p}<0.039]$ となり、有意差が観測できた。(Fig.8)パワー正規化ありの2kHz以上のフォルマントピークの後500HzのRMS指標で、違母音について求めた結果では、学生の中央値が15.32で、プロの中央値が10.82となり、 X^2 検定を行うと $[X^2=17.78 \text{ df}=1 \text{ p}<0.001]$ となり、有意差が観測できた。(Fig.9)パワー正規化ありの2kHz以上のフォルマントピークの後500HzのRMS指標で、同母音について求めた結果では、学生の中央値が8.971で、プロの中央値が5.852となり、 X^2 検定を行うと $[X^2=2.563 \text{ df}=1 \text{ p}<0.109]$ となり、有意差が観測できなかった。(Fig.10)

5. 考察

それぞれの統計結果から、有意差が見られた組み合わせについては、連続した2音間における、SFの2区間の差が小さいということであり、その差が小さいと言うことは、音高の変化や母音の変化においてSFの出方が大きく変化しないことを意味している。つまり、それはSFが安定して出せている様子を窺わせるものであり、今回の学生群が理想として目指すべき歌唱がプロ群の歌唱であるとすれば、SFの安定性を増すことが鍵となるという示唆が得られた。音高と母音が変化するという事は声帯と声道の形状が変化することであり、そのような状況にもかかわらず、この「音高の変化」と「母音の変化」の2音間において、学生群よりもプロ群の方が、SFを維持し、より安定させているという結果は興味深いものである。

第1区間と第2区間が同一母音である場合には、パワー正規化無しでの2-4 kHz帯域のRMS指標と、正規化ありの2kHz以上のフォルマントピークの前後500HzのRMS指標では歌手の経験レベルによる群間の有意差が観察されなかった。つまり、学生群とプロ群の間では差が出にくいということがうかがえる。全体を通して見ても、同母音に関しては、違母音の場合よりも中央値の差が小さい。これは、母音が同じ場合は、声道の形状を保つことが可能であるために、母音が変化する場合よりも難易度が下がるからだと思う。これはつまり、ある程度の技術を要する学生たちは、同母音間における音高の変化においては、プロに近い状態でSFを安定させることができているが、より難易度の上昇した違母音間の音高の変化については、差があるということである。このことから、優れた歌手は、母音の違い、音高の違いを伴う歌唱の間で、SFが安定していることが歌唱者の技量レベルの指標としては、さらに望ましいということが示唆された。

6. 今後の課題

今回の研究では、学生7名とプロの歌手3名で行った。今後はさらに歌唱被験者を増やして分析していきたい。また今回は、歌唱被験者を熟達度によってレベル分けしたが、熟達度ではなく録音データを使った知覚評価実験を行い、その結果と音響特徴量実験の結果を相関比較することにより、歌手のフォルマントの安定性が、人間の聴覚心理にどのように働きかけるのかを知ることが出来ると期待できる。

この実験では、歌唱における2音間の「音高の変化」と「母音の変化」という局所的な部分において、歌手のフォルマントの安定性を見ることができた。しかし最終的には、楽曲を通じた歌唱全体において、どのように歌手のフォルマントが存在しているかということを知ることが目標である。今回の実験が、その解明に至る最初の一步になったのではないだろうか。

7. 謝辞

この研究は科学研究費助成金・挑戦的萌芽16K13185ならびに、カワイサウンド技術・音楽振興財団 第33回サウンド技術振興部門の助成を受けた。

文献

1. Sundberg, J, 榊原健一, 伊藤みか, 小西知子, 林良子訳, 歌声の科学, 東京電気大学出版局, 東京, 2007
2. Husler, F, Rood, Y, 須永義雄, 大熊文子訳, うたうこと, 音楽之友社, 東京, 1987
3. Reid, C.L, 渡辺東吾, ベルカント唱法, その原理と実践, 音楽之友社, 東京, 1987
4. 河原英紀, 片寄晴弘, “高品質音声分析変換システム STRAIGHTを用いたスキヤット生成研究の提案,” 情報処理学会論文誌, Vol.43, No.2, pp208-218, Feb 2002年.