

## 題目

# 歌唱における歌い手のフォルマントの安定性

## ー母音の移行を含む歌唱のスペクトル分析ー

高橋 純

## 概要

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)。先行研究では、楽曲の全体の長時間スペクトルによってのみ研究が行われてきた。しかし、歌唱とは多様なピッチと多様な音色（母音）を時間的に変化させるものであることから、単に平均で大きな値を取ることに加えて、それぞれの歌唱音の「歌い手のフォルマント」が安定していることも重要ではないかということに注目した。

そこで、本研究では「音高」の変化と「母音」の変化に着目し、それらを伴う歌唱課題を作成した。そして、歌唱者によって収録された歌唱録音を、音響計測実験により観察し、「歌い手のフォルマント」がどのような時間的なふるまいを行なっているのかということを調べた。そして、そこから「歌い手のフォルマント」に関する音響特徴量を抽出し、さらに並行して行った印象評価実験より得られた、すべての歌唱課題に対する主観評価値との関係を相関分析した。その結果、「歌い手のフォルマント」の占有率が高く、そしてその変化量が少ない方が、聞き手の聴覚印象において高い評価を得ることが明らかになった。

# Stability of Singer's Formant in Singing

—Spectral Analysis of Singing including the transition of vowel—

Jun TAKAHASHI

## 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 pitches which are assumed to be essential to "sing" melodies conveying lyrics. In this research, we focused on the change of "pitch" and the change of "vowel" and created a singing task with them. 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. We analyzed vocal singings provided by a various levels of singers and compared the resultant analysis data with the results of subjective evaluations achieved by an independent experiment. Based on the correlation analysis between the objective and subjective measures, we found that the occupancy rate of the formant of the singer is high, and that change amount or less gain higher evaluation in the auditory impression of the listener.

平成29年度博士論文

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—

2018年3月

京都市立芸術大学大学院

音楽研究科 博士(後期)課程 音楽専攻 声楽領域

高橋 純

# 序論

研究の背景と目的  
音声の定義      ベル・カント唱法

# 本論

## 第1章

### 発声器官

呼吸器官  
声帯  
声道

## 第2章

### 発生の理論

ソース・フィルター理論  
母音の生成  
歌い手のフォルマント

## 第3章

### 「歌い手のフォルマント」の研究と課題

長時間スペクトル分析の限界  
「音質」の変化の少なさ  
STRIGHT分析

## 第4章

### 歌声の音響計測実験

歌唱課題の作成  
「歌い手のフォルマント」の占有率  
「歌い手のフォルマント」の変化量

## 第5章

### 歌声に対する印象評価実験

全ての歌唱課題に対する主観評価値

## 第6章

### 主観評価結果と音響特徴量の回帰分析

「SF占有率の合計」  
「SF占有率の変化量」  
「極大点変化量 正規化あり」

## 第7章

### 総合考察

# 結論

## 目次

<b>序論</b> . . . . .	1
はじめに . . . . .	2
研究の背景 目的 . . . . .	3
音声の定義 (話声と歌声) . . . . .	4
ベルカント唱法と「理想」 . . . . .	8
<b>第1章 発声器官</b> . . . . .	12
1.1呼吸器官 . . . . .	13
1.2歌唱における呼吸 . . . . .	17
1.3声帯 . . . . .	18
1.4声門下圧 . . . . .	20
1.5声道 . . . . .	21
<b>第2章 発声の理論</b> . . . . .	23
2.1ソース・フィルター理論 母音の生成の仕組み . . . . .	24
2.2母音の生成 . . . . .	27
2.3「歌い手のフォルマント」 (Singer's formant) . . . . .	31
<b>第3章 「歌い手のフォルマント」の研究と課題</b> . . . . .	34
3.1 長時間スペクトル分析による研究の限界 . . . . .	35
3.2 「音質」の変化の少なさ . . . . .	37
3.3 観察される音声からソース情報とフィルター情報への分離 . . . . .	39
(STRAIHTのロジックとその有効性について解説)	
<b>第4章 歌声の音響計測実験</b> . . . . .	41
4.1目的 . . . . .	42
4.2実験方法 . . . . .	44

歌唱課題	45
歌唱被験者	47
録音方法	47
4.3特徴量の計算方法	48
分析対象の切り出し	49
4.4占有率	50
二つの方法による歌手のフォルマントの占有率の観察	51
比較周波数帯の算出と計算方法	52
4.5RMS変化量	53
RMSによる比較周波数帯の算出	53
2-4kHzにおけるRMS指標	54
極大点前後500HzのRMS指標	55
パワー正規化	56
<b>第5章 歌声に対する印象評価実験</b>	<b>58</b>
5.1目的	59
5.2実験方法	60
参加者	61
使用機材	61
5.3印象評価実験分析方法	62
<b>第6章 印象評価実験結果と音響特徴量の回帰分析</b>	<b>66</b>
6.1目的	67
6.2回帰分析結果	68
6.3考察	74
6.4まとめ	76

<b>第7章 総合考察</b> . . . . .	77
総合考察 . . . . .	78
結論 . . . . .	86
おわりに . . . . .	87
文献 . . . . .	88
本研究に関する研究業績 . . . . .	92
謝辞 . . . . .	93

## 序論



## はじめに

クラシック音楽において、優れた歌手の歌声とはどのようなものなのか。理想の歌声とは何か。

どの人も歌うための肉体的資産は、生まれつき普通に持っている(Husler,1987)。しかし、すべての人間が優れた歌手の歌声を出せるわけではない。とりわけクラシック音楽における歌手は、その「特別」な歌唱技術によって、自身の歌声の存在感を聴衆に示し、出来るだけ多くの人に音楽を届けることを目的としている。西洋において声楽が興って以来、その歌唱技術は、師から弟子へと受け継がれてきたものであり、現代においてもその伝統を継承し、さらに発展させてゆかなければならない。しかしながら、今日の声楽指導現場において、この「特別」な歌唱技術は何か、「理想」の歌声とは何かという疑問に対して、明確な答えを持つものは少ない。なぜなら、それは非常に感覚的なものであり、また指導者自身の主観的な見地に頼らざるを得ないからである。

近年、科学技術の発展により、人間の発声器官や音声、また音を聞くための聴覚などの研究が進展し、様々な研究がなされるようになった。歌声に関する研究も盛んに行われるようになり、その中で優れた歌手の歌声に共通する音響的特徴が見出されたのである。それは「歌手のフォルマント」と呼ばれるものであり、優れた歌手は、その歌声の周波数を、ある特定の帯域に集約することにより、より人間の耳に可聴しやすく、楽器などの楽音にマスクされにくい歌声にしているというものである。

## 研究の背景と目的

先行研究から見出された「歌い手のフォルマント」は、楽曲全体の長時間スペクトルから見出されたものである。そのため、楽曲において旋律や歌詞の変化に伴い、歌声の音の高さや音韻が変わっていく中で、「歌い手のフォルマント」が、ある一定の条件のもとで局所的に発生しているのか、歌唱全体を通して持続的に存在しているのかは解明されていない。

そこで、本研究では「音高」の変化と「母音」の変化に着目し、それらを伴う歌唱課題を作成する。そして、声楽家及び声楽を学ぶ学生の歌唱者によって収録された歌唱録音を、音響計測実験により観察し、「歌い手のフォルマント」の時間的なふるまいについて調べることにした。具体的には、歌唱課題における「歌い手のフォルマント」の「大きさ」や「変化」に注目し、それらがどのように変化するのか、あるいはしないのかということを観察する。そこから「歌い手のフォルマント」に関する音響特徴量を抽出し、さらに並行して行う印象評価実験より得られた、すべての歌唱課題に対する主観評価値との関係を相関分析する。

これらの実験によって、「歌い手のフォルマント」が歌唱技術にどのように関係し、またそれが聞き手の聴覚印象にどのような評価を得るのかということ、客観的なデータとして示す事が本研究の目的である。

## 音声の定義

本論文において、「音声」というものを定義するならば、どのように言えるだろうか？ Sundbergが自身の著書で「音声 (Voice sounds)」を、「肺からおくられた空気流が声帯による作用を受け、咽頭、口腔、場合によっては鼻腔により変化させられ、作られた音を音声とする (Sundberg,2007)」と述べおり、「音声」というものが、振動する声帯と声道に関連付けられていることを意味している。また人間が音声を生成する時に用いる様々な身体の構造物のことを「発声器官」と呼ぶ。これらの発声器官が相互に連関することで日常耳にするような様々な音声が生まれる。その中には人間が言語によってコミュニケーションを取るための「話声」や歌唱における「歌声」が含まれている。

まず「話声 (Speech)」とは、人間が言語によってコミュニケーションを取るために発せられる音声の意味を指す。また、音声言語とは音響的な符号のひとつであり、人間がある一定の集団の中で、思想、感情、意思などを伝達するために用いるものである。話声には様々な要因により、個別の音声信号の差異が存在する。その要因の一つが、発話習慣の違いである。例えば、言語の中にある母音をとってみても、日本語の母音と、他の外国語における母音では違いがある。あるいは同じ国の種族であっても、その個人の地理的および社会的に育った環境によって異なるのである。

また、もう一つの違いを生む要因が、その個人のもつ「発声器官」の特徴である。「発声器官」は、その個人の発する音声の「音色」を決定させる重要な器官であり、その個人の持つ「発声器官」の特徴の違いが、「音色」の違いをもたらす。

例えば、我々は声を聞くことで多くの場合、その話者が男性か女性かを判断することができる。一般的に女性の方が「高い」声であり、男性の方が「低い」声である。これは男性と女性では発声器官である喉頭や、声道の大きさが異なるためであり、その違いにより発せられた声は異なるのである。男女の差による声の高さは喉頭の大き

さの違いに起因する基本周波数の違いで、音色の違いは声道長の違いに起因する共鳴周波数の違いに由来する。一方で同じ性別であっても、個々人の発声器官の細部の違い（形状的な差）によって、異なる音声が生産される。例えば母音の違いは舌の位置や口唇の開口度の違いにより作られる。歌唱の際のピッチの違いは喉頭音源の基本周波数を変えることで実現される。これら二つの性質において、学習によって獲得可能な後天的な側面が発話習慣と呼ばれ、発声器官の大きさなど先天的な要因による違いとは区別される。具体的に発声器官の男女差は、身体の発達段階に従って徐々に現れ、第2次性徴を経ることによって顕著に発現する。

では「歌声」とはなんだろうか。それを論じるには、まず「話声」と「歌声」の違いは何なのかということを考えることが問題解決の糸口として有効であろう。「話声」と「歌声」はそれぞれ別のものとして研究されてきた。しかし、その間に明確な一線を引くことができるのであろうか。Husler(1987)は「話すこと」と「歌うこと」という二つの現象には、はっきりした一線を画しておくべきであるとしている。

近年では、歌声特有の音響特徴量について様々な研究がなされている。例えば、歌声にしか存在しないヴィブラート、また歌声に顕著に現れる特有の周波数群などである。また「話声」と「歌声」と言う線引きは、人間の心理的な要素にも依存している。私たちは日常生活において、その音声の話声なのか、歌声なのかを聞き分けることが出来る。齋藤(2008)らの研究では、歌声らしさの心理的特徴として、音声の「揺れ」「響き」「明瞭さ」の知覚が、歌声として知覚することに大きな影響を与えていることが分かっている。もちろん、オペラにおけるレチタティーヴォなどの例外はあるが、私たちは知覚的な特徴によって、その音声がいずれかであるのかを判別しているのである。

「歌声」は人類の長い歴史の中で、いつ生まれたのか。Huslerは「言語を話すことは人間の第1の本性に対して設えられた創造物である」と唱え、「言語の根源は、人間

が目で知覚したことの語源的な構造に由来し、大脳の聴覚領域よりも視覚領域においてははるかに多く成立した」と考えた。そして「話をするために設けられた機構よりも、はるか以前に歌うための機構があり、それはことばを作ろうという意思に占領されてしまった可能性が大いにある」と述べている。つまり人間は、言葉を持つ以前から歌声を持っており、人間が言語を使って言葉を発するようになるにつれて、それは衰退していったのではないかというのである。もちろん、ここでいう歌声は、今日私たちが耳にしている歌詞や旋律を伴った歌唱の歌声ではなく、「幼児が話し始める前に旋律的な、あるいは歌唱的ともいえるような声の放出、純叙情的な気分の発散」

(Husler,1987)を意味し、人間の歌声を、その内なる感情や意思の表現と捉えるならば、言葉よりも先に「歌声」というものが存在したと考える方が、自然であるとHuslerは述べているのである。また、自然界における他の動物に目を向けるとどうだろうか。

Huslerの説に従えば、歌声とは動物の鳴き声のような生来的な要素から派生したものであり、同じものであると短絡的に考えられるが、実際の歌声は決してそうではない。

歌声と鳴き声とは違いがあり、歌声は鳴き声にはない音響的な特徴を持っているからである。「歌声」の最大の特徴は、音声の高さである基本周波数の自然な下降がないということであり、そのような鳴き声を持つ動物は稀である。詳しくは後述するが、

基本周波数を一定に保つという行為には、声帯の緊張も一定に保ちながら一定の呼気圧を保つというきめ細かい制御が必要である。話声の自然な呼気圧制御では、発話開始から終了までその基本周波数が緩やかに下降する。それは呼気圧が『自然に』低下

していくからである。しかし、歌声では通常一定の高さを保つことが要求される。つまり話声とは異なる制御が必要なのである。そう考えると、「鳴き声」＝「歌声」と

いう考え方には、矛盾が生じるのではないだろうか。確かに私たちは時として、「話声」からでは得ることのできない感動を「歌声」から得ることができる。その理由の一つとして「歌声」が、本来「鳴き声」の果たしていた役割（制限なしに感情や意思

を表現すること) を果たしているからなのかもしれない。北村(2009)らの研究によって、発声器官の中で、声の音質や音色を決定させる声道は、人間の情動によっても変化することが分かっている。つまり、人間の声と感情には大きな関わりがあるということなのである。

## 2. ベル・カント唱法と「理想」

ベル・カント (Bel Canto) とはイタリア語で、美しい歌唱という意味で、一般的には、17世紀や18世紀にオペラにおける歌唱様式そのもの、あるいはその様式を支えた発声の手法のことを指す。特に、イタリア的な歌唱法や歌声 (明るく輝かしい声など) に対して使われることが多く、賛賞の言葉として用いられる。またG.A.ロッシーニ (1792-1868)やG.ドニゼッティ (1797-1848), V.ベッリーニ (1801-1835)が作曲したオペラは、ベル・カントオペラと呼ばれ、ベル・カント唱法によって歌われるべき作品であるとされている。

17世紀にオペラが創案されると、声楽的な表現と歌手たちの妙技を発揮させる場と機会が提供された(Reid1950)。それ以前の声楽音楽はポリフォニー様式であり、厳格な音楽の決まりがあった。しかしオペラによってその決まりから解放された歌手たちは、発声のテクニックや、声の魅力を競い合うようになり、A-B-A形式のダ・カーポアリアが導入されると、歌手の持つ声楽的な技量をいかんなく発揮させ、即興演奏によって音楽的な趣味の良さを表現させる(Reid1950)ために、第3部分は歌手の自由な装飾歌唱が行われるようになった。その過程の中で、ベル・カント唱法が確立されていったのである。このことからベル・カントの時代は装飾歌唱の時代と言い換えてもよく、ベル・カントの本質は装飾的歌唱法そのものである(水谷2006)とも言われている。しかしその一方で、ベル・カントはイタリア的、イタリア性といった意味でも用いられ、19世紀以降のG.ヴェルディ(1813-1901)やG.プッチーニ(1858-1924)などのオペラ作品に対しても、ベル・カントという「理想」としての考え方は生き延びている。今日でも声楽発声における理想的な歌唱法、またはその歌声自体を指す言葉として用いられている。

以上のようにベル・カントは様々な解釈や意味を含んだ言葉であり、厳密な定義は存在しない。そして、それが本来どのようなものであったのかということは、現代で

は録音資料の残らない時代の産物であるため、文献や書簡を基に類推することしかできない。また、音楽の現場ではその実体が時代によって変遷しているにも関わらず、その言葉が安易に使用し続けられている傾向がある。筆者自身は声楽を専門とする演奏家であるが、筆者が知る中でも、様々な歌唱法やその流派が存在し、実際に多種多様な歌声を、今日の演奏会で聞くことができる。その中で、どれが一番理想的なものであるか、筆者が述べたとしても、それはあくまでも筆者自身の主観的な思考や好みによるものにすぎない。それと同じように、すべての演奏家は、発声の技能や歌声について「理想的である」と断言しても、それは客観性を帯びることはできない。このように各自の主観的な理想が存在する現代の音楽界において、ベル・カントという概念は、その理想的形態として定義が不確定なまま放置されている。

歌声の「理想」像の実態を、直接観察ができないまでも、その「理想」が生まれるに至った背景を見ることで実像に多少とも迫ることが出来るかもしれない。

オペラという「見世物」が流行し、歌い手は頻繁に大観衆に対して、オーケストラをバックにして歌うことが要求されていった。おそらく、17世紀から18世紀にかけて(現在もそうであるが)、様々な発声法や歌唱法で歌唱する歌手が存在したに違いない。しかし、全ての歌手が、その声の質を長期にわたって維持できたわけではなく、医学が進歩した現在では、ある程度の治療が可能であろう発声器官の故障や、疲弊も、当時の歌手たちにとっては致命的なものとなりえたことは想像に難くない。その為、歌手には、芸術的、技巧的な発声や歌唱のテクニックだけではなく、なるべく発声器官にかかる負荷を軽減しながら、より効率的に、それらを行うということが求められたのであろう。その過程でより問題を起こしにくいものが自然淘汰されてゆき、それらの歌手たちの持つ発声法や歌唱法がベル・カントと呼ばれ、「理想」とされていったと考えることには一定の合理性がある。また、聴衆もベル・カントという「理想」をつくりだすために大きな役割を果たしてきた。なぜなら、先に述べた芸術的、技巧的な



歌手のテクニックを評価するのは、歌手本人ではなく、むしろ聴衆だからである。何を「美しい」と感じるかは時代による変遷も大きいですが、少なくとも言えることは、他の歌唱者では達成できない声の質を（当時としては）備え、際だった存在としての地位を得なければならなかったことであろう。「普通の人間には出せない高い声を出すことができる」「良く聞こえる声である」「長いフレーズを一息で歌うことができる」などの指標によって、歌声を評価することは特に専門家でなくても判断のできる分かりやすい知覚的特性である。そして、その様々な「指標」において、それらを達成した歌手の声を「理想的」とであるとされた可能性は高い。

しかし、ではその「指標」は本当に正しいのだろうか。あるいはその「指標」をどのように定めるべきなのであろうか。「高い声」とはどのくらい高い音を指すのか。歌手が出すことができる最高音は、性別や声種によって異なる。また同じ声種であれ、その歌手の持つ音域はそれぞれに異なるはずである。その中で最も高い声を出したものが最も優れているのだろうか。あるいは「大きい声」といっても、それはどのような建物の中で、どれくらいの距離で聞いている時の声なのだろうか。あるいはピアノ伴奏、オーケストラ伴奏の違いによっても、歌声の聞こえ方は変わってくる。また「長いフレーズ」といっても、歌われているピッチや、その歌声の音量、声質によって変動するものである。

また、「理想」という言葉に疑問を抱かせるもう一つの要因は、その伝承のされ方にある。先に述べたが、現代には様々な歌唱法やその流派が存在する。ベル・カントが確立されて以来、その伝承は指導者自身の経験と主観的な印象に基づいて行われてきた。Reidは自身の著書中で「昔の教師たちは、このような共鳴に関する音響学的な法則については何も知りませんでした。しかし、彼らはその音が間違いなく共鳴した音質であるかどうかを耳で聴き分けて、的確に判断していたに違いありません」

(1950). と述べている。確かに、優れた演奏家、あるいは音楽教師は共鳴の程度を聞

き分ける力を持っており、それによって声の良し悪しの判断をすることは可能であったと思われる。歌声の音質は、発声器官のその時々の状態を示す、あるいは状態そのものであると考えられ、教師は聴覚的に捉えた歌声に対する主観的な判断に従って指導を行っていた。そして、その構図は、今日の歌唱指導や音楽教育の現場においても同じである。やはり教師自身の経験と主観的な印象によってつくられた曖昧な「理想」を、学習者は追い求めなければならないのである。Reidもこの構図については、

演奏家として名をなすような歌手たちで、発声訓練のメソッドの歴史や発展についての精密な研究をする時間や精力を持てるという人は、めったにいないものではない。レパトリーを準備し、公演し、社会的な責任を果たすという多忙な生活には、普通、こうした特殊な研究をする余裕などないものです。そのため、第一線で活躍してきた歌手たちであった人たちが、人に教えるということに就く時がやってくると、<自分たちの>先生から教わった発声の原理だけを教え、自分たちが生徒として経験した指導手順をそのまま繰り返すことになり、誤りはとめどなく犯され続けられることになります。・・・演奏家としてはいくら偉大であっても、このような恵まれた条件のもとで訓練された歌手たちに、はたして発声のメカニズムについて何を教えることができたでしょうか。(Reid,1950)

と述べ、理想像を曖昧なままに放置している現状を危惧しているのである。

## 第1章 発声器官

## 第1章 発声器官

この章では、過去に行われた研究の中からSundberg(2007), Bordenら(2004), そして今泉(2007)の著書を基にし, 声楽教育や発声指導において行われてきた指導法や筆者の考察を交えて, 発声器官とその働きについてまとめる. 後の章ではSTRAIGHTと呼ばれる分析手法に則った分析データを用いている. この分析手法の妥当性を理解するには音声のソース・フィルタ・モデルに対する基礎知識が必要となり, そのため音声の生成や発声の理論に関する基礎知識を提供することがこの章の目的である. 発生に関わる器官のことを発声器官と総称する. しかし, 個別の器官の本来の機能は必ずしも発声をするためのものではない. 呼吸器官は文字通り, 生命維持のために必要な呼吸をするためのものであり, 声帯をその一部として持つ喉頭の本来の目的は肺に飲食したものが侵入しないための器官である. さらに人間は舌の位置や唇の開閉を調整して母音の違いなどの細やかな声道形状の変化を産むような進化を遂げたが, これも人間が直立二足歩行を選択し, その一方で成熟しても口吻が長く前に突き出さないという形態を持つに至った結果である. これらの総合的な符合はあたかも人間が音声を獲得するために進化したような印象を与える. しかし, それが目的なのか結果なのかについての議論はここではする必要はないであろう. 以下には, 表層的に音声の生成に必要な各部位がどのような構造を持ち, どのような機能を果たしているかを説明する.

### 1.1 呼吸器官

人間が音声を発する際には, 肺から出た空気が声帯を通過する必要となる. これを呼気という. 呼吸は生命の維持に必要な動作があり, それは肺に大気を取り入れる吸気の相と, 肺で酸素を摂取して不要となったガスを排出する呼気の相が存在する. 原理的には吸気流を使って声帯を振動させることも可能であるものの, 大半の場合の発

声はこの呼気相で生じる呼気流が使われる。この項では、一連の動作である「呼吸」を「呼気」と「吸気」に分けて考える。それは「呼気」と「吸気」には、それぞれ使われている身体の筋肉群や機構が異なり、動作における物理的な作用も異なるからである。

呼吸器官は、酸素を体内に取り込む肺があり、その周りを囲むように胸郭がある。肺そのものは筋肉もなく、スポンジのような構造であり、肺の動きや形の変動は、その外枠の胸郭と横隔膜によって受動的に決定される。ただし、風船と同様に肺自体に弾性があり、ある一定以上に膨らませると自然と収縮する。胸郭は肋骨柄、胸骨、肋骨、肋軟骨とからなり、肋骨の間にある筋肉によって、全体が拡大したり、縮小したりすれば、肺の容量もそれに比例して変化する(図1-1)。

また、横隔膜とは、肺の下に位置するドーム型の筋で、上下に動くことによって、肺の形状を変化させている。

まず、「吸気」時には声門が開いて外気を肺の中に流入させる。肋骨の間の筋肉や、横隔膜などの吸気のための筋肉である吸気筋の活動によって肺の中の圧力が大気圧と等しくなり平衡状態となると「吸気」が完了する。「吸気」時には、肋骨の間にある軟骨にあたるものが収縮し、肋骨を持ち上げるように働く(軟骨には動力はなく、動力源はあくまでも肋間筋などの筋肉であり、筋肉は収縮することでしか力を出すことができない)。そうすると、胸郭は前後方向にも、側方向にも拡大する。胸郭が拡がると、肺の内部が相対的に陰圧となり、外気が肺に流入していく。そして、肺の容量が一定より大きくなると、その量を維持するために、吸気筋の緊張を持続させる必要がある。

そして「呼気」時には、先に述べた吸気筋の緊張を緩ませると、肺と胸郭の「弾性復元力」のため、肺内の空気は急速に呼出される。またそのほかにも、肋軟骨のよじれが戻ろうとする「ねじれの復元」と、拳上した肋骨を下げようとする「重力」の作

用も働く。人間が無意識の呼吸に行ってる安静呼吸では、出て行く空気の量は少なく、一回の呼吸において全肺気量の約7%しか出入りしていない。

また、発話時においては、安静呼吸とは異なり、より吸気筋の作用を強め、そして無意識のうちに行っていた呼吸を、意識的にコントロールする。また、吸気に使う時間もより短くなり、呼気の時間的な割合を増やす。また、歌唱時においては、呼気における能動的な力だけでは不十分であり、持続的で能動的な呼気調節が必要となる。最初のうちは吸気筋を使い、一定以上の呼気を使うと、呼気筋の力を使う。吸気筋は、呼気を歌唱に必要な容量をゆっくりと呼出するために作用し、呼気筋は、弾性復元力によって胸郭の容量がもとに戻ったあとも、さらに容量を縮めて呼気を続けるために働くのである。呼気と吸気の力が等しくなる肺気量のことを機能的残気量 FRC(functional residual capacity)と呼び、肺がこのFRCから外れて拡大または縮小をすると、先に述べた作用によって元の状態に戻ろうとする受動的な力が働く。

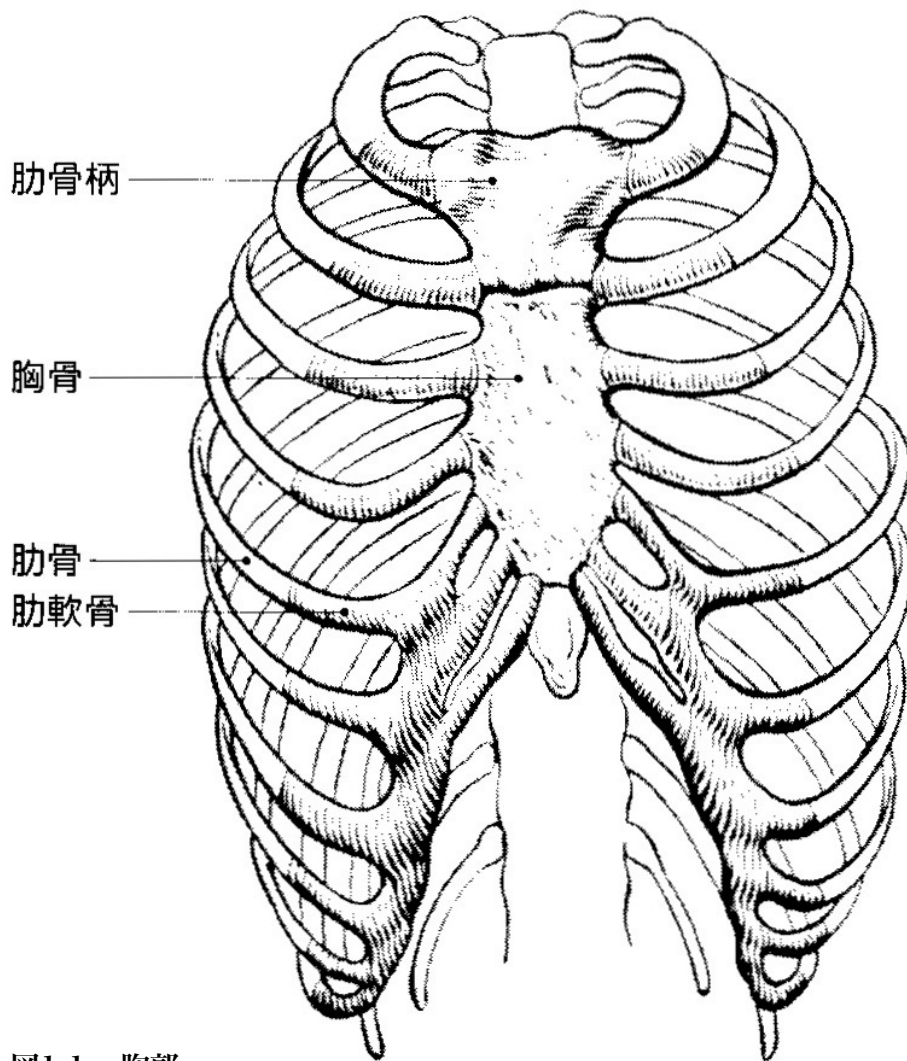


图1-1 胸郭  
(Gloria J. Borden, 2004)

## 1.2 歌唱における呼吸

現在の声楽教育や発声指導においては、「腹式呼吸」という言葉を代表に、呼吸の方法が大きく歌声に関わっているとされている。「歌唱時に側腹部を広げたまま維持すべきである」「横隔膜を上方に押し上げるように歌唱すべきである」反対に「横隔膜を下方に押し下げて歌うべきである」など、人間の生理的な作用と、イメージがかけ合わさった表現を用いられることが多く、いずれの表現も、歌唱時におけるブレスコントロールの重要性を指摘している。しかし一方で、Reidのベル・カントの研究によると、17世紀から18世紀における声楽教師たちの見解は「呼吸作法というのは、根本的な発声の原理が確立されると、自然にそれと活動を共にするものだ。」と考えられていた。そのため、当時の声楽教師たちは呼吸の仕方に特別な追及をしていなかったのである。また歌唱時における呼吸法の研究では、腹壁を外側に広げる方法を「ベリーアウト」、内側にへこます方法を「ベリーイン」と呼び、両方の場合における呼吸器官の動作を観測することによって、どちらが歌唱にとって適切な方法であるかを検証する実験がされている。(Sundberg,2007)

「腹式呼吸」に対して、悪い意味で使われて来た「胸式呼吸」という言葉がある。「胸で息は吸ってはいけない」などという指導はよく耳にするが、実際は歌唱において胸郭は積極的に参加している。このように歌唱における呼吸法には、様々な見解があり、さらなる研究の必要性が窺えるが、本論文において、呼吸法の是非をこれ以上言及することはしない。しかし少なくとも、先述したように歌唱には安定した呼気の供給が必要であり、それを可能にするためには、受動的な安静呼気とは異なり、持続的で能動的な呼気調節が必要となる。

実際に、歌い手の一回の呼吸における歌唱時間は長く、最近の研究ではFRC値の5～10%の肺気までを使用していることが判明している。つまり、歌い手はより肺を縮ませて呼気圧を保つ必要に迫られるのである。また安定した呼気を実現するためには、



呼気時において呼気筋と同時に吸気筋を使わなくてはならないと述べたが、その点については声楽教師であるHuslerも「横隔膜は歌唱呼気の時も、その固有の傾向、すなわち吸気傾向を捨てない。」と述べており、歌唱には呼気筋と吸気筋の協力が必要であり、高度な呼気調節が必要であるということが窺える。

### 1.3 声帯

声帯は人間の喉頭の中にある甲状軟骨の内部にあり、前後方向に走る弾力に富んだ組織で、筋肉、靭帯および粘膜からできている(図1-2)。

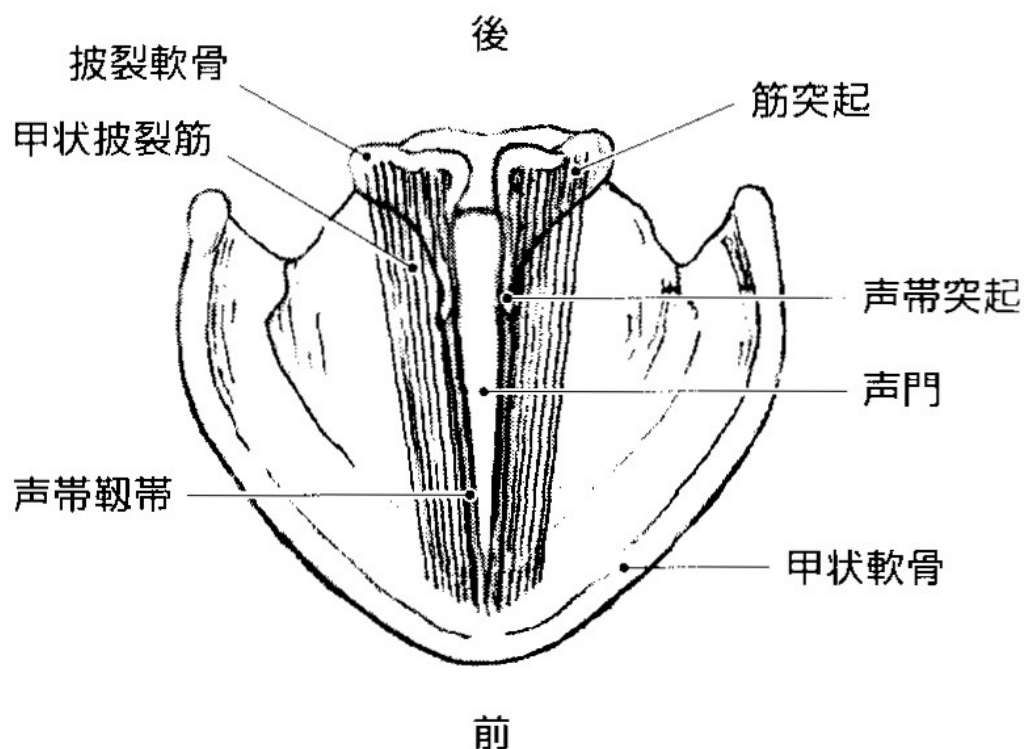


図1-2 上方から見た声帯

(Gloria J. Borden, 2004)

声帯自体の弾性や張力は変化させることが可能であり、その厚みや長さの形状も変化させることができる。二枚の声帯は左右方向に開閉し、その度合いも調整することが可能であり、甲状軟骨自体を上下に動かすことによって、声帯も上下方向に移動させることができる。この声帯の様々な動きによって、声の大きさ、高さ、そしてその音色などを変化させることができるのである。呼吸時には、声帯は左右方向に開いており、呼気と吸気が通過するだけである。この左右に開いた声帯の隙間を声門と呼ぶ。声門が中心に向かって閉じている時に、肺から出た呼気が声帯を通過すると、声帯は呼気圧の空気力学的な要因によって開き、次に声帯そのものの弾性と、声門を空気が通過する時の吸引力によって閉じることを繰り返し振動する。この振動を声帯振動と呼ぶ。先に述べた声門を空気が通過する時の吸引力はベルヌーイ現象と呼ばれる。ベルヌーイ現象とは、ある物質が狭いところを通過すると、その物質の速さは増し、また周りの物質との圧力の変化により吸引力が働くのである。運転している自動車の横を、大きなトラックが追い越してゆくと、トラックの方向へ自動車が吸い寄せられるのもこの原理と同じである。

また、物質が振動すると音になることは周知の事実だが、先に述べた声帯振動が声を生み出すのである。声帯は、開いて呼気が流れる相と声帯が閉じて呼気流が堰き止められる相が交互に訪れることで、声門の上に気圧の疎密の状態が交互に存在し、この疎密波こそが音の物理的実態である。これを「喉頭音源」と呼ぶ。喉頭で生じた疎密波はその後声道を伝わり、この過程で必ず声道の共鳴を反映した変形を受け、大気中に放射される。この大気中に放射された音が我々の耳に到達し、音声として知覚される。

## 1.4 声門下圧

歌唱時における呼吸の重要性と、声帯振動について先述したが、ではなぜ歌唱には安定した呼気が必要なのであろうか。それは呼気によって生まれる声帯の下部の圧力、すなわち声門下圧が、声帯振動に大きく関わっているからである。まず、呼気によって声門下圧が上昇すると、声門が押し開かれる。その瞬間にまた声門は吸引力によって閉じ、気流は止まる。この繰り返しにより、声帯振動が起こるのだが、この必要条件として声門下圧が声門上圧よりも高くなければならない。声門下圧と声門上圧が平衡に達すると、声帯振動は止まってしまうのである。発声の周波数は喉頭筋と声門下圧の両方によって決定される。また、声門下圧は声の大きさに関係し、一般的に声門下圧が高まると、声の大きさは大きくなる。また、声門下圧は、歌声のピッチにも大きく関係している。ピッチ自体は、声帯の長さ、張力に依存するが、一般的には声が高くなるにつれて声門下圧も上昇するとされている。しかし訓練された歌手では、必ずしもピッチの上行と声門下圧の上昇が比例するわけではなく、声門下圧を制御することによって、歌声の大きさやピッチをコントロールしていると思われる。また、声楽教育や発声指導においては「声帯を鳴らす」「声の鳴り」などという表現で、喉頭筋と声門下圧についての、技術的なアプローチがなされている。Husler(1987)は声の出し始めのことを「アインザッツ(Einsatz)」と呼び、声の出し始めは声帯を傷つけない方法で行われるべきだと述べている。間違った発声指導でよくありがちなものは、声を出し始める前に、ぴったりと寄せ合った声帯の下に、空気をせき止める「声門打撃」という方法や、反対に声門を開きすぎる「気息的」な方法があり、どちらも偏った方法で、推奨すべきではないとしている。科学的な見地から言及するならば、発声には喉頭筋と声門下圧のバランスが重要であると言える。

## 1.5 声道

声道とは、声門から口唇までの喉頭より上方にある空間のことを示す。その中でも主だった空間は、咽頭、口腔であり、そして場合によっては鼻腔が含まれる。また喉頭腔や、気管内の空間も声道の一部として考えられている(図4).

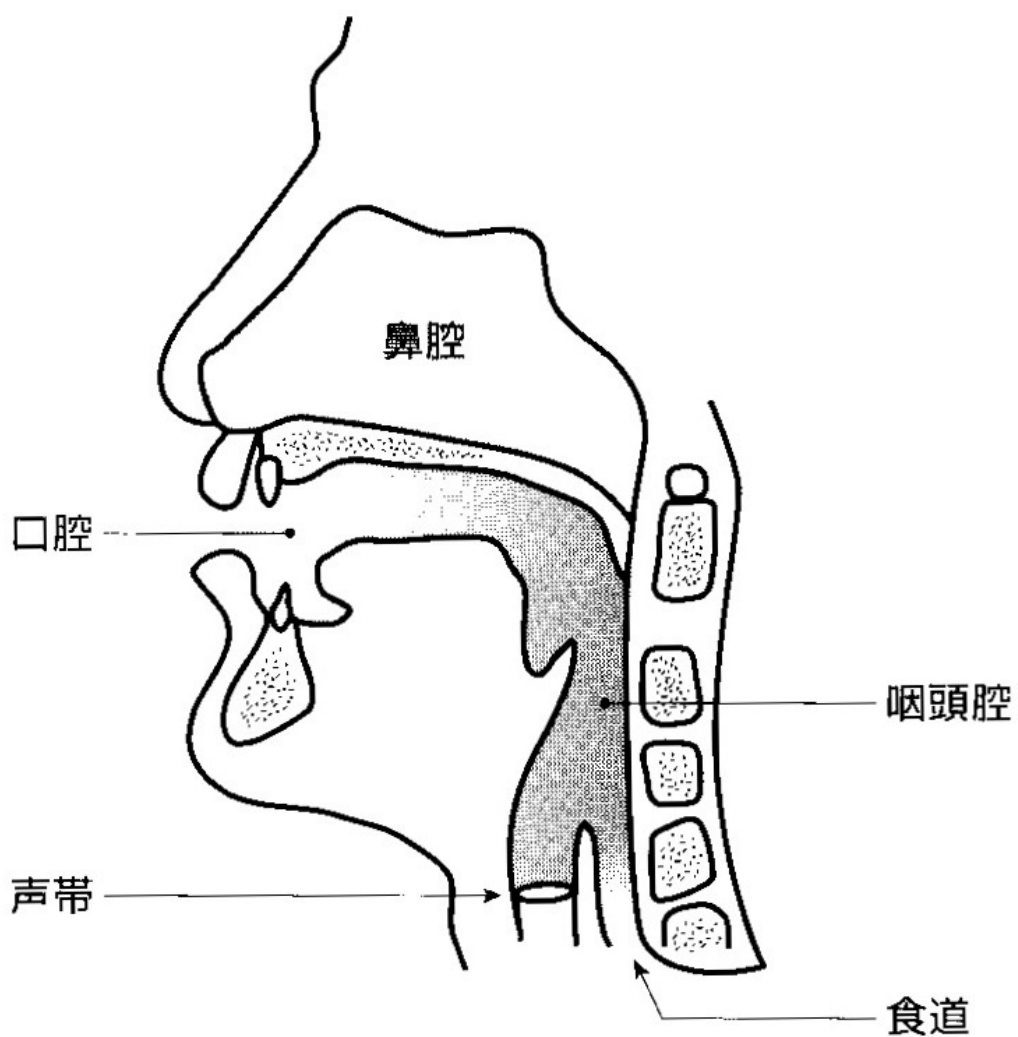


図1-3 声道の断面 (Gloria J. Borden, 2004)

咽頭は声道の奥の方の部分で、筋肉で覆われた管でできている。口腔は私たちが自分の目で観察することができ、口腔の奥の上方に軟口蓋がある。軟口蓋を挙上させると、鼻腔への連絡が閉ざされ、垂下させると口腔と鼻腔が繋がる。また、口腔の中には舌が存在する。舌は筋肉のかたまりで、3次元的な動きをすることかできる。また、舌は外から観察できる部分は、一部分で、実際は舌根と呼ばれる咽頭の奥の部分まで繋がっている。また口唇は、声道の一番外側の部分であり、後述する声道の共鳴特性に重要な役割を担っている。

## 第2章 発声の理論

## 2.1 ソース・フィルター理論 (Source-filter Theory)

発声器官は呼吸器官、声帯、声道の3つの単位で構成されている (Sundberg,2007) .  
呼吸器官である肺から出た空気が、声帯の声門を通過することにより、喉頭音源  
(voice source) になり、その喉頭音源が声道によって調音され音声となる。

Sundbergは呼吸器を圧縮機、声帯を発振器、声道を共鳴機という工学的な概念に置き換えて、それぞれ独立した器官を、その機能で考えることにより、音声の発生の仕組みを説明している。(図2-1)

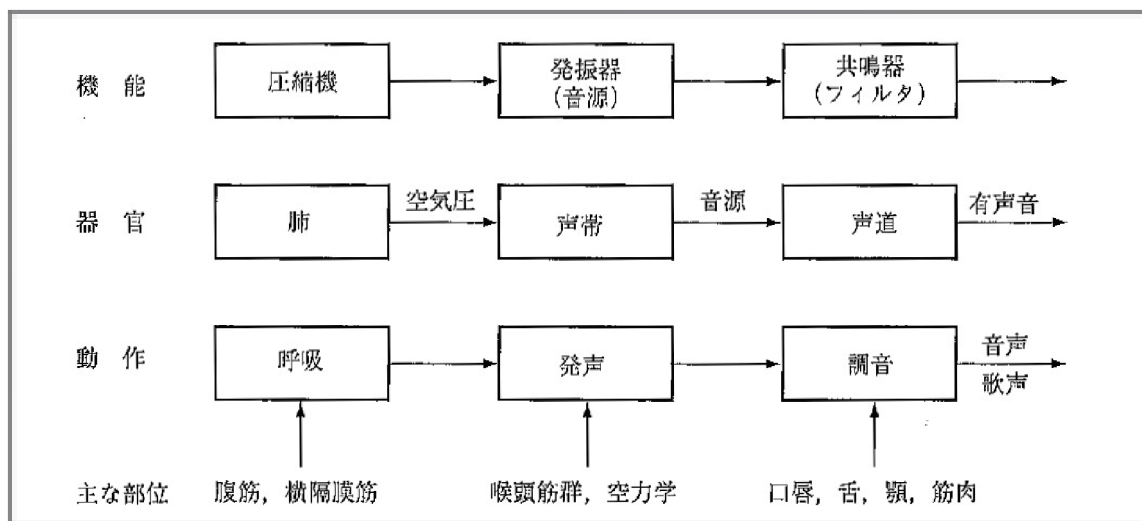


図2-1 ソース・フィルター理論

呼吸器官は加圧状態の空気を供給するコンプレッサとして働く。発声器官は声門を通過する空気流を音源となる空気のパルス列へと変換する発振器として動作する声帯からなる。声道は共鳴器として動作する。(Sundberg, 2007)

このような実体を音源とフィルターの2つの要素によって構成されるとモデル化できるとする理論のことを「ソース・フィルター理論」という。

先述したように、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

源スペクトルに、声道の断面積から算出された声道伝達特性が掛け合わされることにより、歌声のスペクトルがつくられる。簡略化した図を(図2-2)に示す。

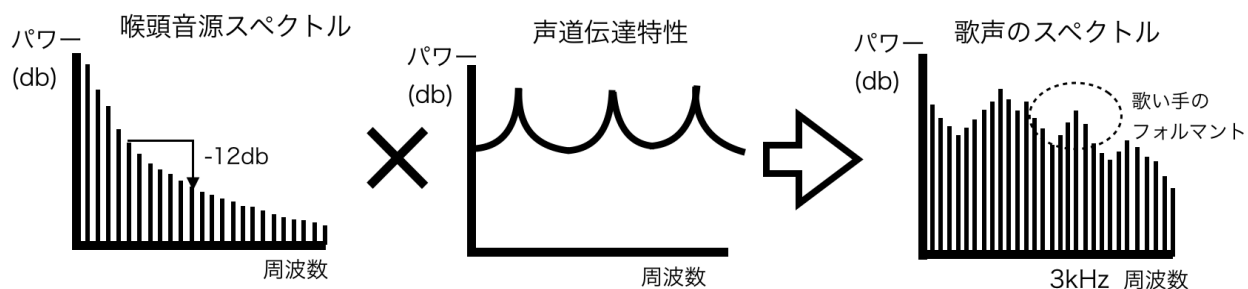


図2-2 歌声のスペクトルができる仕組み

声帯から発信された喉頭音源スペクトルは、声道伝達特性によって調音されて歌声のスペクトルになる。

声道は声門から口唇までの喉頭より上方にある空間のことを指し、一端のみが開いた管として考えられ、その共鳴は閉管の共鳴と考えられている。普通の管とは異なり、表面は粘膜で覆われ、その断面積も口腔の開閉具合や、舌の形によって様々に変化する。声道の形状は、声の音色を決定づけるものであり、声道の形状の変化に伴って、共鳴管の共振特性が変化し、そのために周波数スペクトルが変化する。また実際には、喉頭音源と声道の共鳴という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する。声帯の張力や厚みなどの物理特性が変化することで、呼気圧によって励振する声帯振動の基本周波数が変化し、これが声の高さの変化を引き起こす。声帯が引き伸ばされて薄くなり、張力が高まると喉頭音源の周波数は高くなり、反対に短く分厚くなると、喉頭音源の周波数は低くなる。喉頭音源が高くなると、実際の声も高くなり、低くなると声も低くなるのである。例えるなら弦と同じ原理であり、細くて張力の高い弦は高い音が鳴り、太く



て張力の低い弦は低い音になる。また、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。

声帯は肺から供給された呼気によって声門下圧という力を受ける。呼気の圧力が強まるか、あるいは声門の閉じる力が強まると、声門下圧も上昇する。声門下圧が上昇すると、声の大きさや、高さも上昇する。Sundbergの研究において、声門下圧と声の大きさや高さとの比例関係は証明されている。しかし訓練された歌手では、一定の歌声の大きさで上昇する音階を、呼気の気流量をほとんど増加させることなく歌い、逆にこの状況で、呼気の気流量が減少することも観察されている。

## 2.2 母音の生成

異なる母音を発話する際には舌（前舌，後舌，上舌，下舌）の位置が変わる。それは声道の形状が変わることを意味し，その形状の変化は声道断面積関数として記述可能であり，それに従った管を作ることによって母音の区別に必要な音響変化を算出することができる。つまり，母音の音韻性や音質は，声道におけるフォルマントと周波数によって決定されるのである。先述したように歌声を音響的に分析すると，フォルマント(Formant)と呼ばれる周波数群が存在する。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のだが，この第1フォルマントと第2フォルマントが，母音を生成するために重要な要因となる。

声道とは声門から口唇までの喉頭より上部にある空間のことであると先述したが，母音の種別に関わる部分は，主に口唇に近い側の声道の断面積と咽頭腔の断面積，つまり口腔内の面積と舌の位置の関係で決定されている。母音が変わるということは，声道断面積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。このような音響管の形状の変化が声道伝達特性の変化を引き起こし，喉頭音源に対してフォルマントの異なる組合せを生じるわけである。

これらの物理的变化はお互いに影響を受け，それに伴う喉頭の位置や，声道の形状の変化することにより第1フォルマントと第2フォルマントが変化する。(図2-3)

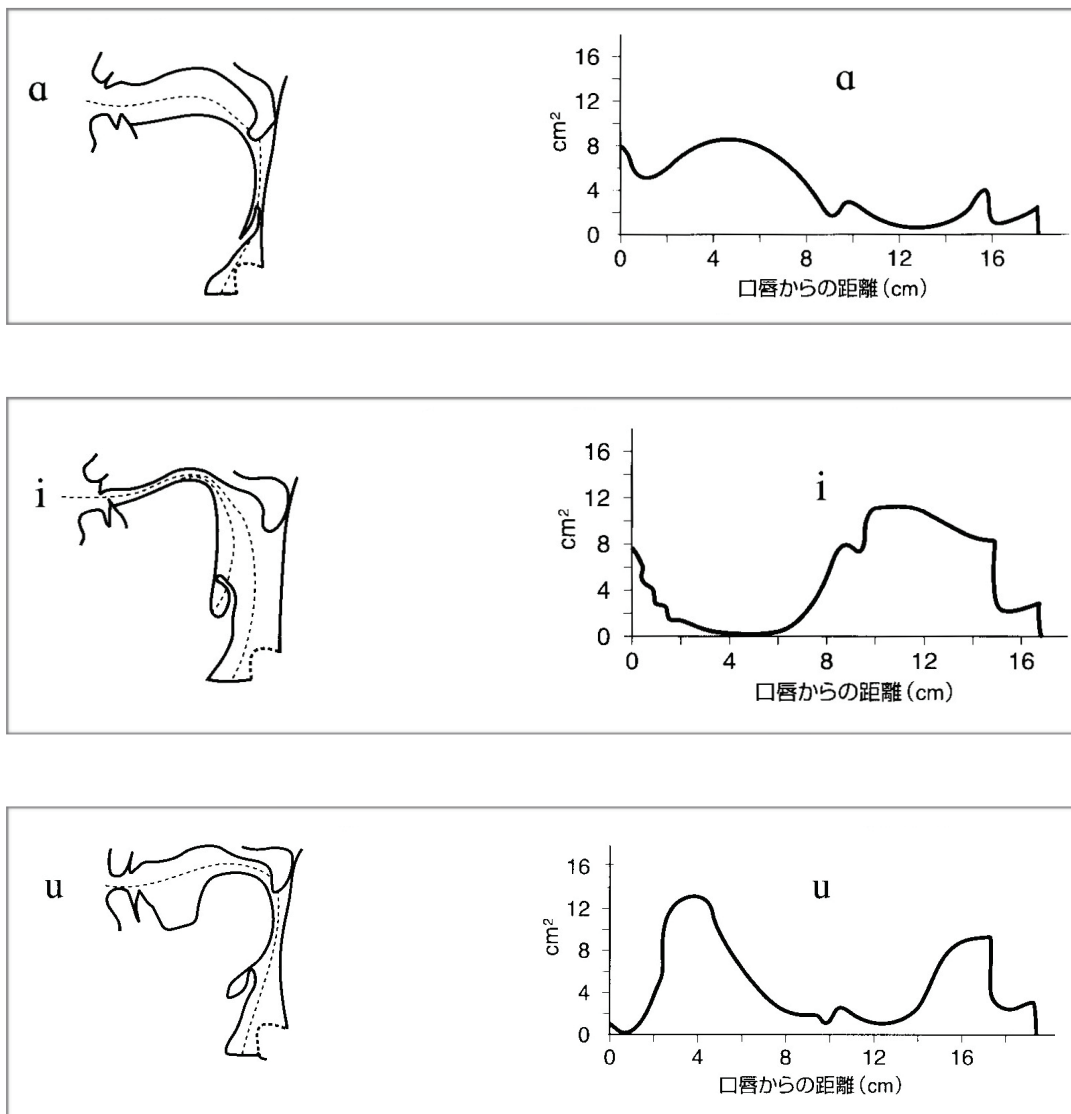


図2-3 「ア」「イ」「ウ」における声道の形状と声道断面積

(Gloria J. Borden, 2004)

Sundbergの研究から第1フォルマントは、主に顎の開きの変化によって決定され、第2フォルマントは、舌の形状の変化に大きく影響しているとされている。例えば、口を縦に開いて舌の形状を低い位置で平らにして声を発すると、「ア」という母音になる。これは、第1フォルマントが800Hz付近にピークを持ち、第2フォルマントは1000Hzにピークを持つようになることから、「ア」という母音になる。次に、舌を高い位置にし、顎を比較的閉じて、咽頭を広げるように声を発すると、第1フォルマントは300Hz付近にピークを持ち、第2フォルマントのピークは2500Hzまでピークが上昇する。この第1フォルマントと第2フォルマントの関係が「イ」という母音を生成する。具体的なフォルマントのピークの周波数域には個人差があるが、概ね（図2-4）のようなフォルマントの関係性によって母音が決定される。

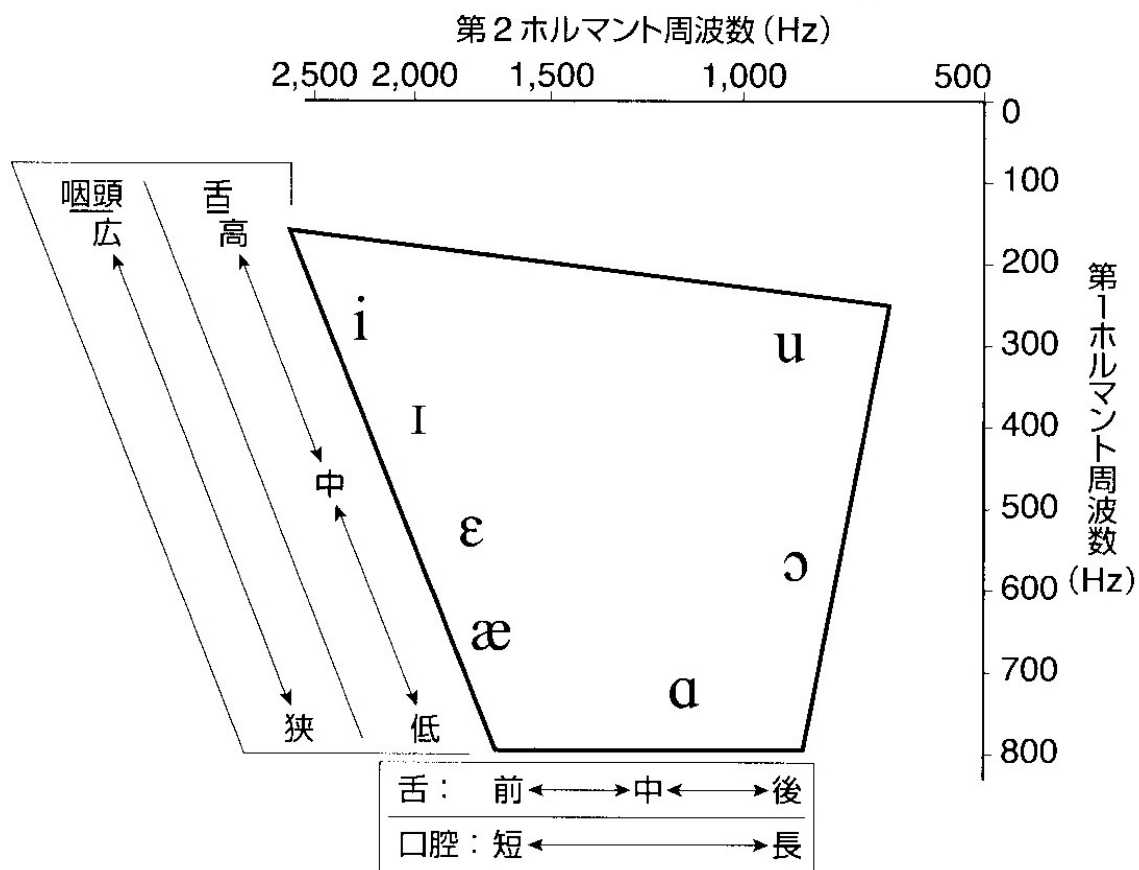


図2-4

様々な母音についての舌の位置と、声道の大きさ、第1、第2フォルマント周波数との相関関係図  
(Gloria J. Borden, 2004)

### 2.3 「歌手のフォルマント」

科学者や研究者などによって、歌声に関する様々な研究がなされてきた。発声器官やその機能、そして発せられた歌声を生理学、音響学的に研究し、その演奏技能の解明や、歌声の分析がなされてきたのである。そして近代においては、科学技術の発展によって、人間の発声器官や音声、また音を聞くための聴覚などの研究が益々盛んになり、より正確で精密な研究がなされるに至っている。その研究の中でも、歌声について「理想的な声」の一つの指標となりうる音響的特徴が発見された。それは、「歌手のフォルマント」(Singer's formant)という音響的特徴である。

歌声に限らず「音声」は有声音と無声音に大別できる。無声音とは、フ・ク・チ・プ・シュ・ス (f/k/ch/p/sh/s(ce)) などの、声帯の振動を伴わない子音のことを指し、音響信号として周期性のない音である。反対に有声音は、全ての母音と声帯の振動を伴う子音 (b/d/gのような破裂音、v/zのような摩擦音、m/nのような鼻音、そしてl/rのような流音) を指す。有声音のうち、母音と鼻音は周期性を保てる。この周期性に対して我々はピッチを知覚する。旋律とはそのピッチの変化によって表現されるため、歌唱音声を考える上でまず周期的な母音を対象とした研究がなされた。先述したように、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。母音という範疇に入る音にはフォルマント(Formant)と呼ばれる周波数群が存在する。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。「歌手のフォルマント」は主に第3、第4、第5フォルマントが融合することで成り、そのエネルギーが3kHz付近に集約されることで生成されていると考えられている。3kHz付近は、人間が最も可聴しやすい周波数域であるため、人間の耳にはより聞こえやすい声になる。齋藤(2008)の研究における歌声特有の音響特徴量の分析では、3kHzの周波数帯における強い高周波成分が「響き」の聴覚印象に寄与することが確認されている。また、優れた歌手の声がオーケストラと共に歌って

も聞こえる理由もこの「歌手のフォルマント」の存在があるからと考えられている。

図2-5は歌手の歌声，その歌手の話し声，ならびにオーケストラの音の長時間スペクトルを比較したものである。

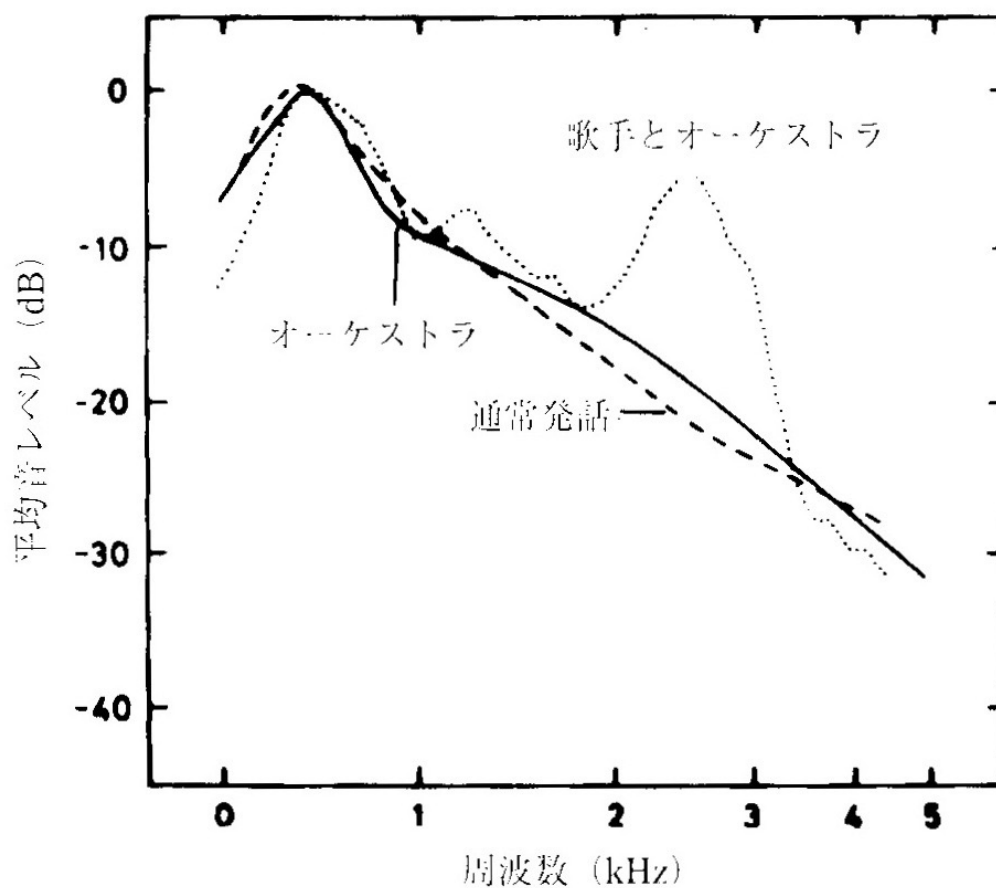


図2-5 交響楽でソリストがいるとき(実践)といないとき(点線)，及び通常話し声(破線)での長時間スペクトル。「歌手のフォルマント」が，オーケストラにソリストがいるときといない時の大きな違いを示している。(Sundberg, 2007)

オーケストラの音には3 kHz付近に共振がそれほど顕著に現れないため、その付近の周波数帯域にパワーの盛り上がりが存在することはオーケストラの音以外の音が存在すること、つまりこの場合はベル・カント唱法による歌声が存在することの「証拠」として働く。その結果、歌声はオーケストラの音にマスクされにくいという傾向を生む。同じ人間の話し声ではこの3 kHz付近のピークが出現していないことから、この特性はその個人の普段の状態から出現するものではなく、歌唱時に特異的に出現するものであると考えられる。すなわち楽器にはないフォルマントのピークを歌声に与えることにより、オーケストラと共に歌っても聴衆が歌声を聞き取りやすい条件を作っていることになる。

このようにクラシック歌唱の訓練をしっかりと受けた歌唱者の歌声の音響的な特性として「歌い手のフォルマント」の存在があることは分かっている一方で、その発生の生理学的な機序についてはまだ未解明の部分が多く残っている。



### 第3章 「歌手のフォルマント」の研究と課題

### 3.1 長時間スペクトル分析による研究の限界

これまで、歌声や「歌い手のフォルマント」に関する基礎的な理論について述べ、各項において行われている過去の研究や最近の研究を紹介してきた。この項では、さらに近年盛んに行われている歌声や「歌い手のフォルマント」における研究について報告し、そこから見えてくる課題について論じる。

近年のデジタル信号処理の普及と、それに伴うコストの低減によって、人間の発声器官や音声、また音を聞くための聴覚などの研究が進展し、より正確で精密な研究がなされるに至っている。それに加えて、音声の加工・合成技術についての研究も注目されており、歌声を中心としたメディアコンテンツの普及に大きな発展をもたらしている。

では、「歌い手のフォルマント」に関する研究はどうだろうか。先行研究において、「歌い手のフォルマント」を含む高域スペクトルが、個人性に寄与していると言うことは示されている。また、話声と歌声を判別する要因として、3 kHzにおけるフォルマントのピークである「歌い手のフォルマント」や、声の揺れであるヴィブラートが、歌声として知覚するための要因になっていると言うことは分かっている。齋藤、後藤(2009)らの研究では、歌唱指導によって、3kHz付近のフォルマントピークが高まり、ヴィブラートが変化するという結果を示した。少なくとも、この2要因は歌唱技術に関わっていると言える。しかし、「歌い手のフォルマント」と歌唱技術との具体的な関わりについての研究はまだ少ない。国内外の先行研究において、「歌い手のフォルマント」の存在が理想的な声質、とりわけクラシック音楽におけるベル・カント唱法の豊かな声質に必要な不可欠な音響特性であるという研究はなされている。しかし、先行研究によって見出された「歌い手のフォルマント」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だけであり、「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のが現状である。

例えば、楽曲の中には歌詞や旋律が存在し、それに伴い歌手は、母音や音高を変化させてゆく。しかしその瞬間において、「歌手のフォルマント」がどのように存在し、関わっているのかとすることはわかっていない。つまり、全体を通せば「歌手のフォルマント」という存在は確認できるが、具体的な歌唱技術に「歌手のフォルマント」がどのように関係しているのかは、まだ十分に解明されていないのである。さらなる研究によって、この「歌手のフォルマント」の生起や特性の時間的なふるまいが解明されれば、「歌手のフォルマント」が歌唱技術にどのように関わっているのかという疑問に答を見出すことができ、歌唱における演奏技能の解明にもつながるのではないだろうか。

### 3.2 「音質」の変化の少なさ

声楽における歌唱技術指導の上で、頻繁に注意喚起され、最も重要視されているものの一つに「響きの統一」というものがある。これは、歌唱において発せられる歌声の音色や響きを出来るだけ一定に保つことを意味しており、それを実現することによって、旋律をレガートに歌う事が出来、より遠くまで声を届けて喉の疲弊を軽減する事ができるというものである。この歌唱技術に関しては、「響きの統一」という言葉だけではなく、「母音の統一」「響きの確立」「マスケラ(伊:maschera=仮面)で歌う」等、様々な表現で示されているが、本論文においてその多様性について言及することはない。だが少なくとも、歌声の音が変わる（ピッチや母音が変わる）場合において「音質」を揃えるという事が、歌唱技術の中で重要であるということは明白である。しかし、このように重要視されているにもかかわらず、その音質変化の程度に着目した研究例はまだ少ない。先に述べたように、先行研究では長時間スペクトル分析によつてしか「歌い手のフォルマント」を観察できていないために、この「音質の変化」については、十分な研究がなされていないのが現状である。つまり、楽曲全体の平均スペクトルの中に含まれる「歌い手のフォルマント」の存在は確認されているが、ある一定の条件のもとで「歌い手のフォルマント」が局所的に発生しているのか、全体を通して持続的に存在しているのかは分からないのである。近年では、大谷（2014）により、ピッチの変化に対する歌声のスペクトルの安定性と聴覚印象の関わりについて調べた、これまでにない視点を導入した研究がなされている。研究結果から、1オクターブの順次進行の歌唱において、歌声の全体のスペクトル変動は、少ない方が聴覚印象において高い評価を得られるという事が明らかにされている。しかし、この研究では主に声区転換における音質の違いに注目した研究であったため、その歌唱課題は全てa母音でのみ行われた。そのため、異なる母音間での歌声のスペクトルの変化については研究されていない。母音が変わるということは、声道の形状が変わることを意味し、

またそのことにより、歌声のスペクトルの構造が変化する。従って「歌い手のフォルマント」も母音の変化に大きく影響を受けてしまう。また、この研究では歌声のスペクトル全体の変動について観察したことから、「歌い手のフォルマント」自体がどのように変動しているのかということには言及されておらず、歌唱におけるピッチや音韻の変化に伴う「歌い手のフォルマント」の時間的なふるまいについては今だ謎のままである。

これらのことから、「歌い手のフォルマント」に関する研究において、さらに求められことは、歌唱における旋律の変化と音韻の変化に着目することであり、それらの変化の中において「音質」がどのように保たれているのか、そして中でも「歌い手のフォルマント」がどのような時間的ふるまいであるのかということを観察する必要があると思われる。

### 3.3 観察される音声からソース情報とフィルター情報への分離

歌声のスペクトルは、第2章のソース・フィルター理論でも述べた通り、ソース情報（喉頭音源スペクトル）とフィルター情報（声道伝達特性）に分けて考えることができる。（図3-1）これは理論上の話ではなく、実際に声帯から発せられている喉頭原音が、喉頭や口腔などの声道に響くことにより、歌声が作られているのである。

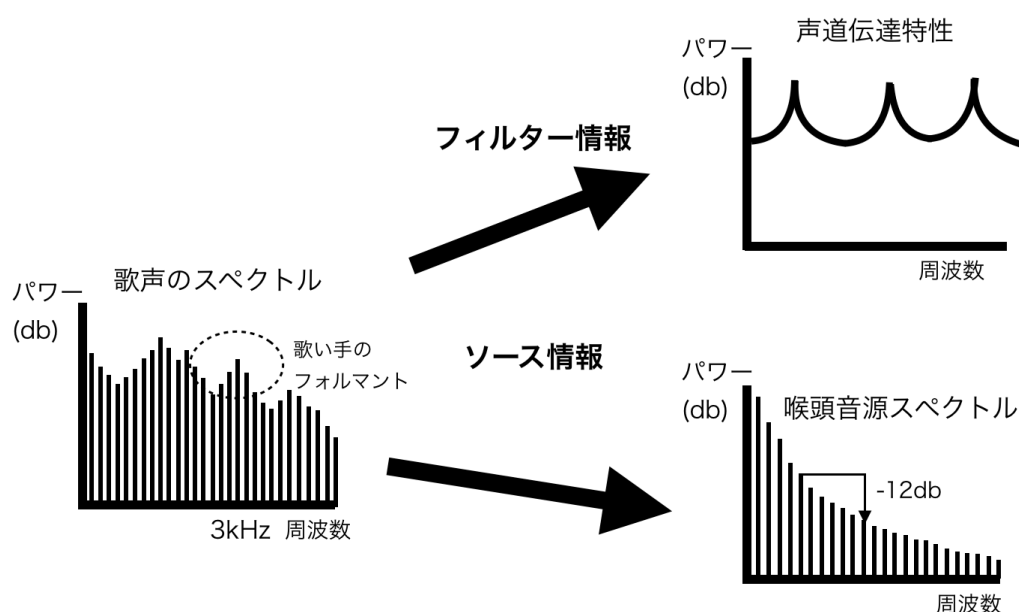


図3-1 ソース情報とフィルター情報への分離

歌声のスペクトルは、ソース情報（喉頭音源スペクトル）とフィルター情報（声道伝達特性）に分けることができる。

音声の研究では、サンプリングされた刺激をフーリエ分析して、そのパワースペクトルを見る方法がしばしば行われてきた。しかし、フーリエ分析で得られたパワースペクトルでは、基本周波数（F0）の情報を含んだままであり、音高が変化することによってF0が変化すれば、歌声のスペクトル全体が、F0の影響を受けて変わってしまう。（図3-2）その状態で二つのスペクトルの違いを観察することは難しく、歌声において高次

フォルマントの変化について観察する場合には不適切である。つまり、フーリエ分析によって観測できるスペクトルの調波構造は、F0の変化によって、ずれてしまったものであり、そのずれによって、同じ音色であるにもかかわらず数値上は異なるものになってしまうのである。それらの問題を解決するために、近年、新たな音声分析の方法として河原(1999)が開発した、STRAIGHT分析法が注目されている。このSTRAIGHTは音楽知覚研究を支援するために発明されたもので、STRAIGHTでは、パワースペクトルに含まれるF0の周期倍音の影響を、F0に適応した処理を用いることで分離して取り除き、F0による干渉のないスペクトル包絡として取り出すことを可能にした。よって、STRAIGHTで観測できるパワースペクトルは、平滑化スペクトルと呼ばれ、F0の変化による影響を受けていないスペクトル包絡を観察することができる。

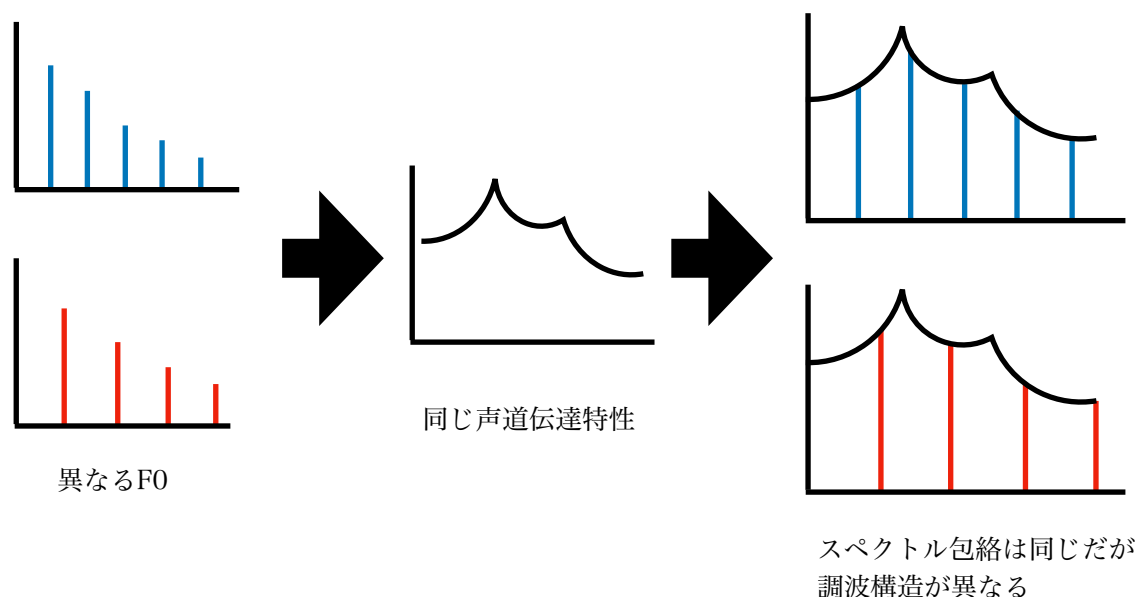


図3-2 F0の影響によるフーリエスペクトルのズレ

F0のが変化したことで同じスペクトル包絡（音色）を持ちながらも調波構造が異なる

## 第4章 歌声の音響計測実験



## 4.1 目的

第4章では、実施した音響計測実験についての詳細を記述する。音響計測実験の目的は、歌唱における「歌い手のフォルマント」に関する特徴量を抽出し、歌唱の安定性を示す特徴量の算出方法として妥当性の高いものを見出すことである。

第3章の第1項において述べた通り、先行研究によって見出された「歌い手のフォルマント」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、「歌い手のフォルマント」がある一定の条件のもとで局所的に発生しているのか、歌唱全体を通して持続的に存在しているのかは分かっていない。楽曲中の歌詞や旋律に伴い、母音や音高が変化する中で、「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。

歌唱には「旋律・歌詞」という2要素があり(図4-1)、これをほぼ独立に変化させることが歌い手には求められる。旋律における「音高の変化」と歌詞に含まれる「母音の変化」を同時に達成するためには、声道の形状や声帯の形状を巧みに制御しなければならない。声道の形状の変化は、日常的な会話の音声の中でも起きている事象であるが、さらに歌唱においては、日常的な会話では起きない音高の変化、つまり基本周波数(F0)の動きや、歌唱における明確な母音の変化のために、声道や声帯の形状の急峻な変化が求められる。それらのことが連続した2音間に大きな声質の差を生みやすいと要因として働き、その音質の差の激しさは多くの場合、歌唱能力の熟達度の低さと見なされる。即ち、歌唱においては、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、「歌い手のフォルマント」に代表される安定した声の質を保つことが要求されるのである。前項で述べたように、歌声における音高の変化には声帯の長さや張力が関わっており、母音の変化には声道の形状が関わっている。音響的な観点から言えば、音高の変化とは歌声のスペクトルの中に含まれる喉頭音源の基本周波数(F0)が変化することであり、母音の変化は、歌声のスペクト

ル群であるフォルマントの、第1フォルマントと第2フォルマントの関係性によって決定される。ソースフィルター理論においては、喉頭音源（ソース）と声道伝達特性（フィルタ）に分けて考えられ、理論上ではソースとフィルターは独立し、一方の特性が変化しないという条件のもとで、もう一方の特性を変化させるということが可能である。しかし、実際の発声器官においてはそれぞれの器官が独立して変化するというのではなく、喉頭の位置が上下したり、声帯や声道の形状を変化させることにより、筋組織がお互いに影響を受けてしまうのである。これらのことから音高や母音の変化を伴う歌唱において一定の声質を保つということは困難であると言える。

しかし、その状況において「歌手のフォルマント」に代表される安定した声質を保つことが出来るということは、歌手の熟達に関する特徴量（データにある特徴を数値化したもの）であると考えられる。

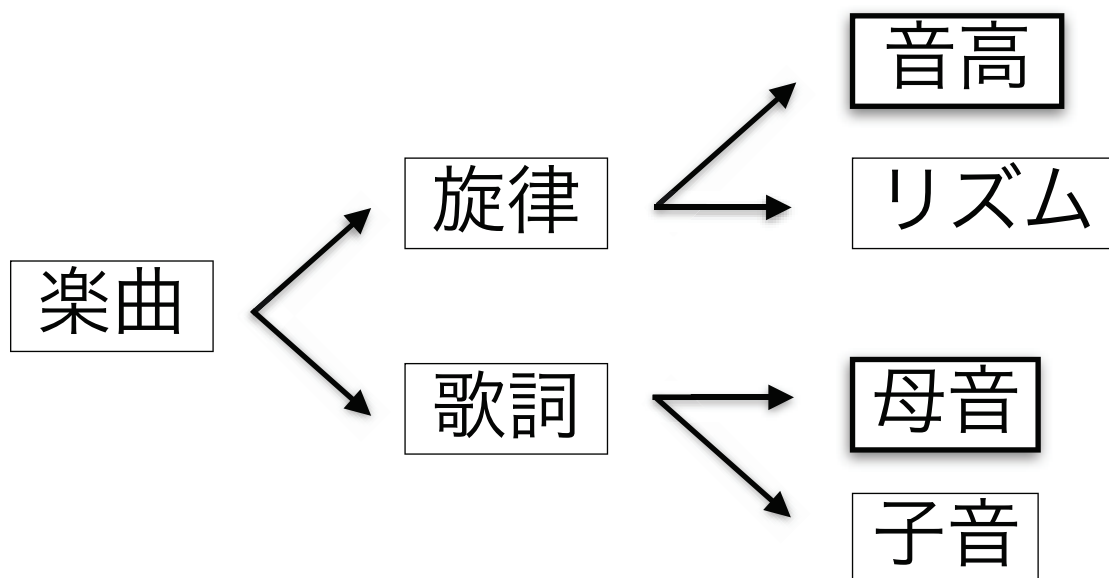


図4-1 楽曲の要素

楽曲の要素を旋律と歌詞に分けて、旋律を音高とリズム、歌詞を母音と子音に分ける。

## 4.2 実験方法

次の3点を軸として研究を進める。(図4-2)

1. 「母音・音高」の要素を直交配置した歌唱課題の作成.
2. 熟達度の異なる歌唱者による課題曲の歌唱収録と音声分析.
3. 音響特徴量の抽出.

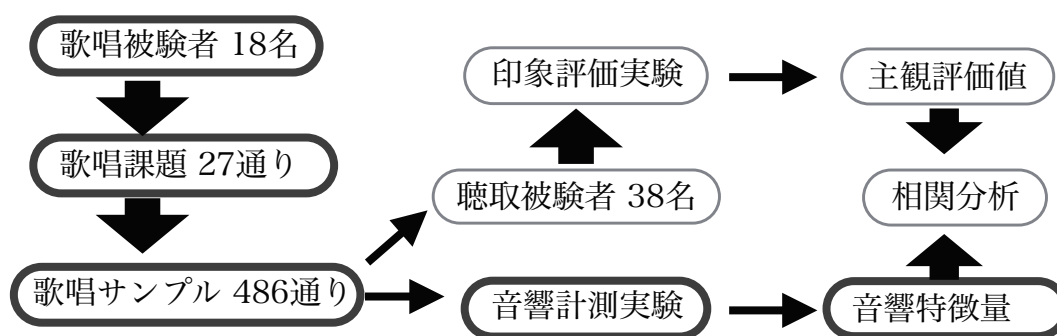


図4-2 実験の全体の中で第4章の流れを示すフロチャート

太枠の項目が第4章の流れを示す。「音高」の変化と「母音」変化を直交配置した歌唱課題を作成し、音響分析を行う。そして、その中から、「歌い手のフォルマント」に関する音響特徴量を抽出する。

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。つまり、喉頭音源スペクトルと声道伝達特性が掛け合わされることにより歌声のスペクトルとなるのである。

優れた歌い手ほど、この声帯と声道が変化する状況で、「歌い手のフォルマント」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「歌い手のフォルマント」付近のスペクトルを観察することが、本研究の手法上の特徴である。

## 歌唱課題

本実験は、歌唱における「音高」の変化と「母音」の変化に着目し、ある程度制御された歌唱課題を用いて収録された音声进行分析する事をねらいとしている。そのため、「音高」の変化においては、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。ここでは、隣り合う順次進行ではなく、ある程度の「音高の変化」が顕著に示す事ができるように跳躍進行を選んだ。また、跳躍進行は順次進行よりも歌唱技術においては難易度が高く、歌唱者の熟達度による差が出やすいと考えたからである。そして、本実験は「母音の変化」と「音高の変化」の歌声进行分析することから、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。(図4-3)

(e, oに関しては中間母音と呼ばれ、主要3母音の間に存在する中間的な母音であることから、今回の実験では省略した。「母音の変化」を顕著に示すためである。)歌唱被験者には、各自の歌いやすい主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。パッサッジョに関しては、歌唱技術の中でも「響きの統一」に関係するもので、異なる声区間において、それぞれの音色を出来るだけ変えないように行き来する事が良いとされている。これも「音高」の変化において、難易度を上げて、歌唱者の熟達度による顕著な差を示すためである。しかし、歌唱者の中には明確なパッサッジョの意識をもたない者もいた。それは、技術的な才能に恵まれ、意識せずにそれらを実現させている場合と、それとは反対にまだ歌唱技術として習得できていない場合の2通り存在

したが、その場合はそれぞれの声部で通説的にパッセージョであるとされている音を基準に選定することとした。

(1)同音程同士

(2)主音から完全五度上への上行跳躍進行

(3)主音から完全四度下への下行跳躍進行

**音型**

例 主音がラの音の場合

1音目                  2音目

**母音の組み合わせ**

1音目	2音目	
a	a	} 9通り
	i	
	u	
i	i	
	a	
	u	
u	u	
	a	
	i	

主要三母音である'a', 'i', 'u'の順列9通り  
(aa, ai, au, ii, ia, iu, uu, ua, ui)を組み合わせて

合計27通り (3×9) の歌唱課題を設定した。

**図4-3 歌唱課題**

「音高の変化」と「母音の変化」を直交配置した歌唱課題。

## 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、その歌唱者の熟達度と歌唱データとを比較分析することが本研究の目的である。そのため、声楽の中度経験者や高度経験者を歌唱被験者とした。

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7名、テノール4名）11人、高度経験者には、京都市立芸術大学大学院声楽教員及び同大学卒業生を含む、より経験年数の高いプロのバリトン歌手4名とテノール歌手3名の合計18人に、先に述べた歌唱課題を課し録音した。プロ歌手の選定基準としては、演奏会やオペラ公演の出演などプロとしての活動履歴がある歌手を選出し、経験年数が10年を超えており、コンクール入賞歴なども選定材料とした。

先行研究から、女性の歌声は基本周波数が高いために「歌手のフォルマント」の観測が難しいことがわかっている。そのため本実験では男性の歌声のみを扱うこととした。

## 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。ヘッドセット・マイクフォンの利用により、マイクと歌唱者の口元の距離や方向性の変化は最小限に抑えられると考えた。その程度には個人差があるものの、多くの歌唱者は歌唱時にある程度の姿勢の変化や顔の向きの変化を伴う。スタンド・マイクによる収録では、この姿勢変化によって口元とマイクの位置関係が変化し、それは収録された音声に少なからず

影響を与える。実験の目的上、できる限り歌唱者の口元での音声資料を収録することが望ましいと考えた。スタンドマイクを使用して歌唱者にできるだけ姿勢制御をしないように教示して収録する手法も検討したものの、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強いることとで不必要な緊張を起す可能性が考えられたためこの手法とした。

### 4.3 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(河原, 1999)によって平滑化スペクトルを求めた。従来の音声研究で用いられているフーリエ分析では、基本周波数 ( $F_0$ ) を含んだままのパワースペクトルであるため、先の述べた通り、今回のように音高の変化や母音の変化を含む歌唱課題を用いた実験には適さない。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。そのため、歌声のスペクトルを、ソース情報とフィルター情報に分離することができ、 $F_0$ の影響を受けないスペクトル包絡を観察できるのである。今回の実験のように、「音高」の変化についてスペクトル分析を行う場合、 $F_0$ による調波構造のズレは、仮に異なるピッチにおいて同じスペクトル包絡を持っていても、数値上では異なるものであるということになり、分析結果に大きな影響を及ぼす。これらのことから、今回の実験ではSTRAIGHT分析法を選択した。

### 分析対象の切り出し

STRIGHT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秒間の定常部の一つ目を第1区間、二つ目を第2区間とし、それぞれの区間のスペクトルを分析対象とした。(図4-4)

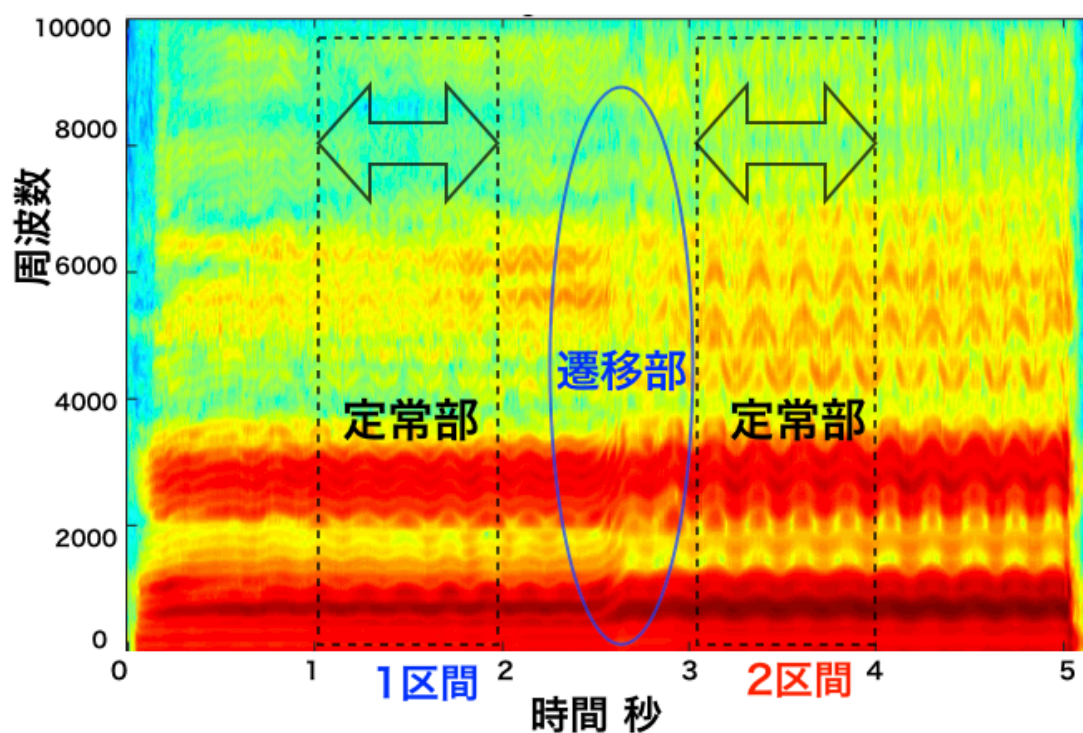


図4-4 分析対象の切り出し

それぞれの区間から、定常部から一秒間のスペクトルを切り出した。



#### 4.4 占有率

先行研究から、「歌い手のフォルマント」が優れた歌手の歌声に共通する音響的特徴であることがわかっている。そこで本実験では、「歌い手のフォルマント」の大きさに注目して、その特徴量を見出そうと考えた。本実験では、中度経験者以上の歌唱者を選定しているため、「歌い手のフォルマント」のピークがないというサンプルは存在しないが、そのピークの大きさには様々なものが存在した。それらのサンプルにおいて「歌い手のフォルマント」の大きさを、一つ一つ目視によって確認することも不可能ではないが、今回の実験に用いたサンプル数から考えるとその方法は適さないと考えられる。

そこで今回の実験では、「歌い手のフォルマント」自体が、歌声のスペクトル全体の中でどのくらいの割合を占めているのかという「歌い手のフォルマント」の「占有率」を見るという方法を選択し、その大きさを量ることにした。占有率という計算方法ではピーク推定が必要なフォルマント推定に比べて、頑健性を持つ結果が得られると期待できるからである。

また、大谷・津崎らの研究から(2014)から、歌声のスペクトルの変動の大きさは、歌声の安定性を考える上で大きな要素となりうるということも示唆されていることから、さらに「歌い手のフォルマント」の「変化」に着目し、「歌い手のフォルマント」の占有率の変化についても検討しようと考えた。

## 二つの方法による「歌手のフォルマント」の占有率の観察

「歌手のフォルマント」の占有率を観察するにあたり、第1区間と第2区間において計算されたそれぞれの占有率をどのように観察するべきか考えた。第2章で述べた通り、3 kHz付近という周波数帯域が、人間の可聴音域のなかで最も可聴しやすい帯域であること、また他の楽器にはない周波数帯域のピークを持つことにより、他の楽器にマスクされにくいということが、「歌手のフォルマント」の利点であった。そう考えると、単純にその帯域のパワーが大きいことが良いと思われるが、ただ大きいだけではなく、歌声全体のエネルギーの中で、「歌手のフォルマント」の占有率が大きい方が効率が良く、そのことにより声の酷使を防ぎながら歌えるということの方が重要であると考えられる。そのため、それぞれの区間における占有率を足すことにより、「音高・母音の変化」に対する「歌手のフォルマント」の占有率の大きさを見ることにする。

そして、「歌手のフォルマント」の安定性という観点で考えるならば、それぞれの区間において「歌手のフォルマント」の占有率がどのように変化しているかの変化量を見る必要があると考えた。これは「音高・母音の変化」を直交配置した歌唱課題の中で、それぞれの区間において「歌手のフォルマント」の占有率が変わらないほど、一定の音色や響きを保つことができているということの意味するからである。また、声楽家である筆者の経験や声楽における通説、あるいはベルカント唱法の理念によると、優れた歌手ほど異なる音域や母音において、歌声の音色や響きを一定に保つことが出来るとされているからである。そのため、優れた歌手ほど、歌声のスペクトル全体の中での「歌手のフォルマント」の割合が、均一化されているのではないかという仮説を立てた。

これらのことから、それぞれの区間の「SF占有率」と「SF占有率変化量」の二つの特徴量を見ることとする。(SF=Singer's formant)

### 比較周波数帯の算出と計算方法

「歌い手のフォルマント」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。「歌い手のフォルマント」は、主に第3, 第4, 第5フォルマントが3 kHz付近に集約されることで生成されている(Sundberg, 2007). そのため、この3 kHzを基点として前後1000Hzにあたる2 kHzから4 kHzまでの周波数域について観測することにした. 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「歌い手のフォルマント」の変動が、おおよそ2 kHzから4 kHzまでの間で起こっていることが窺われたことからこの区間を選定した. また2kHz以上は、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2kHzまでにあることがわかっていることも理由としてあげられる.

占有率 $PRO_{sf}$ の計算は次式(1)による.

$$PRO_{sf} = 10 \log_{10} \left( \frac{\int_2^4 p_f df}{\int_0^{22.05} p_f df} \right) \quad (1)$$

ここで $p_f$ は周波数  $f$  (単位: kHz)に対するパワーである. この式で対数を取る値の分母は全体のパワーであり、分子は2-4 kHzの部分のパワーである. 両者の比、即ち全パワーに対する「歌い手のフォルマント」近辺のパワーの割合を占有率とする. 今回は、その常用対数を10倍し、dB単位で表記する. 全体のパワーは常に部分のパワーを上回るので占有率のdB値は常に負値となる. 理論上は2-4 kHzのみにしかパワーがないときに0 dBとなり、これが最大値である. 値が大きいほど、つまり負値の絶対値が小さいほど占有率は高いこととなる.

## 4.5 RMS変化量

### RMSによる比較周波数帯の算出

「歌い手のフォルマント」を、先に述べた「占有率」とは異なる算出方法を用いて比較する方法として、3kHz付近の帯域における2音間のRMS（二乗平均平方根）が考えられる(図4-5)。それぞれの区間における「歌い手のフォルマント」の差分をRMSによって算出し、2音間の「歌い手のフォルマント」の変化量をグラフから観察する事ができるのである。今回の実験は、主に「歌い手のフォルマント」の「占有率」についての観察であるが、副次的にこのRMS算出方法を用いることにより、占有率とは異なる視点で「歌い手のフォルマント」を観察する。以下に、RMS指標における4つの算出方法を示す。

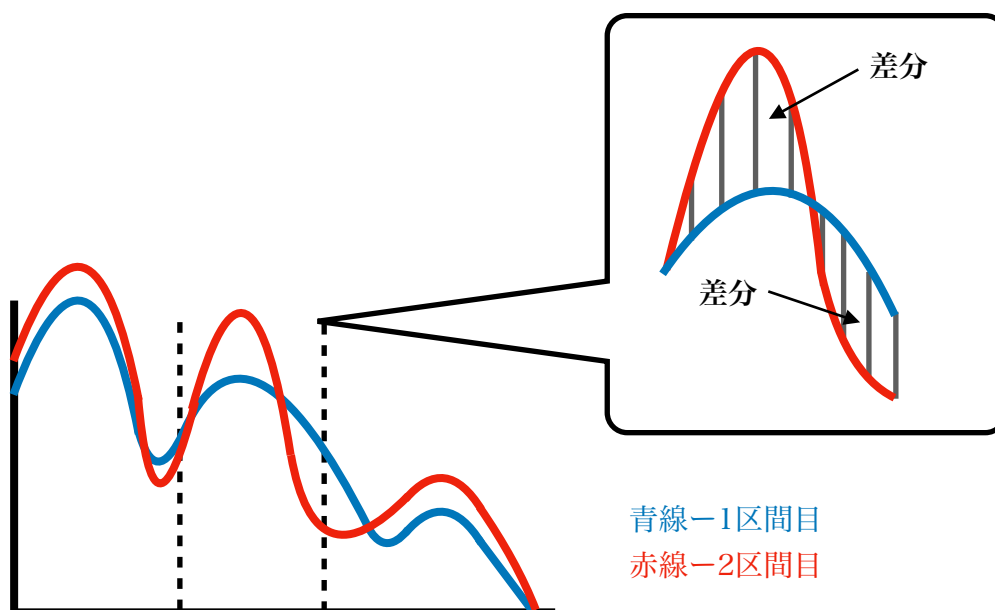


図4-5 RMS変化量の計算方法

2区間目の値から1区間目の値を引いたそれぞれ差分に対して二乗平均平方根を行う。

## 2-4 kHzにおけるRMS指標

まず単純に、3 kHzを中心とした2 kHzから4 kHzまでの周波数域について、各周波数ビンでのRMSを求めた。これは占有率の比較周波数帯と同様であり、実際のデータを目視した結果、1区間と2区間における「歌手のフォルマント」の変動が、およそ2 kHzから4 kHzまでの間で起こっていることがわかれたからである。(図4-6)

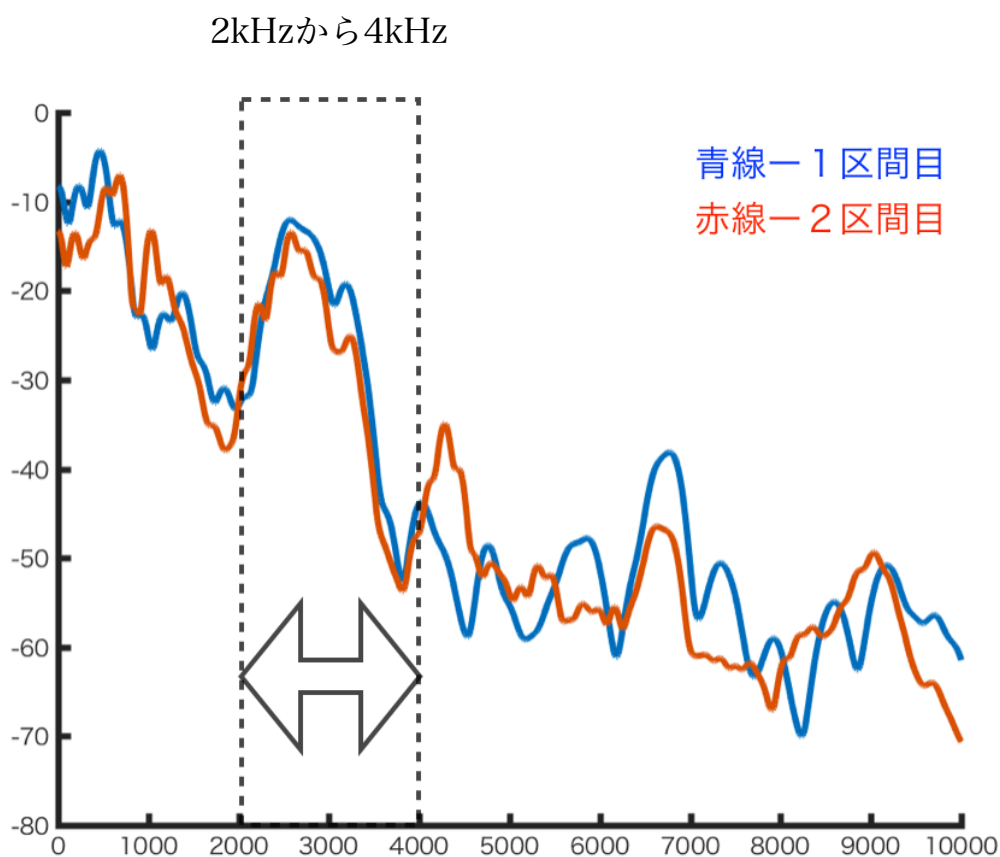


図4-6 RMS比較周波数の算出1

2kHzから4 kHzまでの帯域について観測した。

### 極大点前後500HzのRMS指標

そしてもう一つの2区間の変動の指標は、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。先述したように、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2kHzまでにあることがわかっている。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「歌手のフォルマント」が選ばれることになる。事前の検討により、「歌手のフォルマント」の最大値を中心にする、その前後500 Hzの範囲に共振の山がほぼ入ることが確認された。従って、1000 Hzという範囲で変化量を見ることとなる。(図4-7)

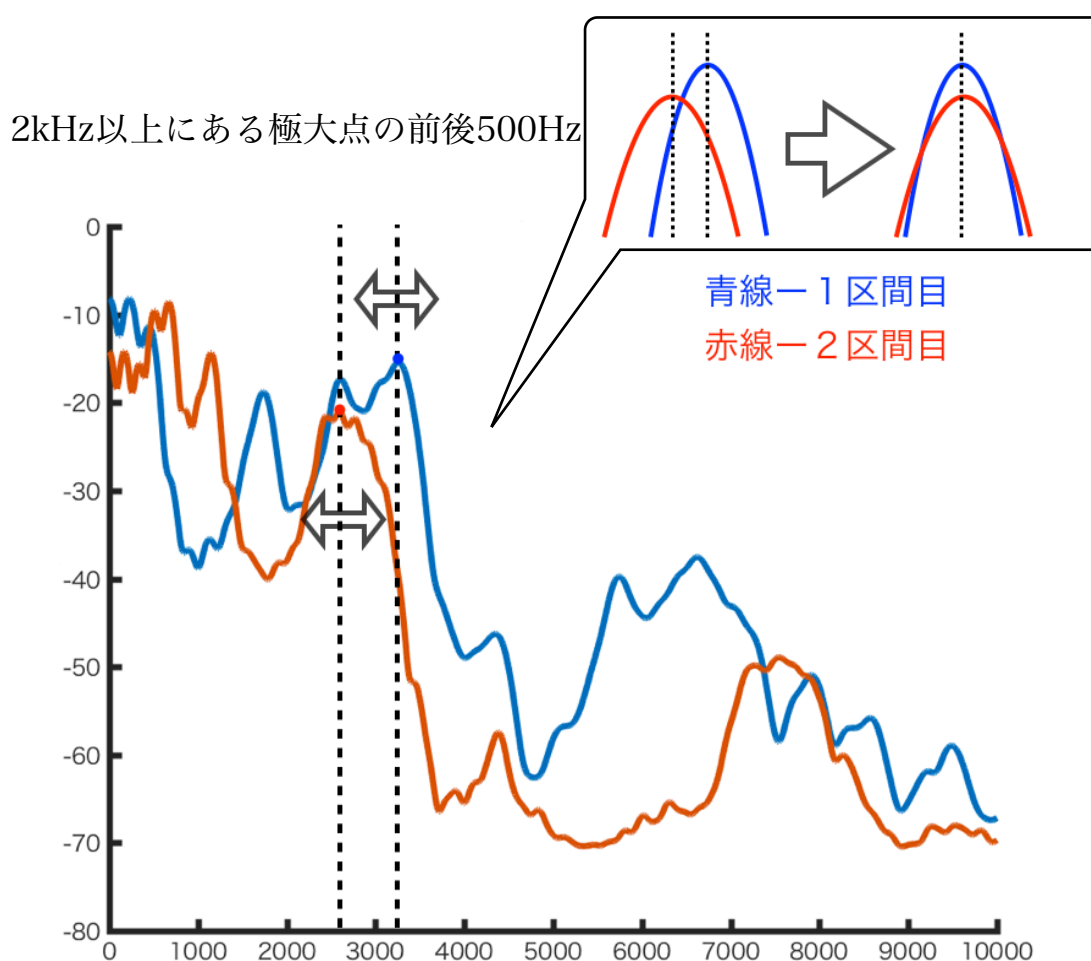


図4-7 RMS比較周波数の算出2

それぞれの区間における極大点の前後500Hzについて観測した。

## パワー正規化

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「歌い手のフォルマント」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「歌い手のフォルマント」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、本実験ではその両方を観察することにした。(図4-8)

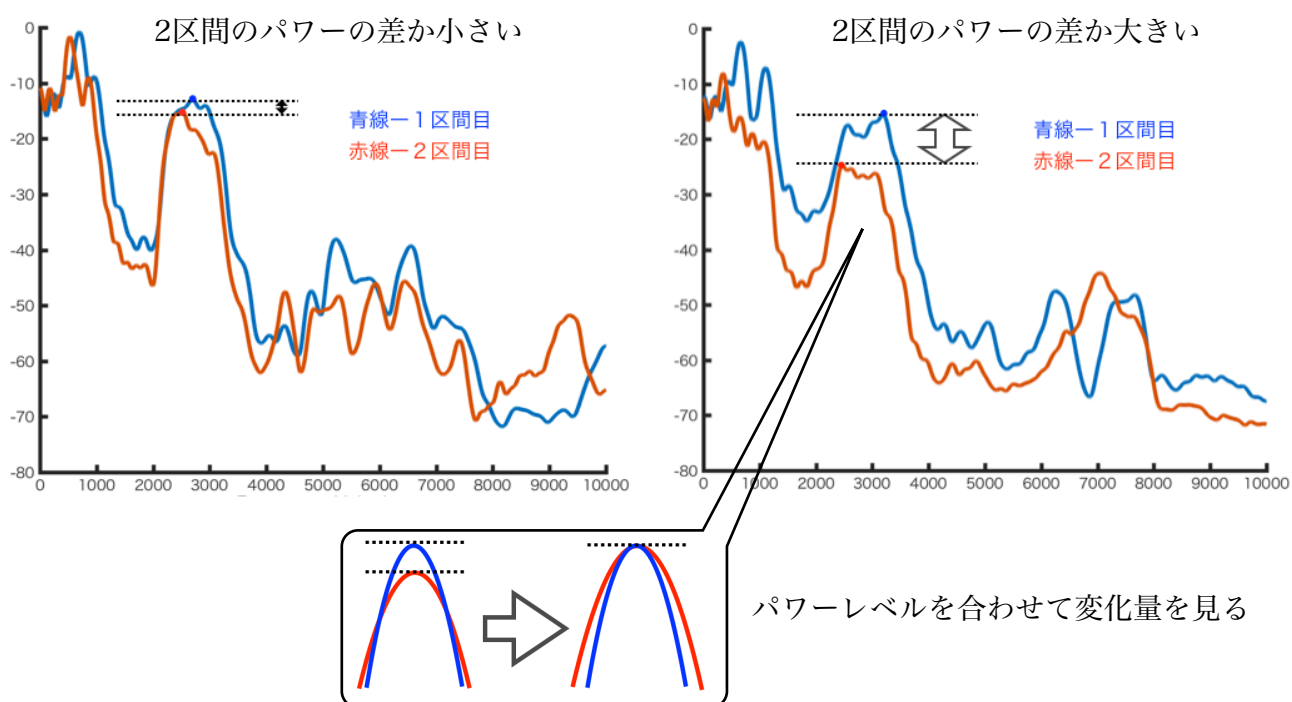


図4-8 パワー正規化

第一区間目と第2区間目のパワーの差を正規化によって合わせる。

よって、第1区間と第2区間の「歌い手のフォルマント」のRMS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（以下「2-4kHz変化量正規化なし」「2-4kHz変化量正規化あり」「極大点変化量 正規化なし」「極大点変化量 正規化あり」とする）が存在することとなる。以下に、それぞれの特徴量の値の歌唱者別平均値を表にする。(表1)

歌唱者	SF占有率	SF占有率 変化量	2-4kHz変化量 正規化なし
A01	-18.55	5.08	13.82
A02	-30.03	10.19	19.03
A03	-35.51	7.63	16.20
A04	-27.09	7.10	16.18
A05	-30.03	12.20	18.74
A06	-37.69	7.98	13.26
A07	-18.13	6.86	14.21
A08	-36.59	8.67	17.51
A09	-36.13	9.50	17.45
A10	-23.52	8.76	15.19
A11	-18.36	7.62	20.09
P01	-10.07	4.47	9.83
P02	-20.23	4.15	16.52
P03	-24.71	5.51	13.97
P04	-27.74	5.26	13.21
P05	-11.07	6.17	14.86
P06	-23.98	9.38	11.74
P07	-34.28	9.36	17.51

表1-A 各特徴量に対する歌唱者別平均値

2-4kHz変化量 正規化あり	極大点変化量 正規化なし	極大点変化量 正規化あり
11.20	10.44	0.36
17.35	16.25	0.41
14.22	12.81	0.39
13.85	13.83	0.36
18.88	16.92	0.29
13.65	12.77	0.27
10.22	12.24	0.33
16.70	17.72	0.29
17.14	14.74	0.36
13.39	16.74	0.32
16.86	13.32	0.36
9.50	8.94	0.30
13.91	11.66	0.42
11.39	10.72	0.38
11.27	12.01	0.28
13.72	11.90	0.36
11.46	14.33	0.33
16.09	17.86	0.37

表1-B 各特徴量に対する歌唱者別平均値



## 第5章 歌声に対する印象評価実験

## 5.1 目的

第5章では、実施した印象評価実験についての詳細を記述する。印象評価実験では、収録した歌唱音声に対して、複数の評価者による主観評価を得ることが目的である。

第4章では、歌声のスペクトル分析を行うために、「音高」の変化と「母音」の変化を直交配置した歌唱課題を作成した。今回の実験では、プロの歌手と声楽を学ぶ学生の歌声を用いており、それぞれの歌唱者の経験年数は異なり、声楽家である筆者の主観的な印象では、その歌唱技術のレベルも様々であった。もちろん、プロの歌手の方が、声楽を学ぶ学生よりも、その技術レベルは高く感じられたが、それはあくまでも筆者自身の主観的な印象でしかないことは言うまでもない。しかし、今回の実験のように、歌声の音響的特徴が、どのように聞き手の聴覚印象に作用し、また、それが歌唱技術にどのように関係しているのかとすることを調べるためには、実験に用いた歌声のサンプルが、客観的にどのような評価を得るものなのかとすることを知る必要がある。先行研究では、実験に用いた歌声に対して、そのほとんどが客観的な評価がないままに分析を行っている。つまり、その歌声がプロの歌手であるとか、あるいは声楽を学んでいない一般人であると言うような曖昧な基準によって分析されているのである。

そこで本章では、歌声のスペクトル分析に用いたサンプル一つ一つに対して、多数の評定者による印象評価実験を行い、全ての歌唱課題に対して、主観評価値を得ることとする。

## 5.2 実験方法

次の2点を軸として研究を進める。(図5-1)

1. 音楽経験者を評定者とした歌唱課題に対する印象評価実験の実施.
2. 印象評価実験により得られたデータから, 各歌唱課題に対する主観評価を取得.

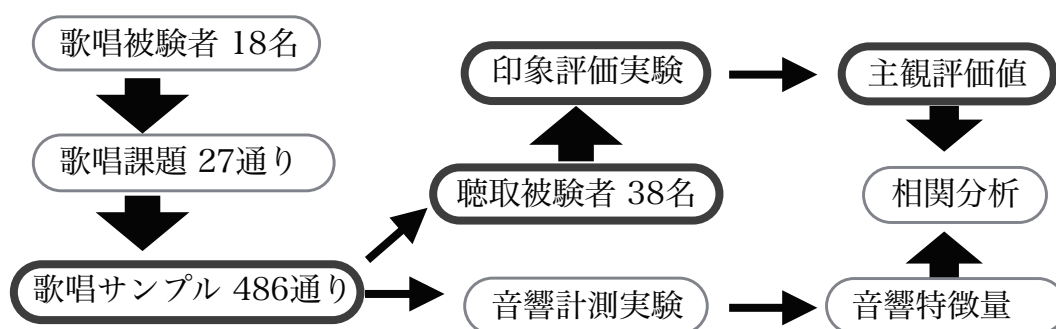


図5-1 実験の全体の中で第5章の流れを示すフロチャート

太枠の項目が第5章の流れを示す。歌唱者によって収録された歌唱課題に対して印象評価実験い、全ての歌唱課題に対する主観評価値を得る。

音楽経験者を評定者とした印象評価実験を行い、全ての歌唱録音データに対して主観評価を得るために、18人の歌唱者から集められた全課題曲の録音を486通り(歌唱者18名、音型3種類×母音の組み合わせ9種類)の音刺激(音のサンプル)とした。評価の基準は「良い発声かどうか、良い声であるか。」とし、聴取実験参加者はそれに対して1点から5点までの5段階評価を行う。それを多数の聴取実験参加者に実施することにより、全ての課題に対する主観評価を得る事が目的である。

聴取実験参加者には、18人の歌唱課題を録音した歌唱者の名前は明かさず、あくまでも聞こえてきた歌声に対して評価するように説明を行った。

また、録音した歌唱課題(486通り)は、歌唱者、音型、母音の組み合わせが全てランダムになるように、聴取実験参加者には、一人一人すべて異なる配置で提示した。印象

評価実験において、各サンプルに点数をつける場合、どうしても前後のサンプルに対して相対的な評価を無意識のうちに行ってしまう傾向にある。つまり、前に聞いた歌声と比べて、今聴いている歌声はどのくらいであろうかという心理が、本人の自覚なしに働いてしまうからである。そこで、そのような相対的判断が作用するということを考慮し、すべての歌唱課題をランダムに配置することによって、それらの問題を解決した。

聴取実験参加者は、PCのディスプレイ(Apple imac G5)に表示されたGUI上の点数を表したボタンをクリックすることによって回答を行なった。

486試行の中で、200試行目を過ぎると休憩を挟み、実験者本人が十分に休むことができたと感じてから再開した。

### **参加者**

聴取実験参加者は京都市立芸術大学音楽学部の学生と卒業生、そしてプロの演奏家の合計38人（声楽19名、ピアノ3名、作曲6名、音楽学4名、管楽器3名、弦楽器3名）である。参加者全員が過去の聴力検査において聴力の問題はないと報告している。

### **使用機材**

聴取実験は無音室（Yamaha AVITECS）の中で行い、PC（Apple imac G5）によって制御されるDigital signal processor(Pacarana FW800 Kyma,Symbolic Sound)により、その出力をヘッドホンアンプ(Luxman P-1)に繋がれたヘッドホン(Sennheiser HD600)を介して両耳に呈示した。

### 5.3 印象評価実験分析方法と結果

得られた回答から、全ての486個（18人の歌唱者×27通りの歌唱課題）の歌唱課題に対する38人分の主観評価の平均値（以下「主観評価値」）を求めた。

次にその結果(表2, 表3)と、それぞれの歌唱者に対する主観評価値の標準偏差を示す。

(図5-2, 図5-3)

歌唱者	U5th, aa	U5th, ai	U5th, au	U5th, ia	U5th, ii	U5th, iu	U5th, ua	U5th, ui	U5th, uu
A01	3.0	3.4	3.9	3.6	3.4	3.9	3.9	3.8	3.8
A02	2.7	2.7	2.6	2.1	1.9	2.7	2.7	2.4	2.4
A03	2.9	2.8	2.9	2.9	1.9	2.7	3.2	2.6	3.0
A04	2.5	2.5	2.4	3.0	2.4	2.5	2.6	2.5	2.4
A05	2.6	3.6	2.4	2.3	2.1	2.1	1.9	2.1	1.9
A06	2.8	2.9	2.7	2.5	2.8	2.8	2.6	2.7	2.6
A07	2.6	2.3	2.5	2.7	2.7	2.5	2.7	2.9	3.0
A08	2.8	3.2	2.6	3.0	2.7	2.6	2.8	2.5	2.3
A09	2.3	2.8	2.6	2.1	2.5	2.3	2.0	2.2	2.5
A10	2.2	2.8	2.8	1.8	2.7	2.9	2.2	2.7	3.1
A11	2.2	1.9	2.1	2.3	1.7	2.2	1.6	1.8	2.0
P01	4.3	4.1	4.1	3.7	3.6	4.3	4.2	3.9	4.2
P02	3.8	4.0	3.9	3.8	3.8	4.1	4.0	3.7	3.8
P03	3.6	3.6	3.5	3.9	3.7	4.1	3.6	3.9	3.6
P04	3.8	3.4	3.9	3.8	3.4	4.0	4.0	3.4	3.9
P05	3.5	3.8	3.8	3.6	3.9	3.8	3.7	3.8	3.5
P06	3.1	2.6	3.4	2.8	3.0	2.8	3.1	2.9	3.2
P07	2.6	3.1	3.3	2.7	3.4	2.8	2.7	3.1	2.8

表2-A 歌唱者別全歌唱課題の主観評価値

D4th, aa	D4th, ai	D4th, au	D4th, ia	D4th, ii	D4th, iu	D4th, ua	D4th, ui	D4th, uu
4.0	3.7	3.9	3.7	3.7	4.0	3.7	3.5	3.9
3.6	3.5	3.5	3.0	3.3	3.3	3.3	3.5	3.2
2.9	3.2	3.1	2.9	3.1	2.7	3.1	3.2	2.8
3.0	3.5	3.2	3.2	3.6	2.9	3.4	3.5	3.0
3.3	3.3	3.2	3.4	3.4	3.2	3.1	3.2	2.9
2.8	2.8	2.9	2.7	3.1	2.7	3.1	3.3	2.8
3.2	2.8	3.1	2.5	2.5	2.3	3.1	2.9	2.6
2.9	2.8	2.8	3.0	3.0	2.7	2.7	2.7	2.7
1.9	2.7	2.6	2.3	2.1	2.5	2.2	2.4	2.5
1.7	2.4	1.9	2.4	2.5	2.1	1.9	2.7	2.0
2.6	2.8	2.7	2.8	2.2	2.4	2.9	2.8	2.4
3.3	4.1	4.1	4.0	3.8	3.9	3.5	4.1	4.3
3.9	4.3	4.1	3.9	3.9	4.1	3.8	3.8	4.3
3.7	3.8	3.5	3.5	3.9	4.0	4.0	3.6	3.6
3.9	3.7	3.5	3.9	3.4	3.2	3.6	2.9	3.2
3.6	3.7	3.1	3.2	3.2	3.2	3.3	3.7	3.3
3.1	3.1	3.4	3.3	3.2	3.0	3.3	3.3	3.5
2.6	3.0	2.8	2.8	2.8	2.8	2.6	3.1	2.8

表2-B 歌唱者別全歌唱課題の主観評価値

Uni, aa	Uni, ai	Uni, au	Uni, ia	Uni, ii	Uni, iu	Uni, ua	Uni, ui	Uni, uu
3.7	3.8	3.9	3.8	3.3	3.6	3.4	3.8	3.7
3.5	3.7	3.5	3.6	3.5	3.7	3.7	3.8	3.6
3.2	3.1	3.4	2.9	2.8	2.9	2.9	2.7	3.0
2.7	3.2	2.9	2.8	2.9	3.0	3.0	3.3	2.6
2.4	3.2	2.8	2.8	3.1	3.0	3.0	3.4	2.7
2.7	2.6	2.8	2.8	2.7	2.8	2.8	2.9	2.8
2.3	2.8	2.8	2.8	2.8	3.1	3.2	3.1	3.0
1.9	2.8	2.4	2.8	2.8	2.7	2.7	2.8	2.5
2.8	2.9	2.8	2.8	2.9	2.6	2.7	2.6	2.6
2.7	2.9	2.3	2.6	3.3	2.5	2.6	2.7	2.7
2.5	2.8	2.5	2.6	2.2	2.5	2.3	2.2	2.1
3.7	4.1	4.3	3.9	3.6	3.8	3.9	4.4	3.9
3.8	4.0	4.1	3.9	3.6	4.1	3.9	4.1	3.6
3.4	3.9	3.7	3.8	3.7	3.4	3.7	4.0	3.6
3.4	3.9	3.7	3.6	3.7	3.7	3.6	3.7	3.4
3.2	3.4	3.2	3.6	3.1	2.9	3.3	3.5	3.4
3.1	3.1	2.9	3.2	3.5	3.0	2.8	3.1	2.3
2.6	3.4	3.1	3.4	2.9	3.1	2.9	3.1	3.0

表2-C 歌唱者別全歌唱課題の主観評価値

歌唱者	平均(主観評価値)
P01	4.0
P02	3.9
P03	3.7
A01	3.7
P04	3.6
P05	3.5
A02	3.2
P06	3.1
P07	2.9
A03	2.9
A04	2.9
A05	2.8
A06	2.8
A07	2.8
A08	2.7
A09	2.5
A10	2.5
A11	2.3

表3 歌唱者別主観評価値の平均

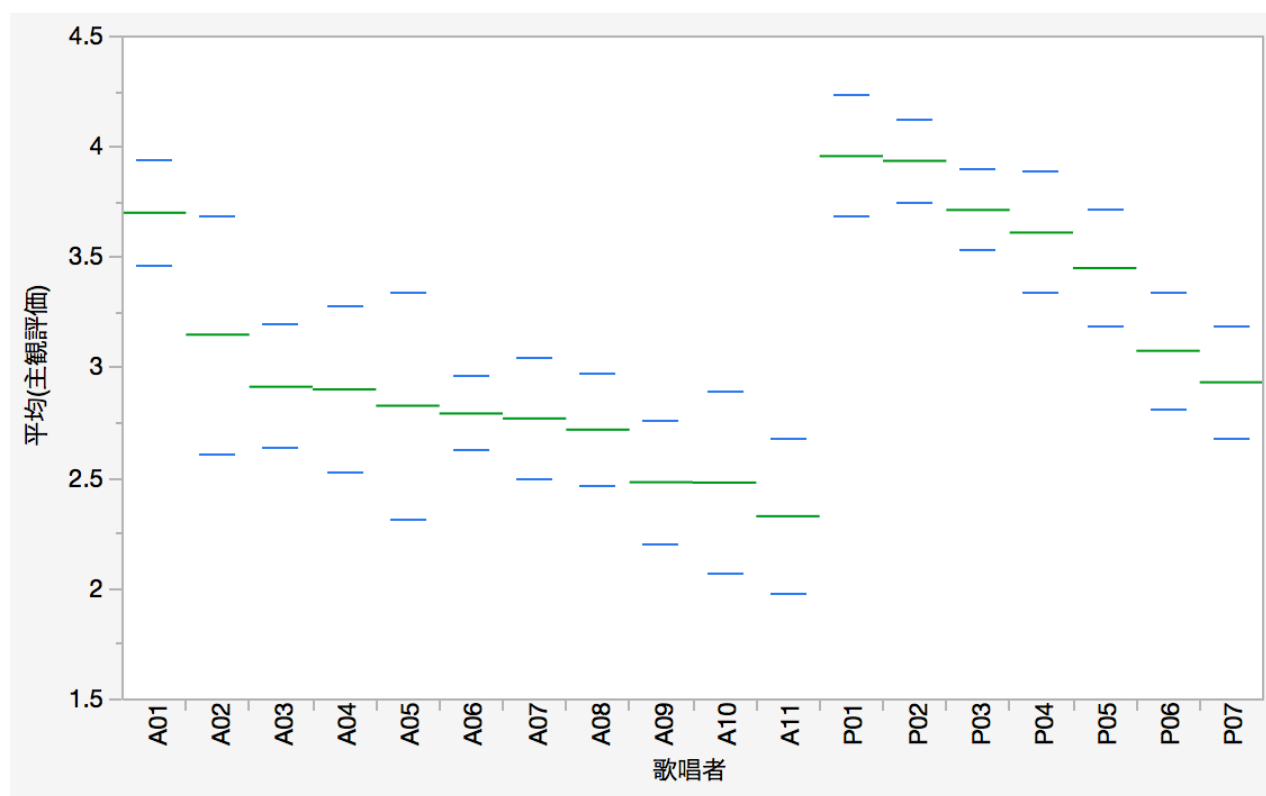


図5-2 歌唱者別主観評価値の平均と標準偏差

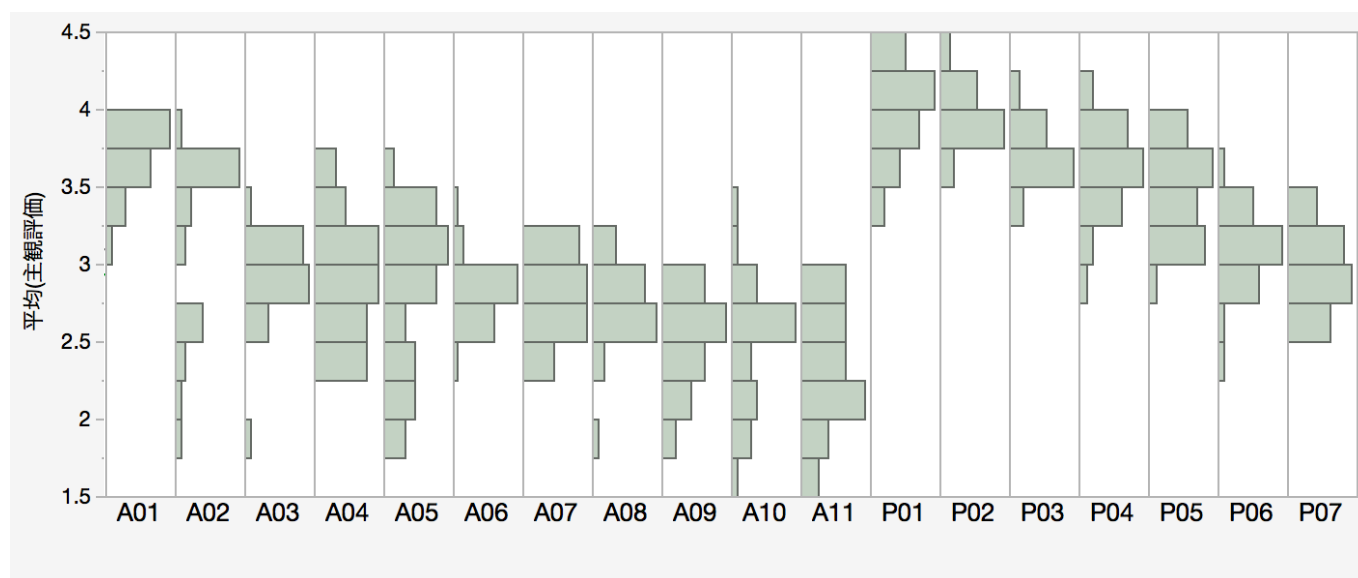


図5-3 歌唱者別主観評価値のヒストグラム



## 第6章 印象評価実験結果と音響特徴量の回帰分析

## 6.1 目的

第6章では、実施した回帰分析についての詳細を記述する。相関分析では、第4章で行った歌声の音響計測実験から抽出した音響特徴量と、第5章で行った印象評価実験から得られた全ての歌唱課題に対する主観評価値を用いて、歌声の音響特徴量とその主観評価の間にある関係を観察する事が目的である。(図6-1)

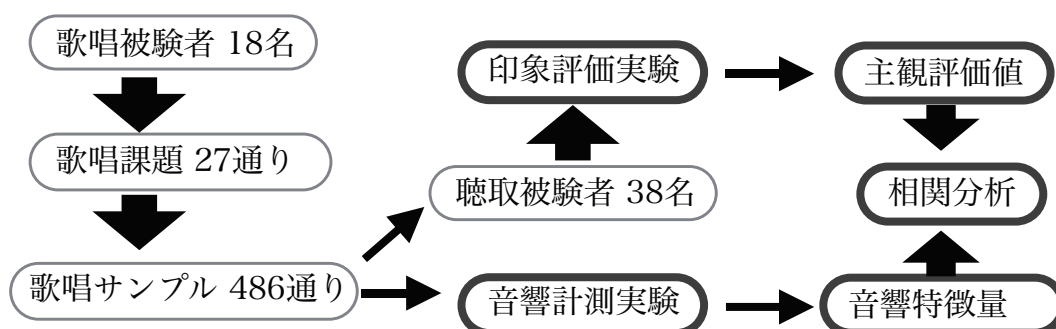


図6-1 実験の全体の中で第6章の流れを示すフロチャート

太枠の項目が第6章の流れを示す。音響計測実験によって抽出した音響特徴量と、印象評価実験から得られた主観評価値を用いて、相関分析を行う。

これまで、第4章において歌声の音響計測実験を行い、そこから「歌手のフォルマント」に関する6つの音響特徴量を抽出し、「歌手のフォルマント」に関する指標とした。また、第5章においては、音楽経験者を評定者とした印象評価実験を行い、歌唱課題全てに対する主観評価値を得た。よって各歌唱課題には6つの特徴量の値と主観評価値の全部で7つの値があり、それらは全て異なった変数である。そして、これらの変数を用いて相関分析を行い、「歌手のフォルマント」と主観評価の関係を調べる。

「相関分析」とは、2変数以上がある場合に、それぞれの間、相関を示す関係があるかどうかを調べるものであり、今回の実験の場合では、歌声の音声分析から抽出した音響特徴量と、印象評価実験により得られた主観評価値との間に、相関関係が見られるかどうかということを検定する。つまり、「歌手のフォルマント」に関する音響

特徴量と主観評価の関係性を見ることにより、「歌い手のフォルマント」が、どのように聞き手の聴覚印象に作用しているのかということ、客観的評価基準によって分析結果を得ることができるのである。先行研究において、「歌い手のフォルマント」の大きさや、高さに関しての研究はなされており、それが聞き手に良い印象を与えるということは分かっている。しかし、今回は「歌い手のフォルマント」を「占有率」という特徴量で算出し、さらに、その「変化」に着目した特徴量の有効性をみることになる。

## 6.2 回帰分析結果

主観評価値を目的変数として、音声分析による物理変数を説明変数とした回帰分析を行なった(図6-1)。主観評価値を目的変数とし、音響特徴量の各組み合わせ(「SF占有率」「SF占有率変化量」「2-4kHz正規化なし」「2-4kHz正規化あり」「極大点変化量正規化なし」「極大点変化量正規化あり」との重回帰分析を行った。今回のように6つの特徴量と主観評価値の合計7変数を用いる回帰分析は「重回帰分析」と呼ばれ、1つの目的変数(主観評価値)に対して、複数の説明変数(音響特徴量)で分析しようとするものである。それに対して、1つの目的変数(主観評価値)に対して、1つの説明変数(音響特徴量)で分析する場合は、「単回帰分析」と呼ばれる。しかし、今回のような複数の音響特徴量に対して、別々に単回帰分析を行ってしまうと、1つの目的変数(主観評価値)に対してそれぞれの音響特徴量が互いに影響しあっている可能性があり、実際にはない相関が表層的に現れ、関係性があると錯覚してしまうことがある。今回の実験では、「歌い手のフォルマント」に対する6つの音響特徴量が、どのように主観評価値と関係しているかということ調べる事が目的である為、まず全ての音響特徴量を用いて重回帰分析を行う。そして、それぞれの特徴量が重回帰分析の中でどのくらいの説明力を持つかを調べるために対数値を求め、その中から主観

評価値との関係性が認められる特徴量を選出する。そしてそれらの変数を用いて、再度重回帰分析を行うことにより、主観評価値と「歌手のフォルマント」に関する音響特徴量の間関係を調べる。

対数値は(2)式として計算される。

$$\text{対数値} = -\log_{10}(P\text{値}) \quad (2)$$

P値とは、帰無仮説(通常は差がないという仮説)が正しい時に、偶然によって観測されたデータ上に差が生じる確率を意味し、観測された差の統計学的信頼性を示す。一般的には、この値が5%未満 ( $p < 0.05$ ) の場合、そのデータは「有意差がある」とされ、5%以上の場合は「有意差がない」とされる。また、対数値の値が1.3以上は、0.05レベル(有意水準5%)で有意となる。(  $1.3 < -\log_{10}(p\text{値})$  となるため)

これらのことをふまえて以下にその結果を示す。







要因	対数値		P値
SF占有率	12.745		0.00000
極大点変化量 正規化あり	2.118		0.00761
SF占有率 変化量	1.080		0.08327
2-4kHz変化量 正規化なし	1.042		0.09076
2-4kHz変化量 正規化あり	0.664		0.21692
極大点変化量 正規化なし	0.034		0.92554

図6-2 6つの音響特徴量の対数値

図6-2において、それぞれの要因の対数値を見ていくと、まず「SF占有率」の対数値が[12.745]で最も高い値を示している。それに次いで「極大点変化量正規化あり」の値が[2.118]という値で、これら上位2つの特徴量が有意となっている。また、3番目にあたる「SF占有率変化量」と4番目にあたる「2-4kHz正規化なし」は、それぞれの値が[1.080]と[1.042]であり、有意傾向にあると言える。また、5番目と6番目の「2-4kHz正規化あり」と「極大点変化量正規化なし」は、それぞれ[0.664]と[0.034]という値で、有意にはならなかった。

これらの結果から、再度重回帰分析を行うために、「歌手のフォルマント」に関する音響特徴量を選出した。まず、6つの特徴量の中で対数値の値が低く有意とならなかった「2-4kHz正規化あり」と「極大点変化量正規化なし」については除外した。そして上位2つの特徴量(「SF占有率」と「極大点変化量正規化あり」)は、すでに有意となっていることから、SFに関する音響特徴量として説明することができるため説明変数として採用した。また、有意傾向にあった「SF占有率変化量」と「2-4kHz正規化なし」についても検討の余地があると考えた。先述したように2つの特徴量は、単純に2kHzから4kHzまでの周波数帯域を切り出して2音間で比較していることから、占有率で比較するのか、差分をRMSによって比較しているのかの違いだけであり、かなり似通った特徴量であると言える。そのため、2つの特徴量の間で対数値の高い「SF占有率変化量」を説明変数として残し、「2-4kHz正規化なし」を除外した。

これらのことから、3つの特徴量「SF占有率」「極大点変化量正規化あり」「SF占有率変化量」を用いて、再度重回帰分析を行なった。以下にその結果を示す。

### 分散分析

要因	自由度	平方和	平均平方	F値
モデル	3	24.82916	8.27639	28.1182
誤差	482	141.87296	0.29434	p値(Prob>F)
全体(修正済み)	485	166.70213		<.0001*

図6-3 重回帰分析

### 効果の検定

要因	パラメータ数	自由度	平方和	F値	p値(Prob>F)
SF占有率	1	1	17.974182	61.0656	<.0001*
SF占有率 変化量	1	1	2.168661	7.3678	0.0069*
極大点変化量 正規化あり	1	1	1.139464	3.8712	0.0497*

図6-4 効果の検定

「SF占有率」「極大点変化量正規化あり」に「SF占有率変化量」について、有意性が見られる。

### パラメータ推定値

項	推定値	標準誤差	t値	p値(Prob> t )	標準β
切片	3.5638216	0.08192	43.50	<.0001*	0
SF占有率	0.0191095	0.002445	7.81	<.0001*	0.338613
SF占有率 変化量	-0.010066	0.003708	-2.71	0.0069*	-0.11827
極大点変化量 正規化あり	0.3003566	0.152656	1.97	0.0497*	0.083357

図6-5 パラメータ推定値

「SF占有率」「極大点変化量正規化あり」については、推定値から正の相関が見られ、「SF占有率変化量」については負の相関が見られる。

重回帰分析を行なった結果、「SF占有率」と「SF占有率変化量」、そして「極大点変化量 正規化あり」のすべてが有意となった(図6-3,6-4,6-5)。

また、有意性のみられた「SF占有率」と「SF占有率変化量」、「極大点変化量 正規化あり」の3つの要素をそれぞれ( $x_1$ ,  $x_2$ ,  $x_3$ )とし、重回帰分析から得られた偏回帰係数を用いて(3)式とし、三次元散布図(図6-6)を作成した。また、主観評価値を色分け(高い:赤 低い:青)し、それぞれの歌唱課題における音声分析の特徴量と主観評価値の関係を見た。散布図からは、「SF占有率」が高く、「SF占有率変化量」が少ない右端に、主観評価値の高い赤い点が集中している事が見て取れる。

$$y=0.019x_1-0.01x_2+0.3x_3 \quad (3)$$

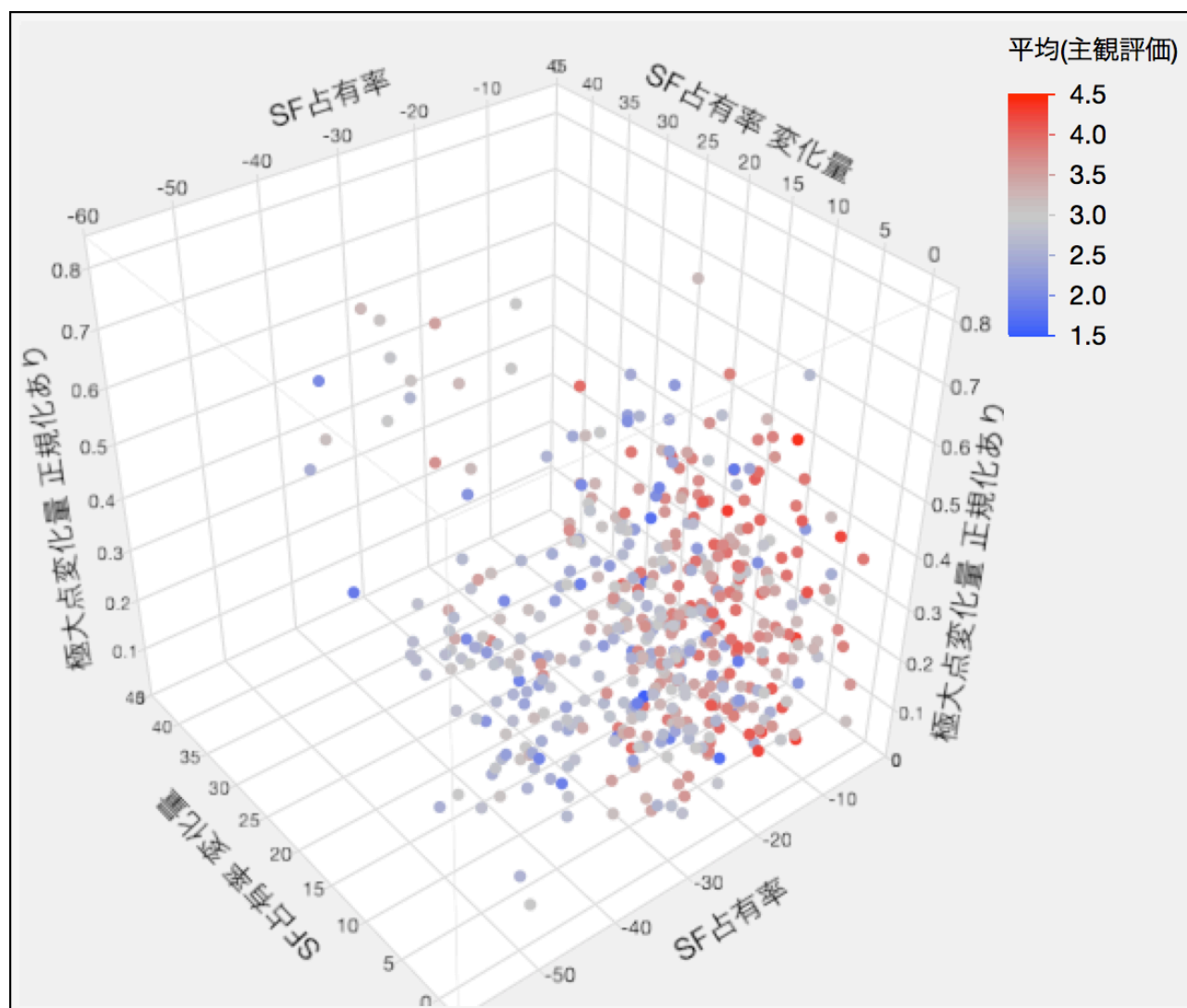


図6-6 三次元散布図

「SF占有率」「SF占有率変化量」「極大点変化量正規化あり」を三方向のベクトルにして、三次元散布図を作成した。主観評価値を色分けすることによって、それぞれの歌唱課題における音声分析の要素と主観評価値の関係が見て取れる。この散布図の角度からは、SF占有率が高く、SF占有率変化量が少ない右端に、主観評価値の高い赤い点が集中している事がわかる。



### 6.3 考察

まず「SF占有率」についての結果から考察する。「SF占有率」と「主観評価値」の要因は、図6-4,図6-5で示された通り有意となった。先行研究から「歌手のフォルマント」は「良い声」「上手さ」といった聴覚印象に寄与することが分かっている。今回の実験によって「SF占有率」と主観評価値との間に有意な関係が見られたことは、それら従来の先行研究と整合的であった。つまり、音高の変化、母音の変化を伴う歌唱において、それぞれの区間における2音間の「歌手のフォルマント」の占有率が高ければ、聴取者に良い聴覚印象を与え、高い評価を得るということである。

次に「SF占有率変化量」の結果について考察する。「SF占有率変化量」と「主観評価値」の要因は、図6-4,図6-5で示された通り有意となった。この結果から、「歌手のフォルマント」の占有率の変化量も、聴覚印象に関わっているということが今回新たに分かったことになる。そして偏回帰係数が負の相関を示している通り、「SF占有率変化量」が少ないほど歌声の主観評価値が高くなることから、音高の変化、母音の変化を伴う歌唱において、それぞれの区間における2音間の「歌手のフォルマント」の占有率の変化量が、より少ないことが聴取者に良い聴覚印象を与え、高い評価を得るということが示された。

そして、歌声のスペクトルのグラフから変化量を算出したRMS変化量の中の、「極大点変化量 正規化あり」について有意差が確認された。「極大点変化量 正規化あり」と「主観評価値」の要因は、図6-4,図6-5で示された通り有意となっている。また、ここでは偏回帰係数が正の値を示している。この結果は、正規化後の極大点変化量が少ない方が聴取者の高い評価を得るということを示し、フォルマントが安定しているほど高い評価を得るという予想とは矛盾しているように思える。そこで、RMS指標の結果について考察してみることにする。

まず、先述したように「2-4kHzの変化量正規化なし」のRMS指標については、「SF占有率変化量」とかなり似通った音響特徴量であった。しかし、極大点を合わせた前後500HzのRMS変化量を正規化した「極大点変化量 正規化あり」は、縦方向(パワー正規化 図4-8参照)と横方向(極大点合わせ 図4-7参照)の2方向において補正されたものであり、「SF占有率変化量」とは全く異なった音響特徴量である。そして、それらのパワー軸方向への正規化と周波数軸方向の補正により、「歌い手のフォルマント」のグラフの山の頂点の形状を捉えられるようになったと考えられる。変化する理由の一つとして、「歌い手のフォルマント」をできるだけ大きく出すためだということが考えられる。フォルマントは、母音や基本周波数の変化に影響される。そこで優れた歌い手は、それらの変化に応じてフォルマントを調整し、「歌い手のフォルマント」を高く維持していると考えれば、「極大点変化量 正規化あり」の正の相関は妥当な結果であると言える。またそう考えるなら、頂点のズレやパワーの違いという2要因を片方ずつ補った、極大点を合わせた前後500HzのRMS変化量「極大点変化量 正規化なし」と2kHzから4kHzのRMS変化量「2-4kHz変化量正規化あり」については、「歌い手のフォルマント」のグラフの山のズレは一方向でしか補正されていないために、「歌い手のフォルマント」の頂点の形を捉える特徴にはならず、主観評価値との関係を示す適切な指標にはならなかったと言える。

よってこれらのことから、「歌い手のフォルマント」の占有率が高く占有率の変化量が少ない時には、「歌い手のフォルマント」のスペクトルの山の頂点の形は変化する方が聴取者に良い聴覚印象を与え、高い評価を得ることが示された。

## 6.4 まとめ

本章では、第4章で行なった音高の変化、母音の変化を含む歌唱課題の音響分析結果と、第5章で行った複数の評定者による印象評価実験を照らし合わせることで、「歌手のフォルマント」の安定性と印象評価との関係を調べた。実験結果から、母音の変化、音高の変化を伴う歌唱課題において、2音間の「歌手のフォルマント」の占有率は、その値が高いほど、印象評価実験において高い評価が得られる事がわかった。

また新たにこの実験で明らかになったことは、「歌手のフォルマント」の占有率は、「母音の変化」「音高の変化」を伴う歌唱課題の2音間において、その変化量が少ない方が、印象評価実験において高い評価を得られるという事である。そして同時に、「歌手のフォルマント」をスペクトルのグラフからRMSによって比較すると、そのフォルマントの山の頂点の形は歌唱課題の2音間ににおいて変化する方が、印象評価実験において評価が高くなるということも示された。

## 第7章 総合考察

## 総合考察

本論文では、歌唱における「歌い手のフォルマント」に注目し、従来の「歌い手のフォルマント」の出方の顕著さに加えて、その安定性に注目した。その際に、実験の条件を実際の歌唱の場合に近づけるために楽曲における旋律の「音高の変化」と歌詞の中の「母音の変化」を伴う歌唱課題を作成し、歌唱者によって録音されたサンプルを用いて音響計測実験を行なった。そして音声分析によって「歌い手のフォルマント」に関する特徴量を見出し、また同時に印象評価実験を行うことによって、それらの歌唱課題に対する主観評価を得た。それにより音声分析によって見出された特徴量が、聞き手の聴覚印象にどのように影響し、また歌唱技術とどのように関わっているのかを調べるためである。「SF占有率」「SF占有率変化量」「2-4kHz正規化なし」「2-4kHz正規化あり」「極大点変化量正規化なし」「極大点変化量正規化あり」という6つの特徴量について印象評価実験から得られた「主観評価値」と相関分析をしたところ、

「SF占有率」と「SF占有率変化量」、そして「極大点変化量正規化あり」に有意性が見られた。

以上のことを踏まえて、本実験で得られた全ての結果について、総合的に考察していく。

本研究では、「音高の変化」と「母音の変化」の2要素を直交配置した歌唱課題において、「歌い手のフォルマント」がどのように存在しているのかを調べるために、学生とプロ歌手を含む18名から採取したの歌声についてスペクトル分析を行った。また「歌い手のフォルマント」の安定性を測るために、「歌い手のフォルマント」自体が、歌声のスペクトル全体の中でどのくらいの割合を占めているのかという「歌い手のフォルマント」の「占有率」に着目した。そこで明らかになったことは、印象評価実験の結果として高評価を受ける歌唱は、歌唱課題の2音間における「歌い手のフォルマント」

の占有率が高い傾向にあるということである。これは従来の先行研究と整合的である。歌い手のフォルマントの存在は、音声研究の分野ではすでに明らかにされてきたことであったが、実際に歌唱指導の現場においては、まだ浸透した知識ではなく、その概念を知る歌唱指導者は少ない。「お腹から声を出す」「体を使って声を出す」等の教示によって、結果的に歌い手のフォルマントが出る歌唱技術を獲得することに繋がることはあるが、あくまでもそれは正しい身体の使い方に対するイメージであり、それをイメージして訓練することと、実際に行われている動作は異なっている可能性が十分にあり得る。

また「占有率」は、歌い手のフォルマントの大きさをうまく指標化できる手法であり、今回の実験においてもその手法の有効性を再確認することができた。この方法は実際のフォルマントを推定するのではなく、あらかじめ決められた周波数帯のパワーを占有率で指標化した。3kHz付近の周波数帯域(今回の実験では2-4kHz)に意味があり、歌声にとって音響的な利点があることから、その範囲内に取まっているパワースペクトルが重要であると考えられ、3kHz付近に限定された周波数帯のパワーを比べることにより、有用な指標を得られたと確信している。ただし、この「占有率」の高さだけでは、本研究の目的である歌唱を通じての「変化」を見たことにはならず、その問いに対して全く調べたことになっていない。つまり、この結果だけでは、ただ大きなパワーで歌えば良いということになってしまい、占有率の「高さ」だけでは、歌唱技術の全てを説明することができないのである。

歌い手のフォルマントの「変化」に注目すると、占有率の高さだけでは説明しきれない部分がある程度説明できることが分かった。本研究で新たに分かったことは、印象評価実験の結果として高評価を受ける歌唱は、歌唱課題の2音間における歌い手のフォルマント占有率の変化量が少ない傾向にあるということである。重回帰分析によって「SF占有率変化量」が「SF占有率」に対して独立性の高い要因であり、「主観評価値」

と有意となったことから、「SF占有率変化量」が単なる占有率の大きさとは独立した要因で十分に聴覚印象に良い効果を与えてる証拠であると言える。声楽において、音高の変化や母音の変化における歌声の音色や響きを統一させることが第一義的な課題であり、それは優れた歌手ならば必ず持ち合わせている歌唱技術の一つである。この「響きの統一」ということと、占有率変化量の少なさということは非常に整合性が高いと言える。

本章の冒頭で考察したように、RMS変化量について「主観評価値」と回帰分析を行ったところ、「極大点変化量正規化あり」において有意性が見られた。その相関は正の相関を示していることから、歌声のスペクトルのグラフから得られた「歌手のフォルマント」の山の頂点の形は、変化した方が聞き手の聴覚印象に高い評価をもたらすということが示された。これは「歌手のフォルマント」の占有率の変化量だけではなく、「歌手のフォルマント」のRMS変化量も、歌唱技術の良好性をより適切に判断することを可能とする指標であるということを示している。但し、正の相関が見られたということは、一見すると「歌手のフォルマント」の安定性という当初の予想と矛盾する結果である。「SF占有率変化量」は、変化が少ない方が評価が高いということであったが、「極大点変化量 正規化あり」では、変化が多いほど評価が高くなるという結果が出ているからだ。このようなことから「極大点変化量 正規化あり」という特徴量と、「歌手のフォルマント」の山の頂点の形が変化する原因についてさらに考察を試みる。

まず、「SF占有率変化量」と「極大点変化量 正規化あり」はそれぞれ異なる特徴を捉えた指標であり、同じ特徴量ではない。「極大点変化量正規化あり」は、歌唱課題における2音間の「歌手のフォルマント」から、重回帰分析によって占有率の高さとその変化量という音響特徴量が取り除かれたものであり、「歌手のフォルマント」の山の頂点の形を示す指標となったのである。そこから分かったことは、優れた歌手

であっても、それぞれの音域や母音において「歌い手のフォルマント」の極大点の位置や、その山の頂点の形は変化しているということである。その理由には、次のようなことが考えられる。例えば、「歌い手のフォルマント」の極大点の周波数やそのスペクトルの山の頂点の形は、第1フォルマントと第2フォルマントの出方によって大きく変化している。つまり、歌唱における母音の変化を明確にするために、よりはっきりした母音で発音すれば、高次フォルマントの位置やその頂点の形は、それに合わせて変動するのである。また、音高の変化によりそれぞれの区間における歌声全体のパワーは変化する。例えば、高い音に跳躍すれば、全体のパワーはそれに伴い大きくなり、低い音に跳躍すれば、全体のパワーは低くなる傾向にある。「歌い手のフォルマント」の山の土台に相当する全体のパワーレベルが、異なる音高間において変動する中で、「歌い手のフォルマント」の占有率をより高く維持し、変化量を低く保つために、基本周波数に応じて「歌い手のフォルマント」の山の頂点の形を変えているとも考えられる。あるいは、頂点の形自体が音色の微妙な変化を生み、それが「歌い手のフォルマント」の占有率の劇的な変化に繋がらない程度であるならば、むしろ聴覚印象に良い効果をもたらすという可能性も考えられる。

以上のように考えるとこの表面的な矛盾は解決できる。しかしながら、これはあくまでも暫定的な解釈に過ぎず、他の可能性も含めて今後さらに探究すべき課題は残されている。

ここまで、本実験で得られた全ての結果について、総合的に考察してきた。本論文の目的は、歌唱における「歌い手のフォルマント」について調べることであり、今回の実験において、「歌い手のフォルマント」の占有率とその変化の少なさは、少なくともクラシック音楽を学ぶ学生やプロの演奏家にとっては高い評価を受ける傾向があることが確認できた。では、なぜ「歌い手のフォルマント」が重要なのであろうか。



そもそもフォルマントというものがどういうものであり、なぜそれは歌声に必要であったのかをもう一度考えてゆく。

まず、第2章で述べたように、フォルマントとは、母音の生成に大きく関わる周波数群のことを意味していた。その周波数群は、声道の形状の声道断面関数によって得られた共鳴特性から成り、それぞれの母音や発生された音の高さによって変化するものである。歌声の研究においては、母音の生成に関わる第1フォルマントと第2フォルマントの研究だけでなく、第3フォルマント以降の、歌声の音色に関わる部分について多くの研究がなされてきた。そして、優れた歌手に共通する音響的特徴として導き出されたものが「歌い手のフォルマント」だったのである。「歌い手のフォルマント」は、3kHz付近の周波数帯域に存在し、その周波数帯域は次のような特性を備えていた。まず、3kHz付近という周波数帯域は、人間の聴覚において最も可聴しやすい音域である。また、楽器から発せられる楽音にはない周波数ピークであり、楽器による伴奏を伴った歌唱においては、歌声を楽音にマスクされない状態で聞き手に届けるために、重要な役割を担う周波数帯域であった。これら2つの利点は、拡声技術がなかった（未熟だった）時代には特に重要な要因であり、より大きな音を出す物に進化してきた楽器を伴奏とした場合に、それに負けない歌声の「存在」感をもたらすための音響的特徴として、「歌い手のフォルマント」は合理性のあるものだった。先行研究では、楽曲中の「歌い手のフォルマント」を、長時間スペクトルの分析によって、その存在を示してきた。しかし、今回の実験結果から示されたように、「歌い手のフォルマント」は楽曲中のある一瞬、あるいは楽曲中の要所に群発的に存在していることによってその効果が表れているのではなく、F0や音韻（音高や母音）の変化の中で、できるだけ安定して保たれていることが重要なのである。それらの変化の中で、声道の形状や声帯の形状を巧みにコントロールし、「歌い手のフォルマント」を維持し続ける技術を習得す

ることは困難であり、その境地に達した歌手は、希少価値の高い「特別」な存在として認められてきたという事実は、納得のいくことである。

序論では、歌声の「理想」像の一例としてベル・カント唱法について述べた。この歌唱法の確立とオペラという音楽劇の進化の間には、先に述べたように密接な関係がある。西洋において声楽が興り、オペラが発展していく中で、演奏家たちは何某かの「理想」を見言い出し、それに向かって進むようになる。何を「理想」とするかは、その時代、その地域によって様々に変容があると考えられるが、少なくとも当時の観客たちが、優れた歌手から発せられる歌声が、常人では実現することのできない「特別」なものとして受容してきたからこそ、その歌唱法の伝統が「理想」として伝授されているはずである。数多くの演奏家や声楽教師たちが、それが何であるかということを探し求め、その一つがベル・カント唱法であったのだ。しかし、この言葉には曖昧な点が多く、歴史的な意味や発声法としての意味合いなど、多様な解釈がある。オペラが発展するにつれて、オーケストラの楽器は大音量を出せるように進化し、大編成になっていった。その一方で、歌手は依然として一人でこの要求を満たさなければならなかったのである。だが、当時は今日のような歪みの少ない拡声システムはなく、その限られた条件の中で聴衆に歌声を届けなければならなかったことから考えると、当時のベル・カント唱法とは、少なくともある一定以上のレベルに達した歌手の歌声を意味していたのだと推測できる。最初の研究者たちは、ベル・カント唱法を解剖学的な見地から解き明かそうとした。歌声は目に見えないが、その発声器官ですら外部からは観察できなかったからである。第1章で取り上げた発声器官の仕組みは、これらの研究者たちの研究の積み重ねによって解明されてきたことの総括である。そして解剖学的な研究がもたらした発声器官の仕組みやその作用は、演奏家や声楽教師たちに沢山の知識をもたらした。しかし、それを発声の「真実」とであると勘違いした者が次のような指導を行うようになったのである。「もっと声帯を伸ばして」「もっと喉頭

を広げて」「もっと横隔膜を下げて」これらの指導法は、一見すれば、それまで雲をつかむような指導によって苦しんできた学習者にとっては「真実」を帯びた言葉だったかもしれない。しかし、実際にそれを実現し、声楽教師の求める歌声を再現しようとしても、それが「理想」の歌声を与えてくれることは決してない。つまり、発声器官の仕組みやその作用を知っていたとしても、それは単なる知識であり、それ以上のものにはならなかったのである。

ここで、もう一度ベル・カントをはじめとする伝統的発声を守り続けてきた演奏家や声楽教師が、どのようにして歌声を鍛錬してきたのかを考えると、その答えは「耳」にある。序論でも述べたが、優れた歌手や声楽教師は、歌声を聞き分ける優れた「耳」を持っていた。つまり、「発せられる声が正しければ、すべての発声器官は正しく作用している」という考えを基に、その「耳」で歌声を判別していたのである。そして声楽教師が聞き分けた「正しい声」「誤った声」の二者択一を繰り返すうちに、「どのように聞こえる声が正しい声なのか」を見出していくのだ。ここで大事なことは、どのような肉知的アプローチを行い、発声器官をどのような状態で用いれば正しい声になるのかということではなく、「発せられる声が正しければ、すべての発声器官は正しく作用している」ということこそ、歌声の技術獲得において重要であると再認識することである。つまり、答えは全てその「歌声」にあり、それらを絶対的に判断することのできる「耳」こそが、「理想」像への唯一の手がかりなのである。しかし、Huslerが、「かつて耳から耳へ、名人から弟子へと受け継がれていた、あの発声教育に関する伝承された、まとまった「知識」は…崩れ、あまりにも多くの流派に分裂した」（Husler, 1987）と述べたように、現代において理想的な声を確実に聞き分け、明確に示すことは、厳しい現状であると言わざるを得ない。

このような問題の中にあって、「耳」に変わる新たな判別方法の一つとして考えられたのが、科学による音声分析なのである。音声分析より得られた音響データは、歌

声を目に見える形で示す事ができ、客観性を帯びたデータとして歌声を示す事ができる。そして、その研究から、理想的な歌声を示すいくつかの音響特徴量が見出された。その中の一つが本論文で研究対象とした「歌手のフォルマント」なのである。少なくとも、この「歌手のフォルマント」は、ベル・カント唱法を含む歌声の「理想」像の一部であると確信を持つていうことができ、歌声を判別する上で重要な役割を果たすということが、今回の研究においても示された。確かに、断片的な一側面でのみの結果を示しているに過ぎないが、これらの研究を続けることにより、確実に「理想」な歌声とはどういうものなのかを、客観的に示していくことが出来ると考えている。そして一人でも多くの演奏家や音楽教師が、これらの科学的な視点を取り入れることによって、よりいっそう歌声に対する認識を深め、演奏や指導の現場に生かすことができれば、これらの研究が、さらに意義のあるものになるのではないかと思う。

## 結論

本論文は、歌声のスペクトルに含まれる「歌い手のフォルマント」に注目し、歌声を音響分析によって観察し、印象評価実験との相関を見ることによってその時間的なふるまいを観察してきた。その結果、2音間における「歌い手のフォルマント」の占有率は、高い方が聴覚印象に高い評価をもたらすという事が示された。この結果は、従来の先行研究と整合的であった。また今回新たに、2音間における「歌い手のフォルマント」の占有率の変化量は少ないほど、聴覚印象に高い評価を与えるという事がわかった。そして、2音間において、歌声のスペクトルのグラフから、「歌い手のフォルマント」の波形の山を比べた結果、その山の形は、変化した方が聴覚印象に良い評価をもたらすということも示された。

先行研究において、歌声から得られた音響的特徴と印象評価実験を相関比較する研究について、今回のような規模で行われた実験の前例は少なく、その中でいくつかの新しい発見があったことには大きな意味があった。中でも、歌唱において優れた歌い手は「歌い手のフォルマント」の占有率を高く保ちながら、できるだけ変化させないように歌っており、そのことが聴覚印象において高い評価を得るという結果を得られたことは、大きな成果であった。

## おわりに

本研究では、優れた歌手の歌声とはどのようなものなのか、理想的な歌声とは何なのかということについて言及してきた。この問いは、西洋において声楽が生まれ現在に到るまでの無数の演奏家や声楽教師、そして声楽を志す全ての学習者にとって切実な問いである。そのような中であって、ここにその「理想」像の一端を目に見える形で示すことができたことは、声楽家である筆者にとって意義のあることである。

本論文は、見えない歌声をどのようにして目に見える形で示すことができるだろうか。全ての声楽を志す者が生涯をかけて追い求める「理想」像とは一体なんなのだろうか。という問いに対して、科学的なアプローチによってその答えを見出そうと試みたものであった。もちろん、本論文で示すことができたのは、奇跡にも近い人間の歌声の、複雑で膨大なメカニズムのごく一部にしか過ぎない。しかし、本論文で示された歌声に関する指標は、演奏家、声楽教師、そして声楽を学ぶ学習者にとって、十分に役立つものになったと自信を持って言うことができる。ここで示された事が、さらなる研究のための礎となり、一人でも多くの声楽を志すものに役立てれば幸いである。

## 文 献

- Raphael, L. J, Harris, K. S, Borden, G. J, “Speech Science Primer- Physiology, Acoustics, and Perception of Speech 5th ed”Lippincott Williams & Wilkins Inc, USA, 2007 (Gloria J. Borden, Katherine S. Harris, Lawrence J. Raphael, 廣瀬肇訳, “新 言葉の科学入門 第2版”医学書院, 東京, 2004)
- Husler, F, Marling, Y. R, “Die physische Natur des Stimmorganes-Anleitung zum Aufschliessen der Singstimme”B. Schott’s Söhne, Mainz, 1965(フレデリック・フースラー, イヴォンヌ・ロッド=マーリング, 須永義雄, 大熊文子訳, “うたうこと”音楽之友社, 東京, 1987)
- Reid, C. L.“BEL CANTO:Principles and Practices”1950 (コーネリウス・L・リード, 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)
- Sundberg, J, “THE SCIENCE OF THE SINGING VOICE” Northern Illinois University Press, 1987(ヨハン・スンドベリ, 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)
- 今泉敏, “言語聴覚士のための音響学”医歯薬出版株式会社, 東京, 2007
- 北村達也, “情動による声道形状変化のMRI観察”日本音声学会全国大会予稿集, pp45-50 (2009)

斎藤毅, 辻直也, 鶴木祐史, 赤木正人, “歌声らしさに影響を与える音響特徴を考慮し

た話声からの歌声合成法”信学技術(SP)2004-30, pp1-6, 2004

斎藤毅, 辻直也, 鶴木祐史, 赤木正人, “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巻5号, pp267-277, 2008

斎藤毅, 後藤真孝, “歌唱指導による音響特徴の変化とその歌唱力評価への影響”信学

技術(EA)2009-18, pp2-6, 2009

水谷彰良, “ベル・カントの歴史様式とその解釈”日本声楽発声学会誌, 第34号, pp33-

43, 2006



## 参考文献

- Caesari. E. H, “The VOICE of the MIND”Robert Hale Limited, England,  
1951. (E. ハーバート-チェザリー, 森下弓子訳, “The VOICE of the MIND”株  
式会社ワニブックス, 東京, 2001)
- Kawahara, H, Masuda-Katsuse, I, Alain de Cheveigné, ”Restructuring speech  
representations using a pitch-adaptive time-frequency smoothing and an  
instantaneous-frequency based F0 extraction: Possible role of a repetitive  
structure in sounds”1999
- 大谷圭介, 津崎実, ”声区転換部を含むオペラ歌唱の音響的特徴—スペクトル変動に見  
る音響的指標について—”音楽知覚認知研究学会誌, Vol.19, No.1, pp3-18, 2013
- 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.
- 北村達也, “物真似タレントによる物真似音声の分析”信学技術(SP)2007-79, pp49-54,  
2007
- 北村達也, “歌声の科学にあたって”日本音響学会誌, 70巻9号pp497-498, 2014

北村達也, 斎藤毅, ”単母音の特徴量の変化が個人性知覚に与える影響”信学技術

(SP)2006-167, pp43-48, 2007

北村達也, 斎藤毅, ”3連続母音に含まれる個人性知覚に寄与する音響特徴量”信学技術

(SP)2006-166, pp37-42, 2007

北村達也, 竹本浩典, 足立整治, パーハム・モクタリ, 本多清志, ”声帯振動に伴う喉

頭腔共鳴の周期性”信学技術(SP)2006-34, pp43-48, 2006

斎藤毅, 鵜木祐史, 赤木正人, ”自然性の高い歌声合成のためのヴィブラート変調周波

数の制御法の検討”電子情報通信学会, 2005

竹本浩典, 北村達也, パーハムモクタリ, 足立整治, 本多清志, ”声門開放時の声道共

鳴特性”電子情報通信学会, 2007

竹本浩典, 北村達也, パーハム・モクタリ, 足立整治, 本多清志, ”声門解放時の声道

共鳴特性”信学技術(SP)2006-165, pp31-36, 2007

吉田友敬, 言語聴覚士の音響学入門, 海文堂出版株式会社, 東京, 2005

## 本研究に関する研究業績

### 〈紀要に発表した論文〉

高橋純 「歌い手のフォルマントについての一考察 -ベル・カント唱法と科学的研究を比較して-」, 京都市立芸術大学音楽学部大学院研究紀要『ハルモニア』, 2017年

### 〈国内学会における発表ポスター/査読なし〉

高橋純 津崎実 「歌唱時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-」, 聴覚心理学会ビギナーズセッション, 2016年

### 〈研究助成〉

第33回カワイサウンド技術・音楽振興財団サウンド技術部門研究助成(共同研究)

「歌唱時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-」 2016年

## 謝辞

本研究を進めるにあたり、ご指導を頂いた、論文指導教官の津崎実教授に心より感謝申し上げます。

また、声楽に関するご指導を頂いた、実技指導教官の小濱妙美教授に心より感謝申し上げます。

それから、副指導教官としてご指導を頂いた、砂原悟教授に心より感謝申し上げます。加えまして、本研究のために快く実験に参加して下さった本学教員の皆様、並びに学生の皆様に深く感謝いたします。

最後に、日常の議論を通じて多くの知識や示唆を頂いた、津崎研究室の皆様に感謝いたします。