

Title	歌声らしさに影響を及ぼす音響的特徴の分析に関する研究
Author(s)	辻, 直也
Citation	
Issue Date	2004-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1790">http://hdl.handle.net/10119/1790</a>
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

修士論文

歌声らしさに影響を及ぼす音響的特徴の分析  
に関する研究

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

辻直也

2004年3月

# 修士論文

## 歌声らしさに影響を及ぼす音響的特徴の分析 に関する研究

指導教官 赤木 正人 教授

審査委員主査 赤木 正人 教授

審査委員 宮原 誠 教授

審査委員 党 建武 助教授

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

210058 辻 直也

提出年月: 2004 年 2 月

## 概要

本稿では、歌声らしさという心理的特徴と物理的特徴の関係について検討するために歌声らしさの3層モデルを提案し、それぞれに関係について検討を行う。モデルの第1層目（歌声らしさ）と第2層目（基本的な心理的特徴）の関係において、MDSと重回帰分析より、声の揺れ、響き、明瞭さは歌声らしさの要因であることが分かった。モデルの第2層目（基本的な心理的特徴）と第3層目（物理的特徴）の関係において、揺れと響きに関連する音響的特徴の抽出を行い、聴取実験から揺れや響きとの対応関係について検証した。その結果、F0と振幅エンベロープの4~6 Hzの変位、又、フォルマントの周波数変調や振幅変調、さらにこれらの関係が同位相であることは揺れに関連する音響的特徴である事が分かった。また、スペクトル包絡における3~4 kHzの見られる顕著なピークと、3~4 kHzの強い高調波成分は響きに関連する音響的特徴であることが分かった。最後に、モデルの第1層目（歌声らしさ）と第3層目（物理的特徴）の関係において、第3層目の揺れや響きに関連する音響的特徴の組み合わせを変えて、第2層目の揺れや響きの心理量を変化させる事により、第1層目の歌声らしさに対する知覚が変化するか検証を行った。その結果、第3層目の揺れや響きに関連する音響的特徴を組み合わせる事により、第1層目の歌声らしさが変化したことから、モデルの各層の対応付けができていたことが示された。

# 目次

第1章	序論	1
1.1	はじめに	1
1.2	背景	1
1.3	本研究の目的・特色	3
1.4	本論文の構成	3
第2章	歌声らしさの心理的特徴	5
2.1	音声データの順位付け	5
2.1.1	使用する音声データ	5
2.1.2	実験 2.1 (音声データの順位付け)	6
2.1.3	音声データの選択	8
2.2	歌声らしさの心理的空間	9
2.2.1	実験 2.2 (歌声らしさ)	9
2.2.2	MDS による歌声らしさの心理的空間	11
2.3	まとめ	14
第3章	歌声らしさの要因の検討	15
3.1	表現語の選択	15
3.1.1	表現語に関する先行研究	15
3.1.2	実験 3.1	15
3.1.3	実験 3.2	16
3.1.4	本研究で使用する表現語の選択	19
3.2	実験 3.3 (表現語)	19
3.3	歌声らしさの空間における各表現語の表す方向	22
3.4	まとめ	24
第4章	歌声らしさの要因に関連する音響的特徴の抽出	25
4.1	揺れに関連する音響的特徴の分析	25
4.1.1	揺れに関連する音響的特徴の調査	25
4.1.2	F0 と振幅エンベロープの変動	27
4.1.3	フォルマントの周波数方向や振幅方向への変動	30

4.2	響きに関連する音響的特徴の分析	36
4.2.1	響きに関連する音響的特徴の調査	36
4.2.2	スペクトル包絡における 3 kHz 付近のピーク	38
4.2.3	3 kHz 付近の強い高調波成分	39
4.3	STRAIGHT を用いた合成法の概要	43
4.4	実験 4.1 ( SYN-BASE の検証 )	46
4.5	揺れに関連する音響的特徴の検証	46
4.5.1	実験 4.2 ( 揺れ )	46
4.6	響きに関連する音響的特徴の検証	51
4.6.1	実験 4.3 ( 響き )	51
4.7	まとめ	57
<b>第 5 章</b>	<b>歌声らしさに影響を及ぼす音響的特徴の検討</b>	<b>58</b>
5.1	歌声らしさに影響を及ぼす音響的特徴の検討	58
5.1.1	実験 5.1 ( 歌声らしさ )	58
5.2	まとめ	60
<b>第 6 章</b>	<b>結論</b>	<b>61</b>
6.1	本論文のまとめ	61
6.2	今後の課題	62

# 目次

1.1	歌声らしさに関する先行研究	2
1.2	歌声らしさの3層モデル	2
2.1	実験システムの構成	6
2.2	評価尺度(話声らしさ-歌声らしさ)	7
2.3	評価尺度(歌声らしさ)	10
2.4	Stress と次元数の関係	12
2.5	歌声らしさの心理的空間と各平面から見た音声データの布置	13
3.1	評価尺度(表現語)	17
3.2	各表現語の評価尺度(揺れの場合)	20
3.3	各表現語の心理的距離	21
3.4	各表現語を表す方向	23
4.1	F0(上)、振幅エンベロープ(中)、スペクトル(下)の分析結果	26
4.2	F0の時間変化と振幅エンベロープの関係	29
4.3	no.03(1位)のフォルマントの変動	31
4.4	no.5(7位)のフォルマントの変動	32
4.5	no.10(9位)のフォルマントの変動	33
4.6	F0とフォルマントの周波数変調や振幅変調の関係	35
4.7	スペクトル(上)、非周期成分の割合(下)の分析結果	37
4.8	時間平均したスペクトル包絡の比較	38
4.9	no.3(上),no.1(中),no.9(下)の非周期成分の割合の時間変化	40
4.10	高調波成分と3 kHz 付近のピークの関係	41
4.11	高調波の時間変化と3 kHz 付近のピークの時間変化	42
4.12	STRAIGHT のを用いた合成手法	44
4.13	SYN-BASE の作成方法	45
4.14	SYN-BASE(上)とSYN-F0(下)のF0	47
4.15	SYN-SPEC の作成方法	48
4.16	SYN-BASE(上)とSYN-EBV(下)の振幅エンベロープ	49
4.17	評価尺度	50
4.18	合成音の揺れの関係	51

4.19 SYN-UPの作成方法 . . . . .	52
4.20 SYN-UPのスペクトルと非周期成分の割合 . . . . .	53
4.21 SYN-MOVEのスペクトル包絡のピークと高調波の操作 . . . . .	54
4.22 SYN-MOVEのスペクトルと非周期成分の割合の操作 . . . . .	54
4.23 SYN-MOVEのスペクトルと非周期成分の割合 . . . . .	55
4.24 評価尺度 . . . . .	56
4.25 合成音の響きの関係 . . . . .	57
5.1 評価尺度 . . . . .	59
5.2 合成音の歌声らしさの関係 . . . . .	60

# 表目次

2.1	実験機材 . . . . .	6
2.2	音声データの順位 . . . . .	8
2.3	選択した音声データ（ 、 、 の記号が同じものは同一発声者） . . . . .	9
2.4	各音声データの心理的距離 . . . . .	11
2.5	Stress の評価 . . . . .	12
3.1	被験者が回答した表現語とその人数 . . . . .	16
3.2	両側尺度 . . . . .	17
3.3	片側尺度 . . . . .	17
3.4	各音声データにおける表現語の評価値 . . . . .	18
3.5	表現語の重相関係数 . . . . .	19
3.6	各表現語における偏回帰係数 . . . . .	22
3.7	各表現語を表す方向の関係と重相関係数 . . . . .	22
4.1	F0 と振幅エンベロープにおける変調周波数と偏移幅の分析結果 . . . . .	27
4.2	F0 の時間的变化と振幅エンベロープの相関値 . . . . .	28

# 第1章 序論

## 1.1 はじめに

歌声は話声にはない特有の音響的特徴を持つ。それ故、歌声特有の音響的特徴は、歌声らしさという心理的特徴に重要な要因であると考えられる。歌声合成のような工学的応用を考えれば、歌声をより歌声らしくするには、音響的特徴の量的な変化や音響的特徴同士の相互関係が歌声らしさに及ぼす影響を考える必要がある。そのためには、歌声らしさと歌声特有の音響的特徴との対応関係が重要となる。しかし、歌声に含まれる音響的特徴は複雑で動的な特性を持つため、それらの対応関係は十分には検討されていないのが現状である。

歌声らしさと歌声特有の音響的特徴との対応関係を明らかにし、それらの音響的特徴が歌声らしさに及ぼす影響について検討する事は、歌声合成における工学的応用への貢献だけでなく、歌声知覚の研究に多くの示唆を与える可能性がある。

## 1.2 背景

歌声らしさに関連する先行研究は、図 1.1 に示す、歌声らしさと物理的特徴との関係に着目したもの (A) と、歌声らしさと心理的特徴との関係に着目したもの (B) に分けて考える事ができる。

はじめに、歌声らしさと物理的特徴の関係に着目したもの (A) について示す。Sundberg は、話声にはない歌声特有の音響的特徴である Singing formant の存在を明らかにした [1]。しかし、歌声特有の音響的特徴については分析されているものの、それらの特徴の量的な違いが歌声らしさに与える影響についての検討は十分には行われていない。一方、齋藤らは歌声における基本周波数の動的変動成分であるヴィブラート、オーバーシュート、予備的変動などが歌声らしさに及ぼす影響について定量的な検討を行い、その重要性を指摘している [2]。

次に、歌声らしさと心理的特徴の関係に着目したもの (B) について示す。西内・大串は、歌声らしさについて分析を行い、複数の基本的な心理的特徴が重要であることを示した [3]。しかし、それらに関連する音響的特徴の分析は十分には行われていないため、その音響的特徴の量的な変化が歌声らしさに与える影響についての検討は行われてはいない。

これらの先行研究の問題点は、歌声らしさという心理的特徴と物理的特徴との対応関係が十分に検討されていなかったことである。なぜならば、西内・大串が指摘するように、

歌声らしさは複数の基本的な心理的特徴から構成されているため、物理的特徴との直接的な対応付けが困難だからである。

そこで、本研究では、先行研究の問題点である歌声らしさと物理的特徴との対応関係について検討を行うために、図 1.2 に示す歌声らしさの 3 層モデルを提案する。

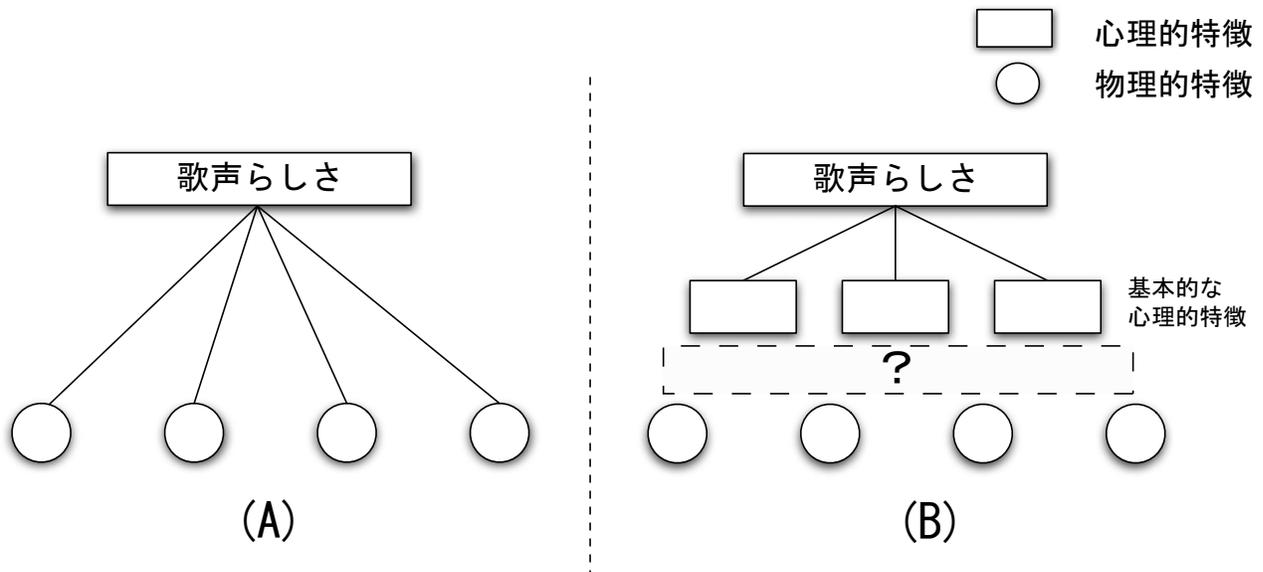


図 1.1: 歌声らしさに関する先行研究

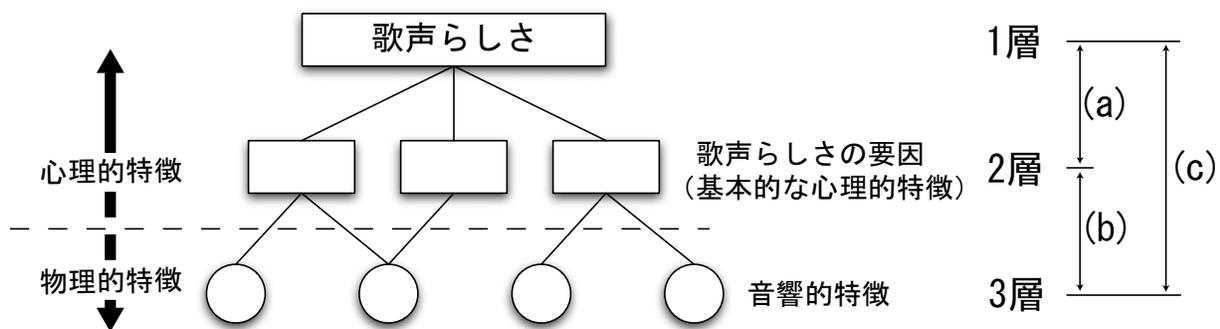


図 1.2: 歌声らしさの 3 層モデル

## 1.3 本研究の目的・特色

本研究の目的は、図 1.2 に示す歌声らしさの 3 層モデルにおいて、歌声らしさという心理的特徴を基本的な心理的特徴で表し、それらに関連する音響的特徴と心理的特徴との対応付けを行うことである。

本研究の特色について以下の 2 点が挙げられる。(1) 歌声らしさという高次の心理的をより低次の基本的な心理的特徴に分解するため、心理的特徴と音響的特徴との対応付けが行いやすい。(2) 音響的特徴の定量的な変化や音響的特徴同士の相互関係が歌声らしさに及ぼす影響などの詳細な検討が可能となる。

## 1.4 本論文の構成

本論文は 6 章で構成される。

### 第 1 章

本論文が対象としている研究分野の背景と問題点を明らかにし、本論文の目的と特色を示す。

### 第 2 章

歌声らしさの心理的特徴について述べる。

はじめに、本研究で分析の対象とする音声データを選択するために、各音声データの歌声らしさの度合いについて順位付けを行う。そして、多次元尺度構成法 (MDS: Multidimensional scaling) により歌声らしさについて分析を行う。

### 第 3 章

歌声らしさの 3 層モデルのうちの第 1 層目の歌声らしさと第 2 層目の歌声らしさの要因について述べる (図 1.2(a) 参照)。

本研究では、歌声らしさの要因には表現語 (音色を表す形容詞) を用いる。そのため、使用する表現語について検討を行う。そして、歌声らしさの心理的空間において、重回帰分析により歌声らしさの要因について検討する。

### 第 4 章

歌声らしさの 3 層モデルのうちの第 2 層目の歌声らしさの要因と第 3 層目のそれらに関連する音響的特徴について述べる (図 1.2(b) 参照)。

歌声らしさの要因である声の揺れや響きに関連する音響的特徴を分析から抽出する。そして、第 2 層目と第 3 層目の関連について合成音による検証を行う。

## 第5章

歌声らしさの3層モデルのうちの第1層目の歌声らしさと第3層目の音響的特徴について述べる(図1.2(c)参照)。

聴取実験により、第4章で明らかにした音響的特徴を付加した合成音を作成し、歌声らしさに及ぼす影響を検討する。

## 第6章

本論文で得られた結果のまとめと今後の展望を述べる。

## 第2章 歌声らしさの心理的特徴

本研究では、各音声データの歌声らしさの度合いは、分析の対象とする音声データの選択や、分析結果を考察する際などにおいて重要な指標となる。そのため、各音声データを発声者や歌唱法で区別せず歌声らしいものから話声らしいものとなるように順位付けを行うことにより、各音声データの歌声らしさの度合いについて調査する。また、歌声らしさという心理的特徴は、どれくらいの心理的特徴から構成されているのか、MDSを用いて分析し検討を行う。

### 2.1 音声データの順位付け

#### 2.1.1 使用する音声データ

使用する音声データは、「日本語を歌、唄、謡う」のCD [4] に収録された話声や歌声の音声データの中から、より多くの歌唱法や発声者が含むように80個の母音/a/を選択した。これらの音声データは、CDに含まれる連続母音/a/,/i/,/u/,/e/,/o/から母音/a/のみを切り出した。例外として、80個のうち5個の音声データについては、母音がつながって発声されているため切り出しが困難であり、5母音の連続母音の音声データを用いた。また、洋楽（ソプラノ、メゾソプラノ、アルト、テノール、バリトン、バス）、わらべ歌、民謡、長唄、小唄、琵琶楽、歌舞伎、能、狂言、地歌、清元節、一中節、山田流箏曲、琉球古典音楽、詩吟、声明、新劇朗読、落語、アナウンサーの計19種類（パートも考慮すると計24種類）の歌唱法と38人の発声者が含まれている。

### 実験システム

図2.1に本研究で行われた全ての聴取実験の実験システムの構成を示す。また、表2.1には使用した実験機材を示す。

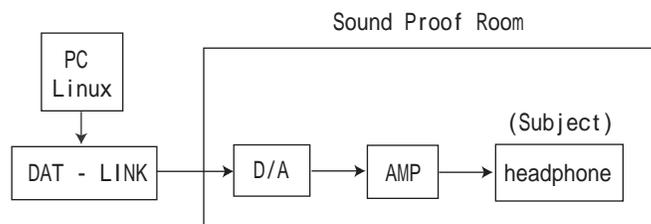


図 2.1: 実験システムの構成

表 2.1: 実験機材

呈示音サーバ	DAT + LINK & Linux
D/A 変換機	STAX DAC-TALENT BD
ヘッドホンアンプ	STAX SRM-1/MK-2
ヘッドホン	STAX SR-404

## 2.1.2 実験 2.1 (音声データの順位付け)

### 実験目的

多数の歌唱法や発声者が含まれた 80 個の音声データの順位付けを行うために聴取実験を行う。

### 実験方法

実験には、呈示した音声データに対して歌声らしいのか、または話声らしいのか図 2.2 に示す 5 段階の両側尺度を用いた。そして、被験者に個々の呈示音に対して絶対評価で回答させた。但し、実験を行う前には、被験者に実験の評価尺度に慣れてもらうためや歌唱法などで判断しないために予備的な訓練を行った。被験者は正常な聴力を有する大学院生 11 名であり、1 人の被験者について 3 回実験を行なった。但し、被験者には、発声者や歌唱法で判断しないように注意を与えた。

### 実験結果・考察

全被験者の評価値を平均し、80 個の音声データを歌声らしいものから話声らしいものとなるように順位付けを行った。その結果を表 2.2 に示す。歌唱法の後に示した番号は、順に、発声者の番号、発声音高の違い (番号が小さい程基本周波数が高い)、s の記号は

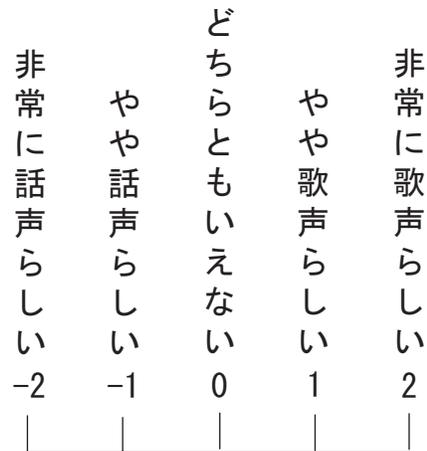


図 2.2: 評価尺度 (話声らしさ - 歌声らしさ)

話声 ( 歌声の場合はない ) を示す。

( 例 ) 表の順位が 1 位のテナー 1.2 は、歌唱法 : テナー、発声者番号 : 1、発声音高は同一発声者の中で 2 番目に高さ、の音声データである。

表の音声データを大きく分けると洋楽的唱法と邦楽的唱法に分かれている傾向がある。また、下位には話声の音声データが並んでいることが分かる。しかし、同じ発声者や歌唱法の音声データにおいて、大きく順位が異っているものがあることから、表の音声データは発声者や歌唱法ではなく、歌声らしさの度合いで順位付けされていると考えられる。また、各音声データにおける評価値についての分散分析の結果より、歌声らしさの順位の上位 35 個と下位 15 個の音声データでは評価値の分散は小さいが、それらの中にある 30 個の音声データでは評価値の分散が大きかった。そのため、2.1.3 節では音声データの選択を行うが、中間に位置する 30 個の中から音声データを選ぶ際には注意が必要であると考えられる。

表 2.2: 音声データの順位

順位	歌唱法	評価値	順位	歌唱法	評価値	順位	歌唱法	評価値	順位	歌唱法	評価値
1	テナー 1 <sub>2</sub>	2.00	16	バス 1 <sub>3</sub>	1.85	40	長唄 1 <sub>2</sub>	0.91	60	地歌 2	-0.67
1	ソプラノ 3	2.00	22	バリトン 1 <sub>1</sub>	1.82	42	詩吟	0.76	62	わらべ歌 1 <sub>1</sub>	-0.91
1	バス 1 <sub>2</sub>	2.00	22	バリトン 1 <sub>2</sub>	1.82	42	清元節 1 <sub>2</sub>	0.76	63	歌舞伎	-0.97
4	バリトン 2	1.977	24	メゾソプラノ 1 <sub>2</sub>	1.79	44	民謡 1 <sub>3</sub>	0.73	64	歌舞伎	-1.36
4	バリトン 1 <sub>4</sub>	1.97	25	民謡 2	1.64	45	声明 09	0.52	65	バス 1 <sub>s</sub>	-1.52
4	ソプラノ 4	1.97	25	メゾソプラノ 2	1.64	46	長唄 1 <sub>1</sub>	0.45	66	テナー 1 <sub>s</sub>	-1.55
4	バス 1 <sub>1</sub>	1.97	25	民謡 1 <sub>1</sub>	1.64	47	能	0.39	67	アルト 1 <sub>s</sub>	-1.73
4	バリトン 1 <sub>3</sub>	1.97	29	わらべ歌 1 <sub>4</sub>	1.45	48	琵琶楽 03 <sub>3</sub>	0.30	68	ソプラノ 1 <sub>s</sub>	-1.76
9	テナー 2	1.94	29	民謡 1 <sub>2</sub>	1.45	48	山田流箏曲	0.30	69	バリトン 1 <sub>s</sub>	-1.82
9	ソプラノ 2	1.94	30	ソプラノ 1 <sub>1</sub>	1.36	50	わらべ歌 1 <sub>2</sub>	0.24	70	新劇	-1.85
11	アルト 1 <sub>3</sub>	1.91	30	わらべ歌 1 <sub>3</sub>	1.36	51	声明 03	0.18	71	民謡 1 <sub>s</sub>	-1.88
11	メゾソプラノ 1 <sub>3</sub>	1.91	30	小唄 1	1.36	52	能	0.09	72	長唄 1 <sub>s</sub>	-1.91
13	ソプラノ 1 <sub>5</sub>	1.88	33	ソプラノ 1 <sub>2</sub>	1.24	52	山田流箏曲	0.09	73	落語	-1.94
13	ソプラノ 1 <sub>3</sub>	1.88	34	長唄 3	1.18	54	琵琶楽 03 <sub>2</sub>	0.03	73	洋楽 1 <sub>s</sub>	-1.94
13	テナー 1 <sub>3</sub>	1.88	35	メゾソプラノ 1 <sub>1</sub>	1.09	55	長唄 1 <sub>4</sub>	0.00	75	アナウンサー 1	-1.97
16	ソプラノ 1 <sub>4</sub>	1.85	36	清元節 1 <sub>3</sub>	1.06	56	琉球	-0.15	75	アナウンサー 2	-1.97
16	アルト 1 <sub>2</sub>	1.85	37	民謡 1 <sub>4</sub>	1.00	57	琵琶楽 03 <sub>1</sub>	-0.21	75	琵琶楽 03 <sub>s</sub>	-1.97
16	テナー 1 <sub>1</sub>	1.85	38	清元節 1 <sub>4</sub>	0.97	58	一中節	-0.33	78	メゾソプラノ 1 <sub>s</sub>	-2.00
16	アルト 1 <sub>1</sub>	1.85	38	清元節 1 <sub>1</sub>	0.97	59	狂言	-0.58	78	わらべ歌 1 <sub>s</sub>	-2.00
16	テナー 3	1.85	40	地歌 1	0.91	60	狂言	-0.67	78	清元節 1 <sub>s</sub>	-2.00

### 2.1.3 音声データの選択

本研究では、歌声らしさという心理的特徴にはどれくらいの心理的特徴から構成されているのか、MDSを用いて分析を行う。MDSを行うためには、聴取実験から歌声らしさについて各音声データ間の心理的距離を求める必要があり、80個の音声データは実験の際に被験者に与える負担が大きいため、音声データの選択を行った。

表 2.2 に示した音声データの順位から、ほぼ全ての被験者が同じ評価をした上位（歌声らしいと回答したもの）から 3 個、下位（話声らしいと回答したもの）から 3 個、中間の順位に位置するものから 5 個、計 11 個を選択した。表 2.3 にこれらの音声データの詳細を示す。但し、これらの音声データを選択した際には、(1) 邦楽、洋楽などの歌唱法に偏りが無いこと、(2) 多数の発声者を含むこと、(3) 同一発声者の話声と歌声を含むことが考慮された。

表 2.3: 選択した音声データ（ 、 、 の記号が同じものは同一発声者）

資料番号	音声	歌唱法	性別	歌声らしさの順位（評価値）
no.01	歌声	テナー	男性	1 (2.00)
no.02	歌声	バリトン	男性	4 (1.97)
no.03	歌声	メゾソプラノ	女性	11 (1.91)
no.04	歌声	民謡	男性	25 (1.64)
no.05	歌声	わらべ歌	女性	30 (1.36)
no.06	歌声	ソプラノ	女性	33 (1.24)
no.07	歌声	長唄	男性	40 (0.91)
no.08	歌声	声明	男性	51 (0.18)
no.09	話声	-	男性	66 (-1.55)
no.10	話声	-	女性	68 (-1.82)
no.11	話声	-	男性	72 (-1.91)

## 2.2 歌声らしさの心理的空間

### 2.2.1 実験 2.2（歌声らしさ）

#### 実験目的

本研究では歌声らしさという心理的特徴が複数の基本的な心理的特徴から構成されているものであると考え。そこで、歌声らしさという心理的特徴を1次元ではなく多次元で扱うためにMDSを利用する。MDSを行うためには、歌声らしさについての各音声データ間の心理的距離を求める必要があり、聴取実験を行う。

#### 実験方法

実験は、一対比較法を用いて、最初に呈示したものをA、後に呈示したものをBとし、AとBのどちらがどれくらい歌声らしいのかを図2.3に示す7段階の評価尺度で被験者に回答させた。また、呈示した刺激は表2.3に示した11個の音声データであり、刺激対の数は逆の順序も含め合計110個とする。被験者は正常な聴力を有する大学院生8名であり、1人の被験者に対して2回実験を行った。

A の方が 非常に 歌声ら しい	A の方が やや 歌声ら しい	ど ちら とも いえ ない	B の方が やや 歌声ら しい	B の方が 非常 に 歌声ら しい
-2	-1	0	1	2

図 2.3: 評価尺度 (歌声らしさ)

## 実験結果・考察

全被験者の評価値から各音声データ間の心理的距離を求めた。この結果を表 2.4 に示す。表に示した三角行列の値は、各音声データ間の心理的距離を表したものである。表より、no.1、no.2、no.3 の音声データ間には歌声らしさには差があまりない事が分かる。また、no.9、no.10、no.11 の音声データ間についても同様である。一方、no.4 ~ no.8 の音声データ間では心理的距離の値が大きく、歌声らしさに差が見れた。

表 2.4: 各音声データの心理的距離

	1	2	3	4	5	6	7	8	9	10	11
1											
2	3										
3	7	2									
4	40	42	47								
5	53	53	56	31							
6	64	68	66	41	22						
7	58	62	60	45	36	34					
8	68	73	69	46	18	19	19				
9	90	94	93	75	61	48	54	41			
10	92	94	93	75	62	46	50	38	0		
11	93	96	93	75	60	57	56	43	0	0	

## 2.2.2 MDSによる歌声らしさの心理的空間

SPSS 11.0J for Windows により、Kruskal の方法を用いた MDS を行った。図 2.4 は Stress と次元数の関係を示す。Stress とは、ある次元数における各音声データ間の心理的距離の適合度を示すものであり、表 2.5 に Stress の評価について示しておく [5]。表 2.5 より Stress の値は 10% 以下である事が望ましく、図 2.4 から Stress の値は良く適合しているといえる 5.4% である 3 次元解を選択した。そして、作成した歌声らしさの心理的空間を図 2.5 に示す。図中の数字は、表 2.3 の資料番号と対応している。また、図 2.5 において布置された音声データをグループ 1 (番号 1,2,3)、グループ 2 (番号 4,5,6,7,8)、グループ 3 (番号 9,10,11) の 3 つのグループに大別することができる。2.1 節で行った音声データの順位付けの結果を考慮すると、これらのグループは、グループ 1 が歌声らしいもの、グループ 3 が歌声らしくないもの (話しらしいもの)、グループ 2 がそれらの間に位置するものであるといえる。また、図 2.4 の 1 次元解における Stress の値が 17.1% と高いことから、歌声らしさという心理的特徴は 1 次元ではなく多次元で扱う必要がある。このことは、被験者が歌声らしさという心理的特徴を複数の要因によって判断していることを示しており、図

1.2 に示した本研究のモデルを示唆する結果であるといえる。

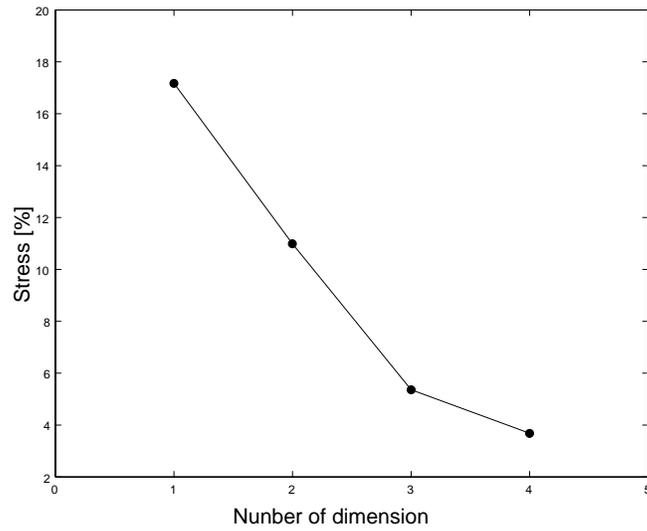


図 2.4: Stress と次元数の関係

表 2.5: Stress の評価

Stress	Stress の度合い
20%	あまりよくない (poor)
10%	まあまあ適合している (fair)
5%	良く適合している (good)
0.5%	非常に適合している (excellent)
0%	完全に適合している (perfect)

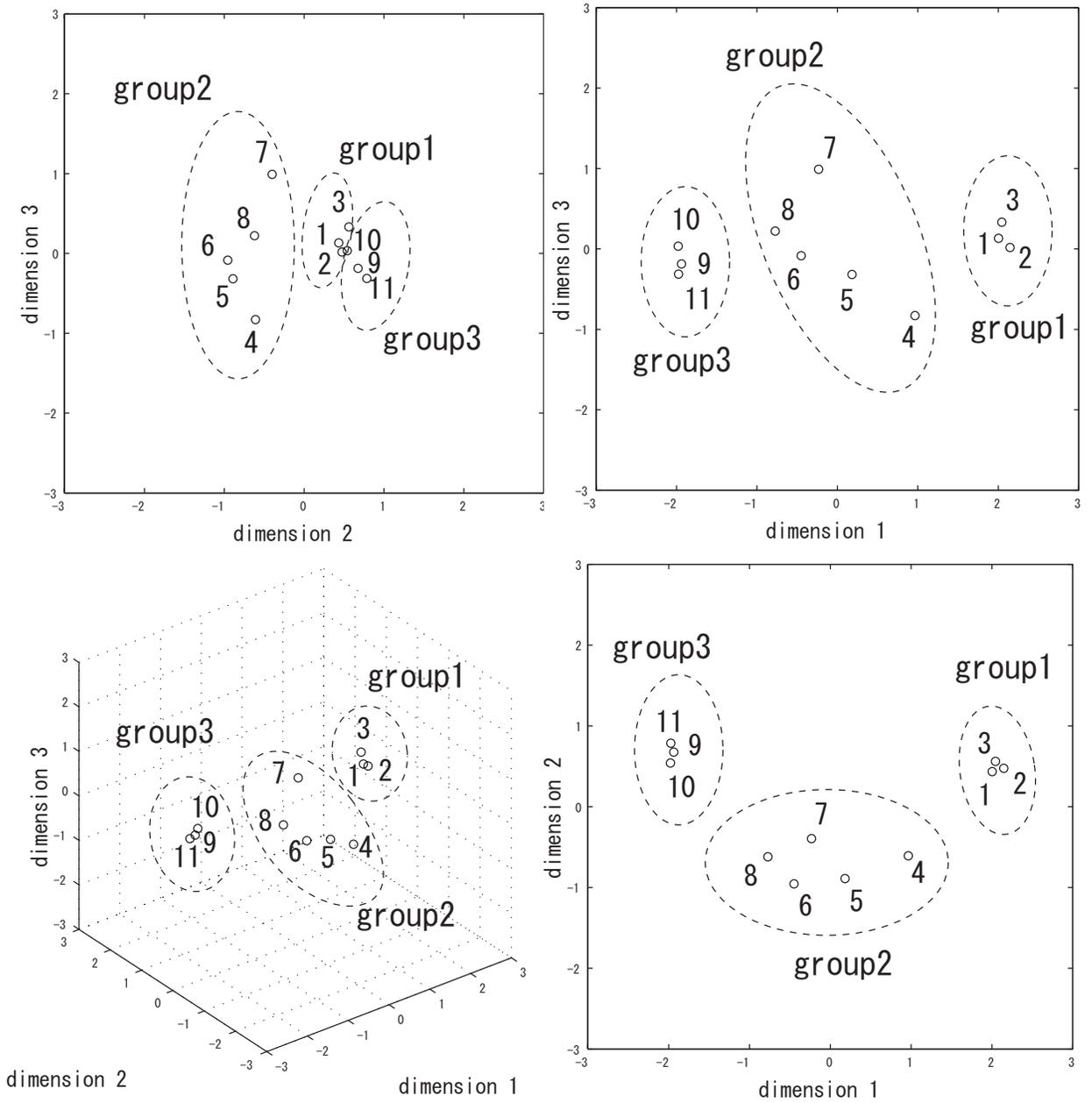


図 2.5: 歌声らしさの心理的空間と各平面から見た音声データの布置

## 2.3 まとめ

本章では、各音声データの歌声らしさの度合いと、歌声らしさの心理的特徴について検討を行った。はじめに、聴取実験から各音声データを歌声らしいものから話声らしいものとなるように順位付けを行い、歌声らしさの度合いについて検討を行った。また、その順位付けの結果から本研究で使用する音声データの選択を行った。

次に、聴取実験とMDSにより歌声らしさの心理的空間を作成した。その結果、音声データの順位付けの結果を考慮すると、歌声らしいもの、歌声らしくないもの（話声らしいもの）、それらの間に位置するものの3つのグループに大別することができた。また、Stressと次元数の関係から、歌声らしさには複数の心理的要因が存在し、少なくとも3つ以上の要因が必要である事が分かった。

## 第3章 歌声らしさの要因の検討

本章では、図 1.2 に示した歌声らしさの3層モデルのうち1層（歌声らしさ）と2層（基本的な心理的特徴）の関係について検討する。はじめに、多数の表現語の中から使用する表現語の検討するために実験 3.1、3.2 を行う。そして、実験 3.3 から各表現語における各音声データの心理的距離を求めることにより、前章で作成した歌声らしさの心理的空間において重回帰分析を行い、歌声らしさの要因を検討する。

### 3.1 表現語の選択

表現語に関する先行研究や実験から、本研究で使用する表現語について検討を行う。

#### 3.1.1 表現語に関する先行研究

表現語に関する先行研究として、上田は日常耳にするあらゆる音に対して使用頻度の高い50個の表現語を挙げている [6]。また、西内・大串らの研究では、歌声を評価する際に良く用いられる表現語として、響き、明瞭さ、音程の正確さ、音色の統一性、好ましいの計5個の表現語が用いられている [3]。本研究で使用する表現語の候補として、これらの先行研究で用いられた表現語を考慮し、本研究で使用する表現語について実験 3.2 より検討を行う。

#### 3.1.2 実験 3.1

##### 実験目的

本研究で行った聴取実験の被験者に対して、歌声の表現語を調査するために実験を行う。

##### 実験方法

前章で行った実験 2.2（歌声らしさ）の後に歌声の表現語について調査を行った。被験者は聴取実験を行った11人である。実験は、呈示した音声データに対して、被験者が歌声らしく感じた心理的特徴を表現語で回答させたものである。

## 実験結果・考察

表 3.1 に被験者によって挙げられた表現語とその人数を示す。実験結果から、被験者は歌声らしさを複数の心理的特徴により判断していることが分かった。また、声の揺れは全ての被験者が回答していることから、歌声らしさに強く影響する表現語であると考えられる。

表 3.1: 被験者が回答した表現語とその人数

表現語	揺れ	響き	長さ	明瞭さ	高さ	張り	裏声
人数	11	10	9	8	7	5	3

### 3.1.3 実験 3.2

#### 実験目的

先行研究において用いられた表現語や実験 3.1 から挙げられた表現語のうち、被験者が表現語の意味を理解できないものや、本研究で使用した 11 個の音声データには当てはまらないものを調査し、本研究で使用する表現語を検討するために聴取実験を行う。

#### 実験方法

実験に使用した表現語について以下に示す。先行研究において、上田が挙げた 50 個の表現語には音声だけではなく日常耳にするあらゆる音を対象としていたため、音声には適切でない表現語が含まれている。また、西内・大串らの先行研究においても、音程の正確さなどは本研究では考慮していない。実験には先行研究で挙げられたそれら以外の表現語と実験 3.1 の結果から 33 個の表現語を選択した。そして、被験者の負担を減らすために、評価尺度を図 3.1 に示す反対の意味を表す表現語は対として両側尺度とし、それ以外は片側尺度とした。その結果、実験には表 3.2、3.3 に示す 21 組 (計 33 個) の表現語を用いた。被験者は前章で行った実験 2.2 (歌声らしさ) を受けた 8 人である。呈示した音声データは表 2.3 に示す 11 個とし、図 2.1 に示した実験システムの構成において防音室内にノート PC を設置する事により、呈示音の聞き直しを可能とした。また、被験者が表現語の意味を理解できない場合は 0 を回答させた。

表 3.2: 両側尺度

明るい - 暗い	軽い - 重い
柔らかい - 固い	鋭い - 滑らか
長い - 短い	豊かな - 貧しい
太い - 細い	澄んだ - 濁った
高い - 低い	綺麗な - 汚い
力強い - 弱い	激しい - 穏やかな

表 3.3: 片側尺度

揺れのある	はっきりした
響きのある	こもっている
張りのある	明瞭さ
深みのある	-
伸びのある	-
迫力のある	-

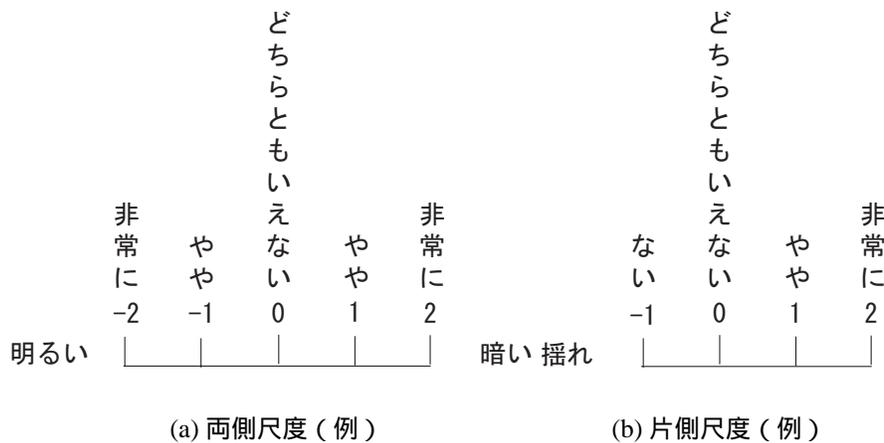


図 3.1: 評価尺度 (表現語)

## 実験結果・考察

実験結果において、多数の被験者が0をつけた表現語は、被験者が表現語の意味を理解できなかったものや、本研究で使用した音声データにはあまり当てはまらない可能性があり、本研究ではこのような表現語は対象外とした。その他にも、被験者の評価値が類似している表現語が見られた。これらの表現語は被験者が同じ意味として評価している可能性があり、代表と考えらる表現語を1つ選択した。その結果、10個の表現語に絞り込むことができた。

次に、この10個の表現語に対して、音声データ毎に被験者の評価値の合計を求めた結果を表3.4に示す。

表 3.4: 各音声データにおける表現語の評価値

	no.1	no.2	no.3	no.4	no.5	no.6	no.7	no.8	no.9	no.10	no.11
明るい(暗い)	7	5	13	8	3	-5	-7	-10	-3	-7	-3
柔らかい(固い)	8	6	4	-5	0	0	-1	-6	-6	-7	-4
太い(細い)	4	8	-4	8	-10	-8	2	8	7	0	4
軽い(重い)	1	0	4	-2	8	3	-1	-7	-1	-4	1
鋭い(滑らか)	-6	-9	8	10	3	-2	-4	-3	4	-1	0
揺れのある	12	13	14	6	-1	-1	10	3	-8	-8	-8
響きのある	8	11	13	7	4	0	2	4	-2	-7	-6
伸びのある	11	13	11	12	0	-4	0	4	-7	-7	-8
長い(短い)	12	9	9	13	8	4	5	6	-13	-13	-12
明瞭さ	9	8	11	13	9	5	-1	-1	4	7	9

表 3.4 の各表現語における評価値を目的変数、図 2.5 の歌声らしさの心理的空間における各音声データの布置を示す 3 次元解を説明変数として重回帰分析を行った。重回帰分析に用いた重回帰式を式 3.1 に示す。 $y$  を目的変数、 $x_1, x_2, x_3$  を説明変数、 $a, b, c$  は偏回帰係数である。式 3.2 における誤差  $e_i$  の二乗和  $E$  が最小となる  $a, b, c$  を求めた。そして、これらの偏回帰係数の値から、各表現語における重相関係数を求めた結果を表 3.5 に示す。但し、重相関係数の値が 1.00 に近いほど、その表現語が歌声らしさを表している事を意味する。

$$y = ax_1 + bx_2 + cx_3 \quad (3.1)$$

$$E = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - ax_{1i} - bx_{2i} - cx_{3i})^2 \quad (3.2)$$

表 3.5: 表現語の重相関係数

表現語	重相関係数
明るい(暗い)	0.92
柔らかい(固い)	0.88
太い(細い)	0.40
軽い(重い)	0.39
鋭い(滑らか)	0.49
揺れのある	0.97
響きのある	0.84
伸びのある	0.93
長い(短い)	0.98
明瞭さ	0.88

表より、声の太さ、軽さ、鋭さの3つは重相関係数の値が低く、歌声らしさとの関係はあまり見られない。以上の結果から本実験で用いる表現語の候補として重相関係数が高い声の明るさ、柔らかさ、揺れ、響き、伸び、長さ、明瞭さの7個が挙げられる。

### 3.1.4 本研究で使用する表現語の選択

実験3.2の結果から明るさ、柔らかさ、揺れ、響き、伸び、長さ、明瞭さの7個の表現語が挙げられた。そして、本研究で作成した歌声らしさの心理的空間は3次元解であることから、少なくとも3つ以上の表現語が必要である。そのため、これら7個の表現語の中で、表3.1の回答数の多い表現語や、西内・大串の先行研究とアンケート結果とに共通する表現語は、歌声らしさに強く関連すると考えられ、声の揺れ、響き、明瞭さの3つの表現語を歌声らしさの要因の候補として選択した。

## 3.2 実験3.3(表現語)

### 実験目的

重回帰分析を用いて各表現語を表す方向を求めるためには、各表現語についての心理的距離が必要である。そこで、各表現語における各音声データ間の心理的距離を求めるために聴取実験を行う。

## 実験方法

実験はシェッフェの一対比較法を用いて、被験者には、各表現語について図 3.2 に示す 5 段階の評価尺度で回答させた。また、呈示した刺激は表 2.3 に示した 11 個の音声データであり、刺激対の数は逆の順序も含め合計 110 個とする。被験者は正常な聴力を有する大学院生 8 名であり、1 人の被験者について 1 回の実験を行った。

A				B
の	A		B	の
方	の	ど	方	方
が	が	ち	が	が
非	が	ら	が	非
常	や	と	や	常
に	や	も	や	に
揺	揺	い	揺	揺
れ	れ	え	れ	れ
て	て	な	て	て
い	い	い	い	い
る	る	る	る	る
-2	-1	0	1	2

図 3.2: 各表現語の評価尺度（揺れの場合）

## 実験結果・考察

実験から得られた全被験者の評価値を浦の変法 [7] によって処理を行い、各表現語における音声データ間の心理的距離を直線上に示した結果を図 3.3 に示す。図において、数値が大きいものほど表現語の心理的特徴が知覚されていることを表す。また、各表現語の音声データの並びを比較すると、各音声データの順序が異っている事から、これらの表現語は異っていることが分かる。このことから、これらの表現語に対応する音響的特徴は異っている可能性が考えられる。

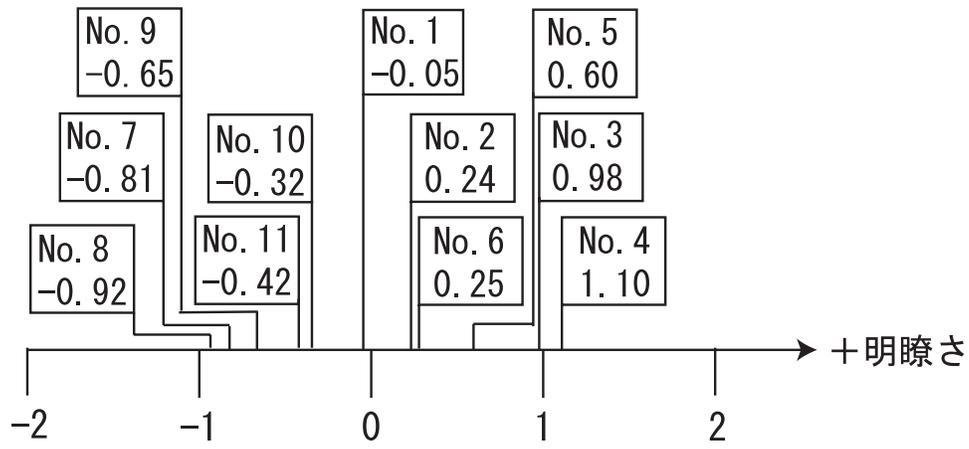
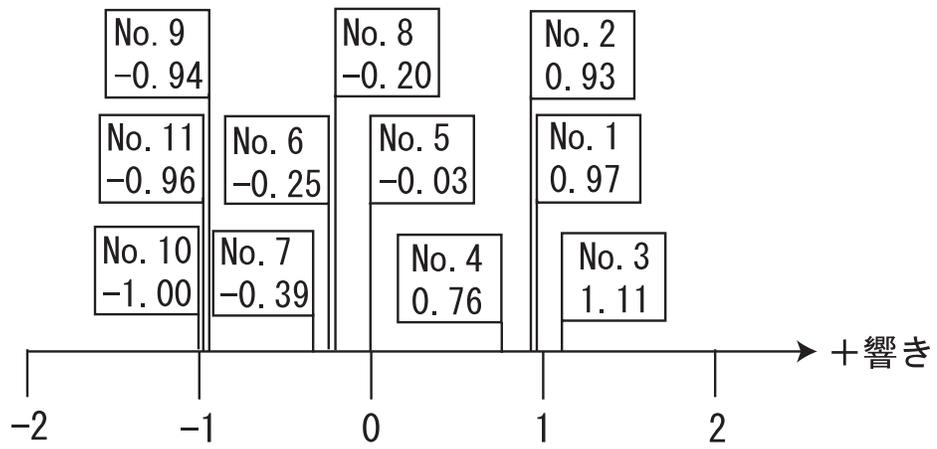
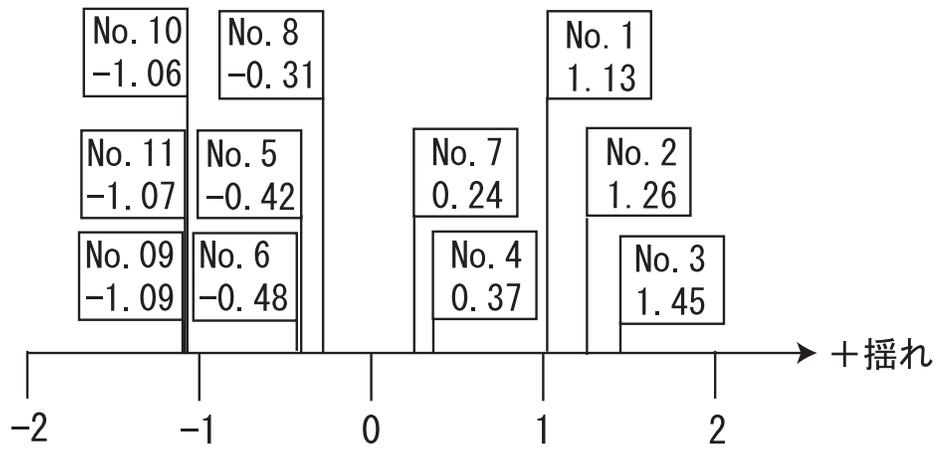


図 3.3: 各表現語の心理的距離

表 3.6: 各表現語における偏回帰係数

	揺れ	響き	明瞭さ
a	0.564	0.501	0.291
b	0.155	0.002	-0.099
c	0.371	-0.201	-0.834

表 3.7: 各表現語を表す方向の関係と重相関係数

表現語		角度 [ ° ]	表現語	重相関係数
揺れ -	響き	56.2	揺れ	0.99
	明瞭さ	105.2	響き	0.99
響き -	明瞭さ	49.3	明瞭さ	0.84

### 3.3 歌声らしさの空間における各表現語の表す方向

図 3.3 の各表現語における心理的距離を目的変数、図 2.5 歌声らしさの心理的空間における各音声データの布置を示す 3 次元解を説明変数として重回帰分析を行った。そして、重回帰分析から得られた偏回帰係数  $a, b, c$  を表 3.6 に示す。表 3.6 に示した偏回帰係数の値から、3 次元空間上における各表現語を表す方向と重相関係数を求めた結果を表 3.7 に示す。但し、重相関係数の値が 1.00 に近いほど、その表現語が歌声らしさを表している事を意味する。表 3.7 より、3 次元空間上における 3 つの表現語の重相関係数の値が高い事や表現語を表す方向がそれぞれ大きく異なっている事から、声の揺れ、響き、明瞭さの表現語は、歌声らしさの要因であると考えられる。

次に、図 2.5 に示した歌声らしさの心理的空間において、1-2 平面、2-3 平面、1-3 平面の各平面から見た場合における音声データの布置と各表現語の方向を図 3.3 に示す。重相関係数の高い表現語の方向についてのみ考察すると、2 節で示した 3 つのグループは、声の揺れや響きの違いを表していると考えられる。

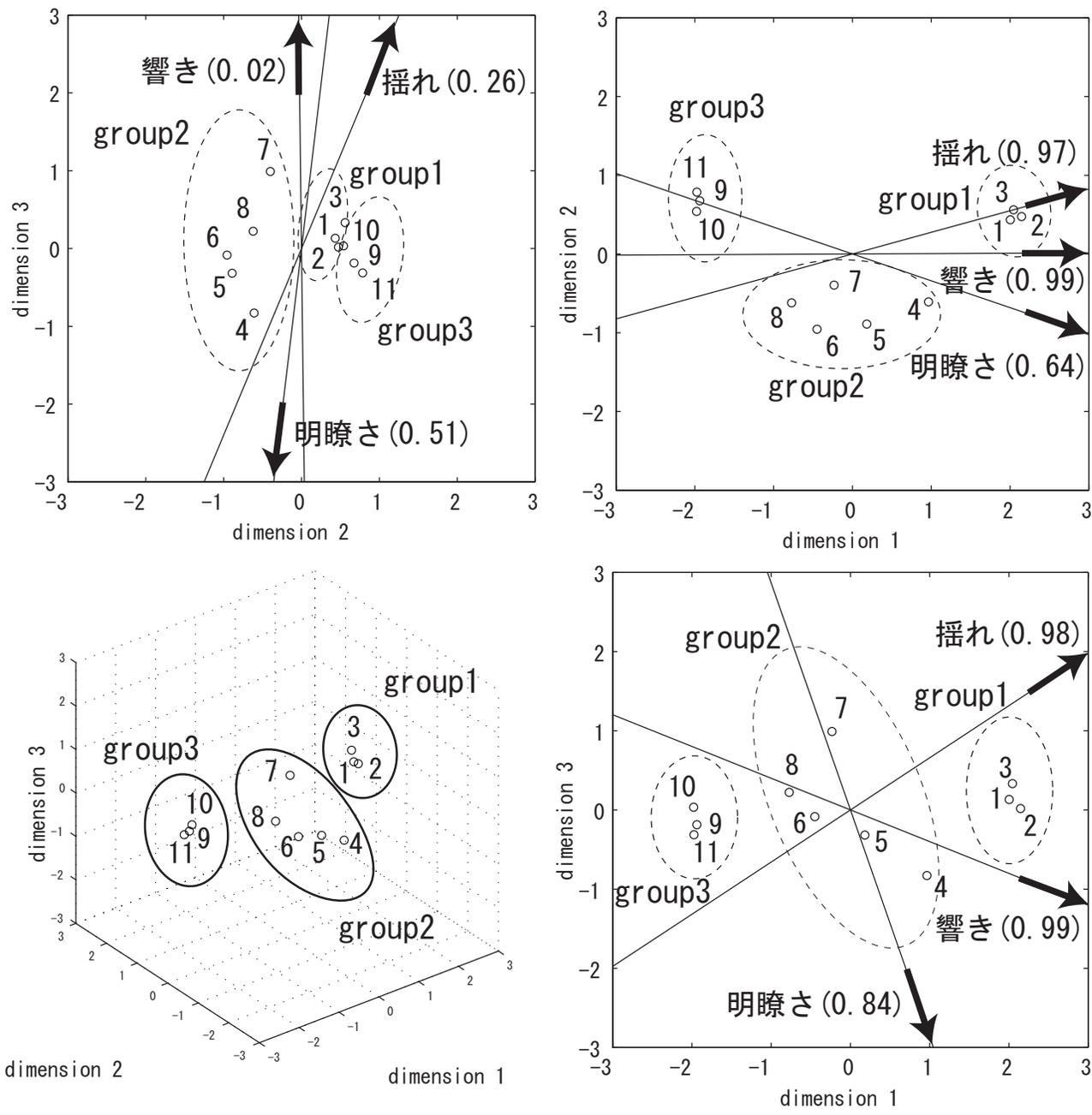


図 3.4: 各表現語を表す方向

### 3.4 まとめ

本章では、歌声らしさの要因について検討を行った。実験 3.1、3.2 より、本研究で使用する表現語の検討を行った。その結果、声の明るさ、柔らかさ、揺れ、響き、伸び、長さ、明瞭さの 7 個の表現語が挙げられた。そして、それらの表現語の中から声の揺れ、響き、明瞭さを選択し、歌声らしさの心理的空間において重回帰分析を行い、各表現語を表す方向の関係について検討を行った。その結果、これら 3 つの表現語は歌声らしさの要因である事が分かった。以上の結果より、図 1.2 に示した歌声らしさの 3 層モデルのうち第 1 層目（歌声らしさ）と第 2 層目（基本的な心理的特徴）の関係が明らかになった。また、本研究では、7 個の表現語の中から 3 個を選択し検討を行ったが、他の 4 個の表現語についても検討する必要がある。

# 第4章 歌声らしさの要因に関連する音響的特徴の抽出

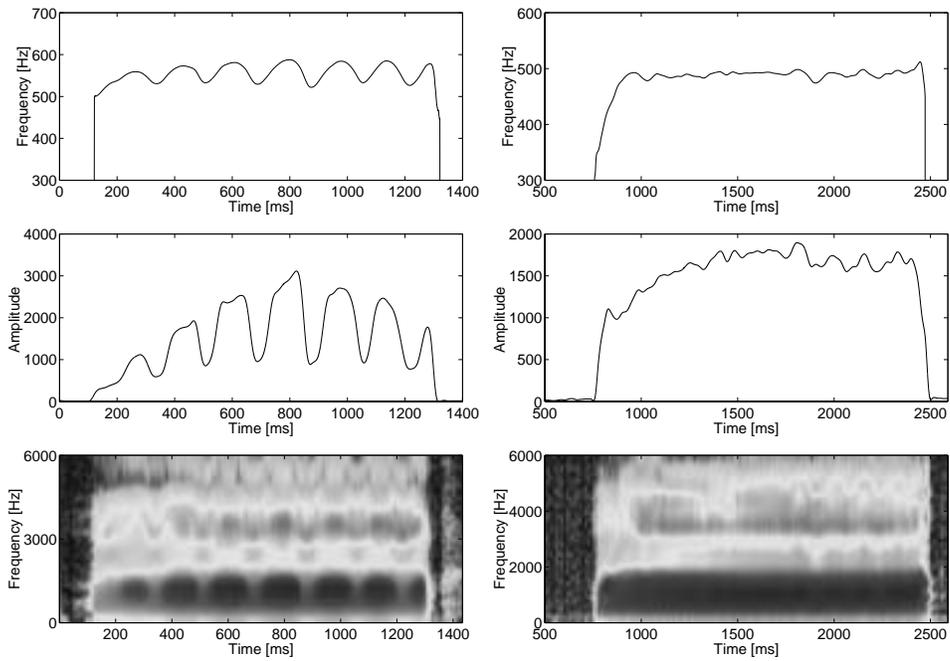
本章では、図 1.2 に示した歌声らしさの3層モデルのうち第2層目（基本的な心理的特徴）と第3層目（物理的特徴）の関係について検討する。前章では、声の揺れ、響き、明瞭さは歌声らしさの要因であることが分かった。これらの心理的特徴のうち、歌声らしさに強く関連すると考えられる声の揺れや響きに関連する音響的特徴を分析から抽出する。そして、それらの音響的特徴と歌声らしさの要因である心理的特徴との関係について合成音により検証を行う。

## 4.1 揺れに関連する音響的特徴の分析

### 4.1.1 揺れに関連する音響的特徴の調査

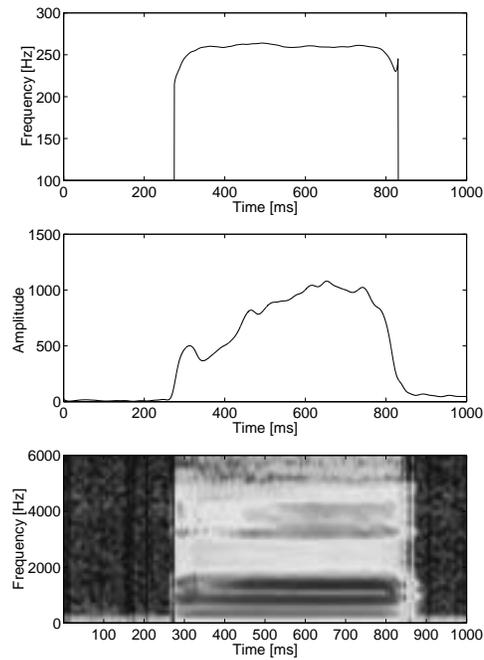
歌声らしさの要因の一つである揺れに関連する音響的特徴について調査を行う。声の揺れという心理的特徴に対応する音響的特徴は、時間的に変動するものと考えられる。そこで、揺れに関する聴取実験により求めた揺れの順位から、順位の1位（no.3）、7位（no.5）、9位（no.10）の音声データの振幅エンベロープと STRAIGHT[5] により基本周波数（F0）およびスペクトルについて分析した結果を図 4.1 に示す。分析には 12 kHz にダウンサンプリングされた音声データを使用し、分析条件は、FFT length 1024 point、frame rate 1 ms とした。

図 4.1 より、F0 および振幅エンベロープには周期的な変動や瞬時的な変動が見られ、特に揺れの順位が高い音声データには周期的な変動が見られる。また、スペクトルにおいても、フォルマントが周波数方向や振幅方向に変動している事が分かる。そこで、これらの変動に着目し、揺れに関連する音響時特徴について分析を行う。また、周期的に変動している区間が同じである事から、これらの音響的特徴は互い影響し合っている可能性が考えられる。そのため、音響的特徴同士の関係についても検討を行う。



(a) no.3 (1 位)

(b) no.5 (7 位)



(c) no.10 (9 位)

図 4.1: F0 (上), 振幅エンベロープ (中), スペクトル (下) の分析結果

表 4.1: F0 と振幅エンベロープにおける変調周波数と偏移幅の分析結果

資料番号	揺れ (心理量)	F0 変調周波数 [Hz]	振幅エンベロープ 変調周波数 [Hz]	F0 偏移幅 [dB]	振幅エンベロープ 偏移幅 [dB]
no.3	1.45	5.4	5.5	$1.6 \times 10^{-1}$	1.7
no.2	1.26	5.1	5.1	$2.1 \times 10^{-1}$	1.2
no.1	1.13	4.5	4.6	$9.2 \times 10^{-2}$	$3.3 \times 10^{-1}$
no.4	0.37	4.4	4.2	$6.9 \times 10^{-2}$	$1.4 \times 10^{-1}$
no.7	0.24	4.1	4.0	$1.3 \times 10^{-2}$	$5.2 \times 10^{-1}$
no.8	-0.31	4.6	4.5	$3.8 \times 10^{-2}$	$3.2 \times 10^{-1}$
no.5	-0.42	5.7	1.1	$4.3 \times 10^{-2}$	$1.4 \times 10^{-1}$
no.6	-0.48	5.6	0.5	$2.1 \times 10^{-2}$	$1.2 \times 10^{-1}$
no.10	-1.06	11.1	5.0	$1.2 \times 10^{-3}$	$7.4 \times 10^{-2}$
no.11	-1.07	7.3	6.8	$5.7 \times 10^{-3}$	$1.6 \times 10^{-1}$
no.9	-1.09	9.6	7.5	$2.6 \times 10^{-3}$	$1.1 \times 10^{-2}$
相関値	-	-	-	0.91	0.79

#### 4.1.2 F0 と振幅エンベロープの変動

図 4.1 において、揺れの順位が高い音声データの F0 と振幅エンベロープには周期的な変動が見られた。そこで、その周期的な変動に着目し、これらの変動の平均値からの変位について周波数成分の分析を行った。分析範囲は音声の定常部とし、no.1 ~ no.8 は 1000 ms とし、no.9、no.10、no.11 は 100、150、200 ms とした。はじめに、各音声データの支配的な変調周波数を求めた。その結果を表 4.1 に示す。但し、表に示した音声データは心理量の揺れの順番に並べてある。表より、揺れの順位が高いものは 4 ~ 6 Hz に支配的な変調周波数を持っている事が分かった。また、齋藤らの報告 [2] によると、これらはヴィブラートに関連するものであると考えられる。

次に、4 ~ 6 Hz の周波数帯域における偏移幅 (振幅スペクトルの和) について求めた結果を表 4.1 に示す。表より、F0 と振幅エンベロープの 4 ~ 6 Hz の偏移幅は、心理量の揺れに関連する並びに類似していることが分かる。そこで、4 ~ 6 Hz の偏移幅と実験 3.3 (表現語) より求めた揺れの間隔尺度の値との相関値を求め、表 4.1 に示す。その結果、4 ~ 6 Hz の偏移幅と心理量との相関値は、F0 で 0.91、振幅エンベロープで 0.79 となり、両者

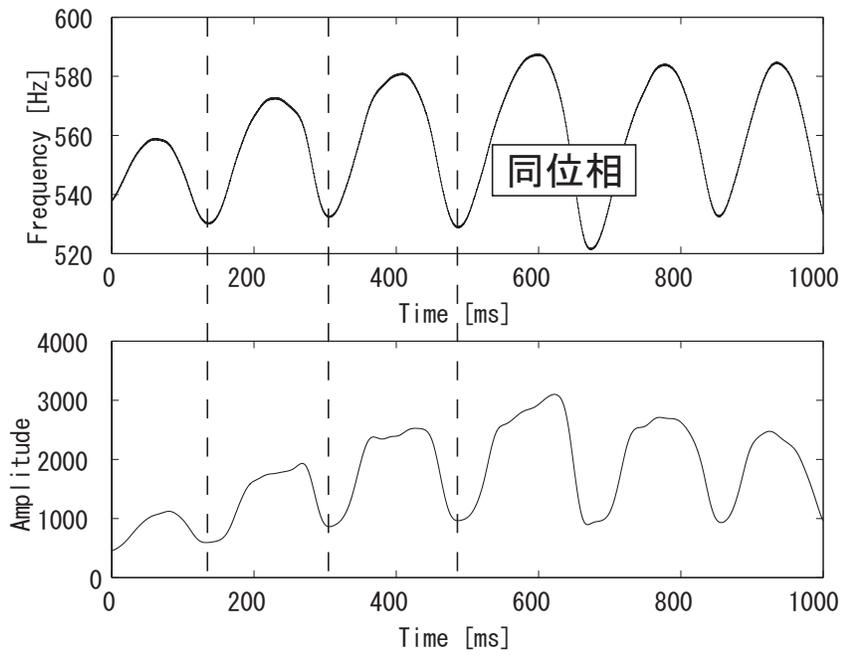
には相関関係があった。また、表から no.7 は心理量との相関値を下げている事が分かる。そこで、各音声データの F0 の時間変化と振幅エンベロープとの相関値を求めた結果を表 4.2 に示す。また、揺れの順位が最も高い no.3 と no.7 の F0 の時間変化と振幅エンベロープを図 4.2 に示す。

その結果、揺れ（心理量）との相関値を下げている no.7 は、F0 と振幅エンベロープとの相関値において-0.86 と負の相関を示し、これらの関係が逆位相である事が分かった。一方、揺れの順位が高いものでは、これらの関係は正の相関を示し同位相であった。そのため、no.7 は他の音声データに比べ F0 や振幅エンベロープにおける 4～6 Hz の偏移幅が大きいにも関わらず、揺れの順位が低くなった可能性が考えられる。

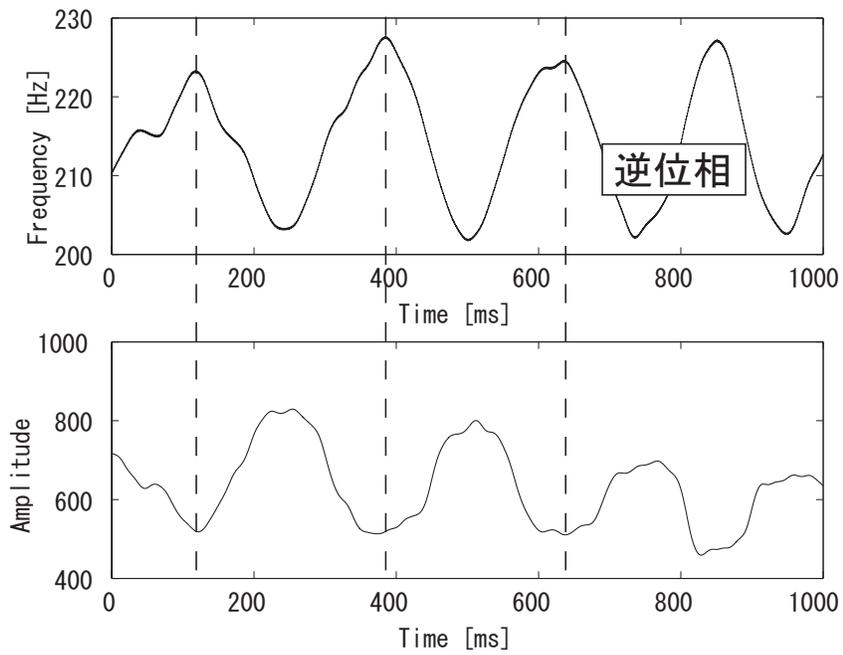
以上の結果より、F0 と振幅エンベロープの 4～6 Hz の変位と、さらにそれらが同位相であることは、揺れに関連する音響的特徴であると考えられる。

表 4.2: F0 の時間的变化と振幅エンベロープの相関値

資料番号	相関値
no.3	0.87
no.2	0.96
no.1	0.82
no.4	0.41
no.7	-0.86
no.8	0.92
no.5	0.76
no.6	0.71
no.10	-0.54
no.11	0.93
no.9	0.35



(a) no.3

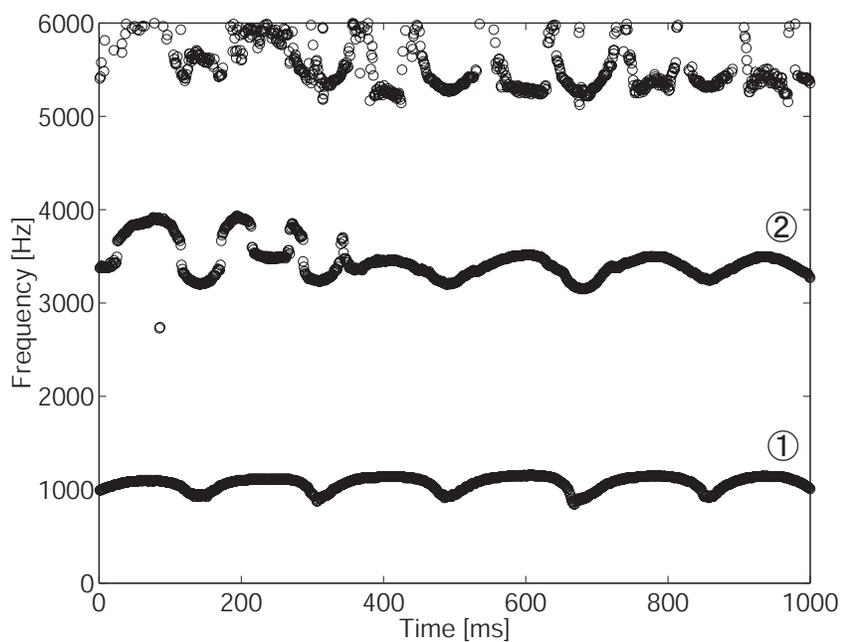


(b) no.7

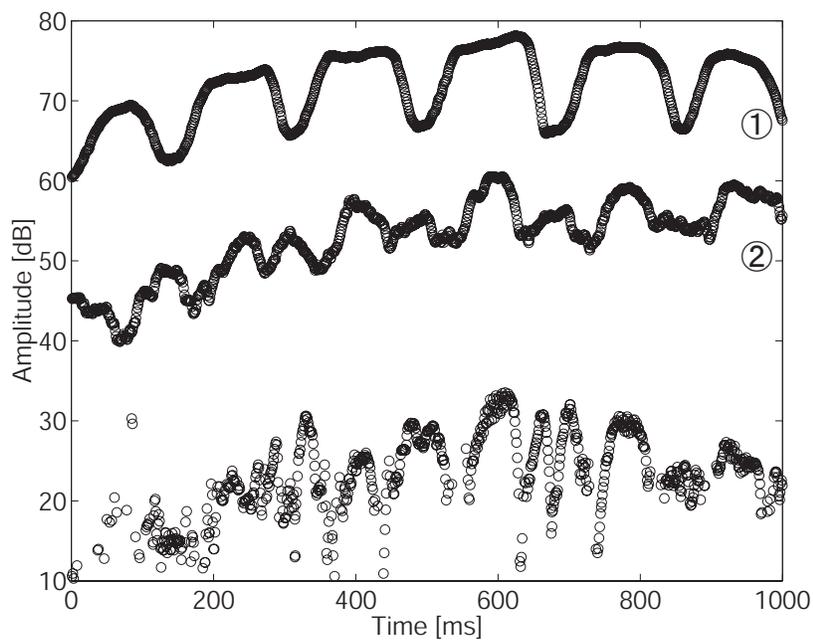
図 4.2: F0 の時間変化と振幅エンベロープの関係

### 4.1.3 フォルマントの周波数方向や振幅方向への変動

4.1.1 節において、分析からフォルマントが周波数方向や振幅方向に変動している事が分かった。そこで、揺れの順位の1位 (no.3)、7位 (no.5)、9位 (no.10) のフォルマントの周波数方向や振幅方向の変動について分析を行った。その結果を、図 4.3 ~ 4.5 に示す。分析範囲は音声の定常部 (no.1 と no.5 は 1000 ms、no.10 は 200 ms) とした。また、図の番号はピークの対応関係を示す。図より、揺れの順位の高い音声データには、周波数方向や振幅方向に周期的な変動がある事が分かった。

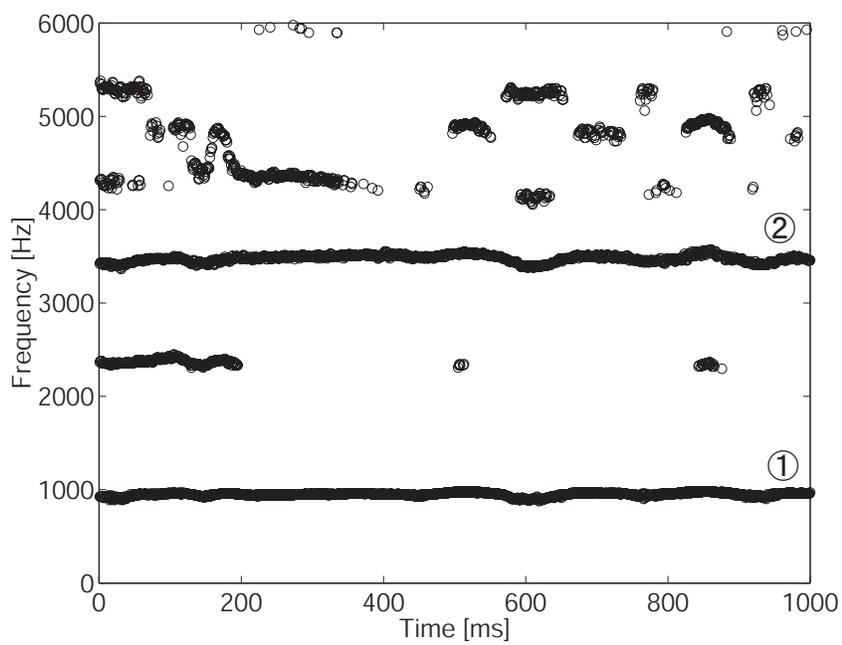


(a) 周波数方向

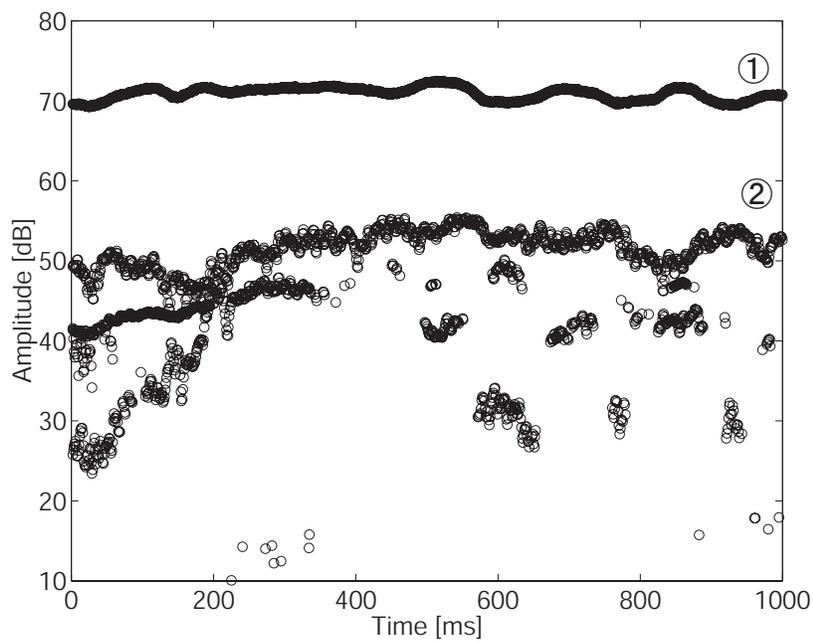


(b) 振幅方向

図 4.3: no.03 (1位) のフォルマントの変動

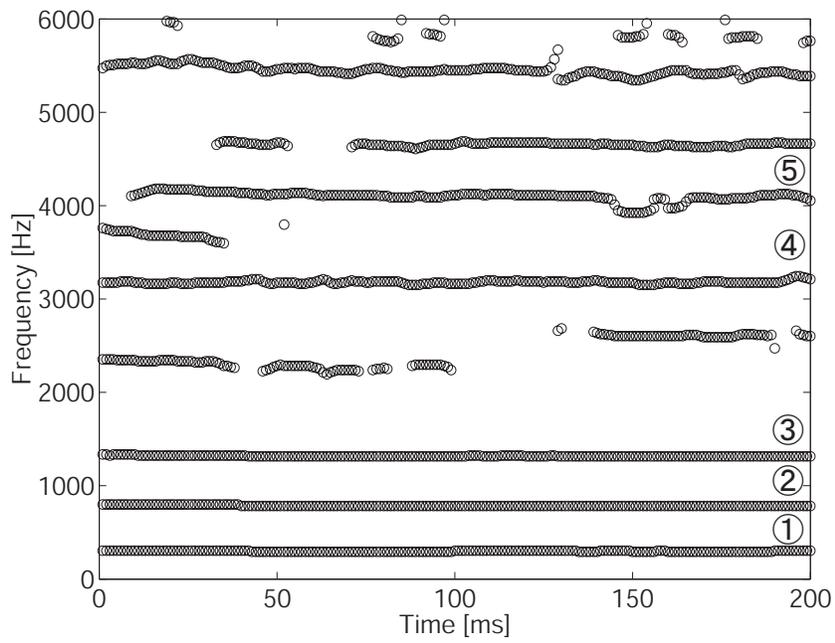


(a) 周波数方向

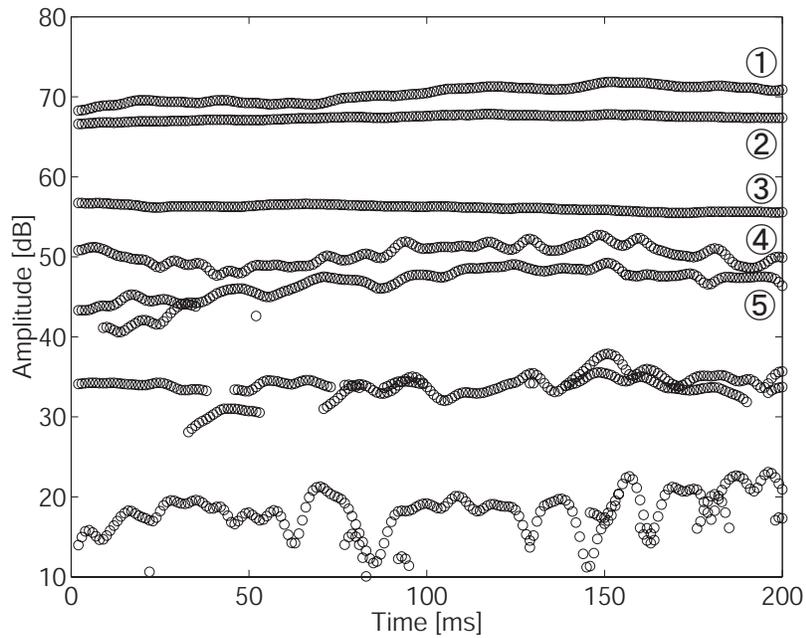


(b) 振幅方向

図 4.4: no.5 (7位) のフォルマントの変動



(a) 周波数方向

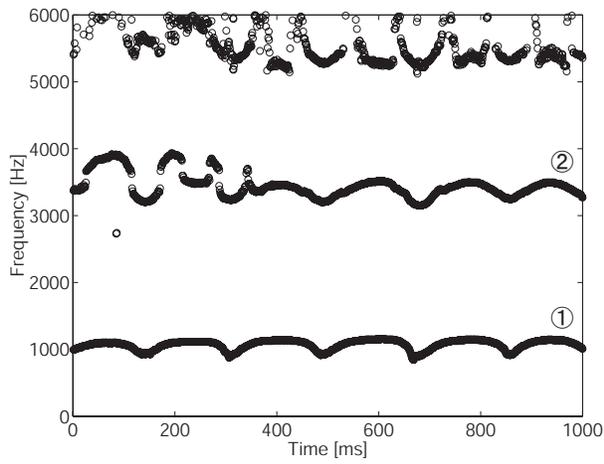


(b) 振幅方向

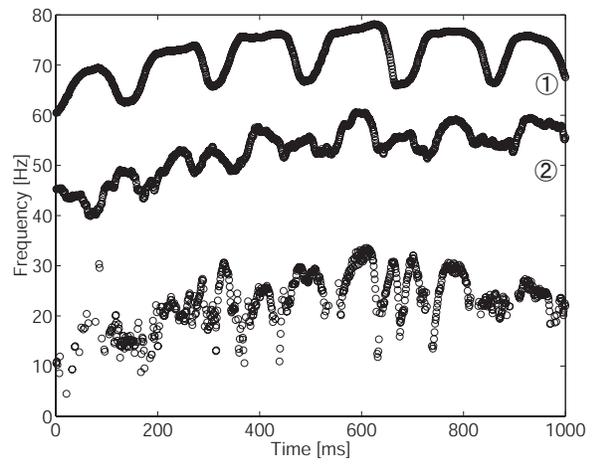
図 4.5: no.10 (9位) のフォルマントの変動

次に、F0 とフォルマントの周波数方向や振幅方向の変動の関係について検討する。図 4.6 に揺れの順位が最も高い no.3 のスペクトルとフォルマントの周波数方向や振幅方向の時間的变化および F0 を示す。

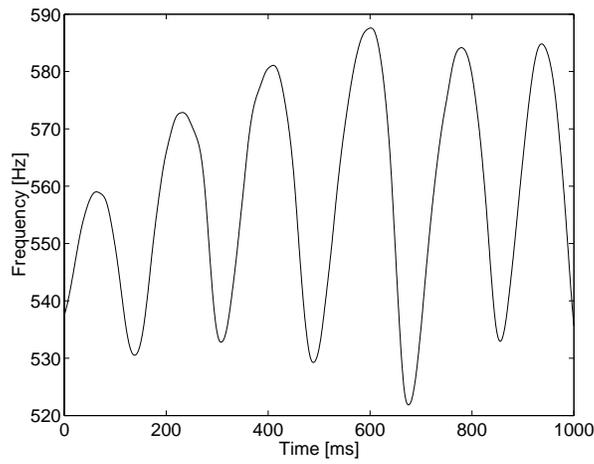
その結果、フォルマントは F0 の変動と同位相で周波数変調や振幅変調している事が分かった。また、このような特徴は揺れの順位の高いものにおいて、F0 が変動している部分に顕著に見られた。以上の結果より、フォルマントの周波数変調や振幅変調は揺れに関連する音響的特徴であり、それらの特徴は F0 と同位相で変動している事が分かった。



(a) 周波数方向



(b) 振幅方向



(c) F0

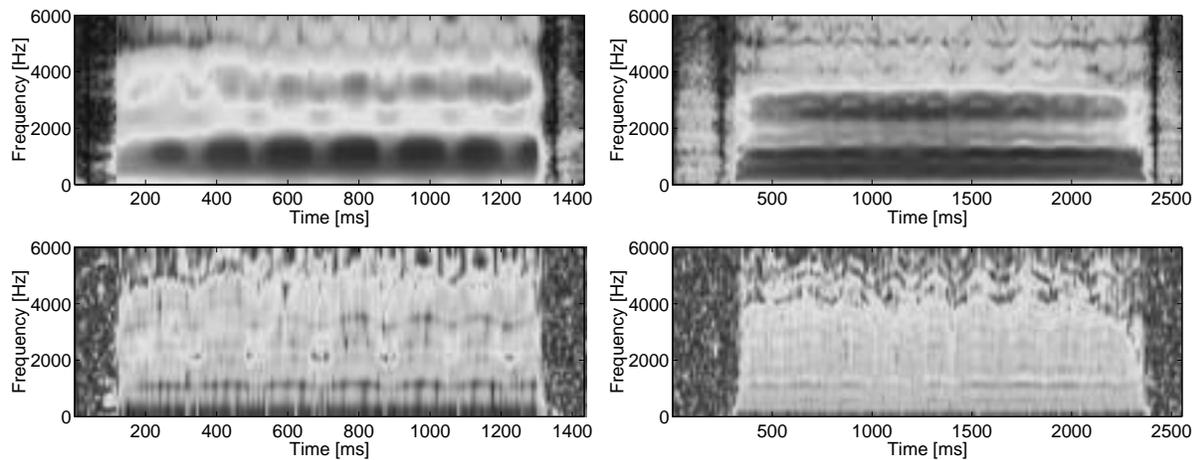
図 4.6: F0 とフォルマントの周波数変調や振幅変調の関係

## 4.2 響きに関連する音響的特徴の分析

### 4.2.1 響きに関連する音響的特徴の調査

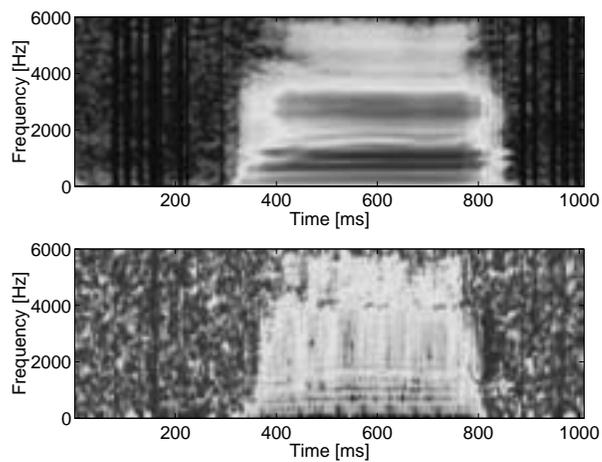
歌声らしさの要因の一つである響きに関連する音響的特徴について調査を行う。Sundberg は、洋楽的唱法において、響きの良い歌声にはスペクトル包絡の 3 kHz 付近にピークがあり、歌声特有の音響的特徴である事を指摘している [1]。また、中山・小林は、洋楽的唱法だけでなく邦楽的唱法においても、3kHz 又は 4 kHz 付近に顕著なピークが見られ、歌声の声質にはそのようなピークが重要である事を指摘している [9]。そこで、スペクトルの 3kHz、又は 4 kHz 付近の周波数帯域に着目し、STRAIGHT により、響きに関する聴取実験により求めた響きの順位から、順位の 1 位 (no.3)、2 位 (no.1)、9 位 (no.9) の音声データのスペクトルと非周期成分の割合 (Noise ratio) について分析を行った結果を図 4.7 に示す。非周期成分の割合は、上側包絡が周期成分、下側包絡が非周期成分を表すことから、これらの差より求まる。また、STRAIGHT は、各周波数毎に非周期成分の割合を求めることが可能であることから、高調波のある周波数では、非周期成分の割合は小さくなる。分析には 12 kHz にダウンサンプリングされた音声データを使用し、分析条件は、FFTLenght1024 point、framerate 1 ms とした。

その結果、響きの順位の高い no.1 の音声データは、響きの順位の低い no.9 と比べて 3 kHz 付近に顕著なピークがある事が分かった。また、響きの順位が最も高い no.3 の音声データには、このようなピークは見られないが、3 kHz 付近の非周期成分に大きなディップが見られ、上側包絡と下側包絡の差が顕著であり、ここには強い高調波の成分があると考えられる。そして、響きの順位の低い no.9 の非周期成分の割合には、3 kHz 付近にこのようなディップは見られない。以上より、この 2 つの音響的特徴に着目して分析を行った。



(a) no.3 (1位)

(b) no.5 (7位)



(c) no.10 (9位)

図 4.7: スペクトル (上) 非周期成分の割合 (下) の分析結果

## 4.2.2 スペクトル包絡における3 kHz 付近のピーク

4.2.1 節において、響きの順位の高い音声データには、3 kHz 付近に顕著なピークがある事が分かった。そこで、STRAIGHT により、各音声データのスペクトルについて分析を行った。

そして、響きの順位の2位 (no.1)、4位 (no.4) 9位 (no.9) の音声データの定常部における時間平均したスペクトル包絡を図 4.8 に示す。但し、no.1 と no.9 は同一発声者である。

その結果、響きの順位の低い no.9 に比べ、no.1、no.4 は3 ~ 4 kHz 付近に顕著なピークがあることが分かる。そして、同一発声者ではピークの位置がほぼ等しい事や、中山・小林により、歌唱法の違いによりピークの位置に違いがある事が指摘されていることから [9]、ピークが表れる周波数の違いは、発声者や歌唱法の違いによる影響によるものと考えられる。以上より、スペクトル包絡における3 ~ 4 kHz の、ピーク値が高い音声データは、響きの上位に位置している事から、この特徴は響きに関連する音響的特徴である可能性が考えられる。しかし、11 個の音声データ全てはスペクトル包絡における3 ~ 4 kHz のピークの高さで並べる事ができなかったが、ピーク値の高いものが上位にくる傾向は見られた。

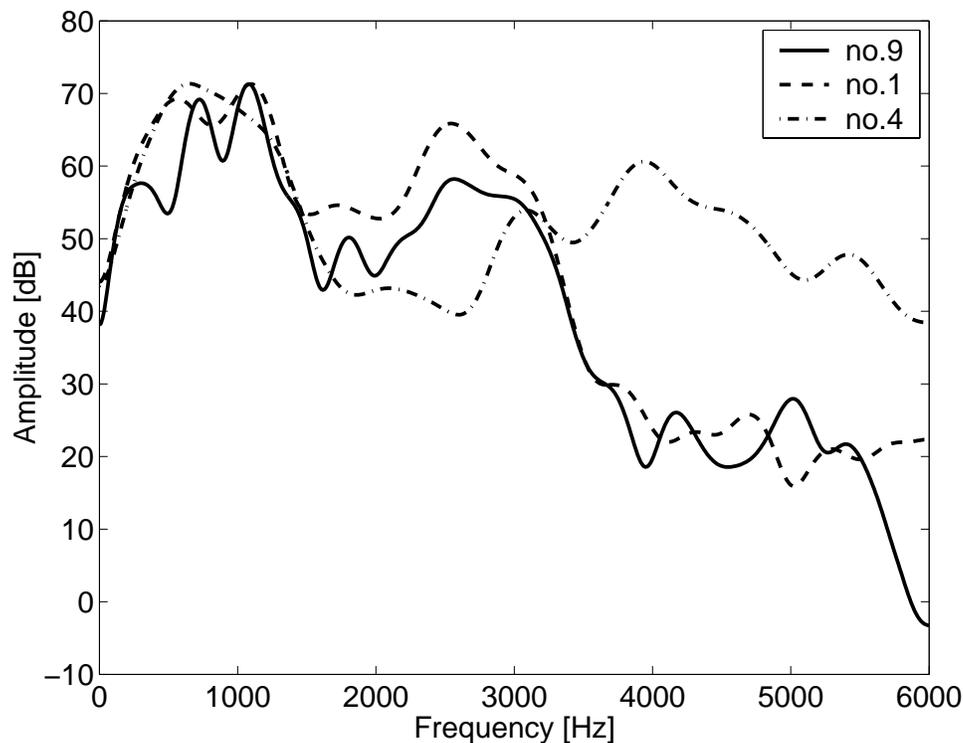


図 4.8: 時間平均したスペクトル包絡の比較

### 4.2.3 3 kHz 付近の強い高調波成分

4.2.1 節において、no.3 には、3 kHz 付近の非周期成分の割合に大きなディップが見られ、上側包絡と下側包絡の差が大きいことから、ここには強い高調波の成分があると考えられる。そこで、STRAIGHT により、響きの順位の高い no.3 (1 位) および no.1 (2 位) と、響きの順位が低い no.9 (9 位) の高調波成分について検討を行うために非周期成分の割合を分析した。また、強い高調波成分によるスペクトル包絡へ影響も検討するためにスペクトルについても分析を行った。図 4.9 にこれらの音声データの非周期成分の割合、図 4.10 に瞬時のスペクトル包絡や非周期成分の割合を示す。

その結果、no.3 の音声データには、3 kHz 付近の非周期成分の割合に顕著なディップが見られ、ここには強い高調波の成分があることが分かる。そして、その高調波とスペクトル包絡のピークが一致している事が分かる。そこで、これらの音声データの 3 kHz 付近の高調波とスペクトル包絡のピークとの時間変化について分析を行った結果を図 4.11 に示す。図において、線は高調波の時間変化を示し丸はスペクトル包絡のピークの時間変化を示す。その結果、no.3 の女性の音声データにおいて、高調波と 3 kHz 付近のピークが重なり変化している事が分かった。以上より、3 kHz 付近のピークと高調波の重なりによる強い高調波成分は、響きに関連する音響的特徴である可能性が考えられる。

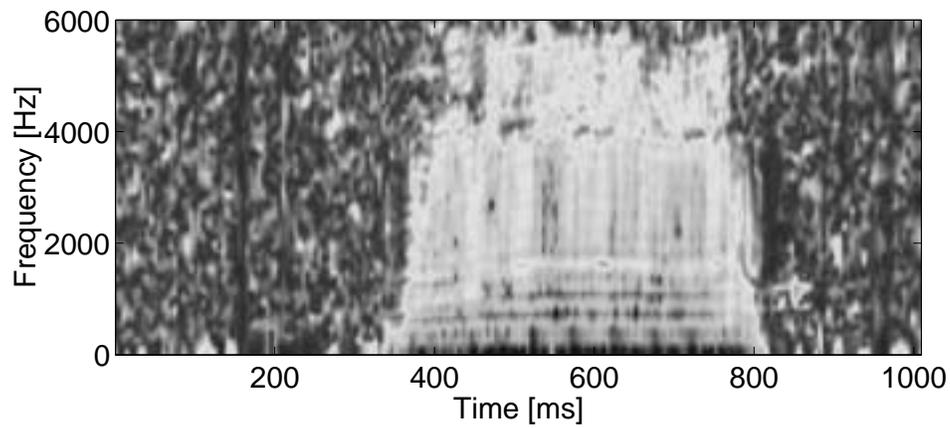
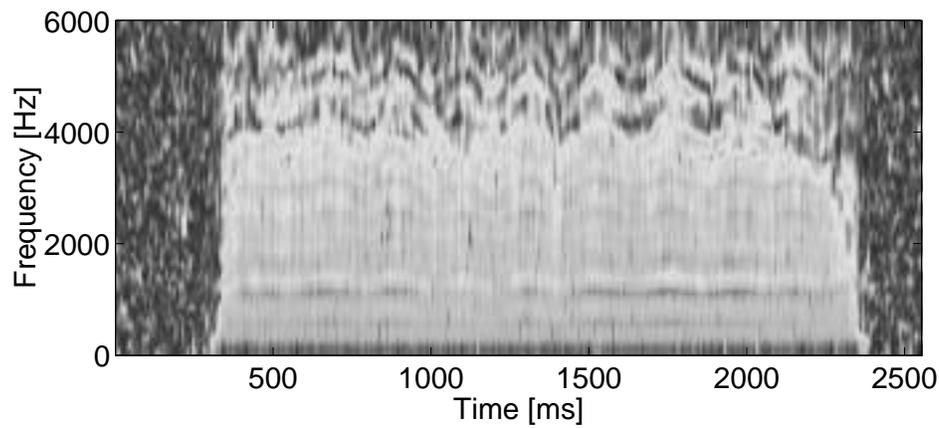
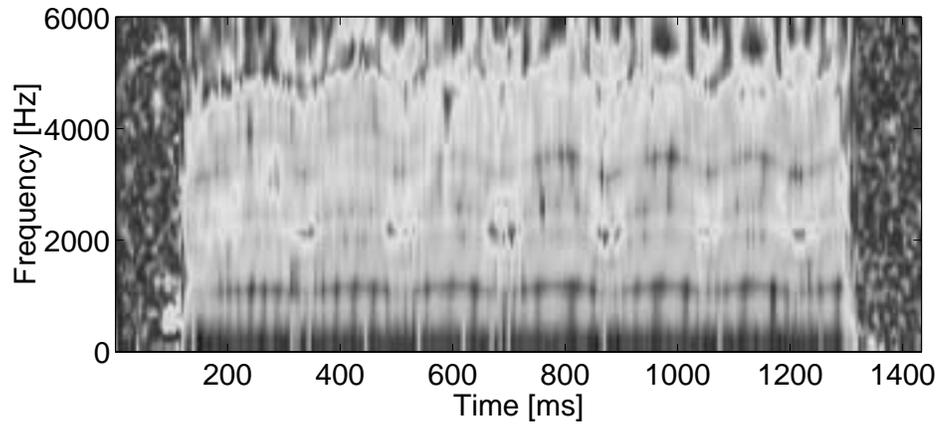


図 4.9: no.3(上),no.1(中),no.9(下) の非周期成分の割合の時間変化

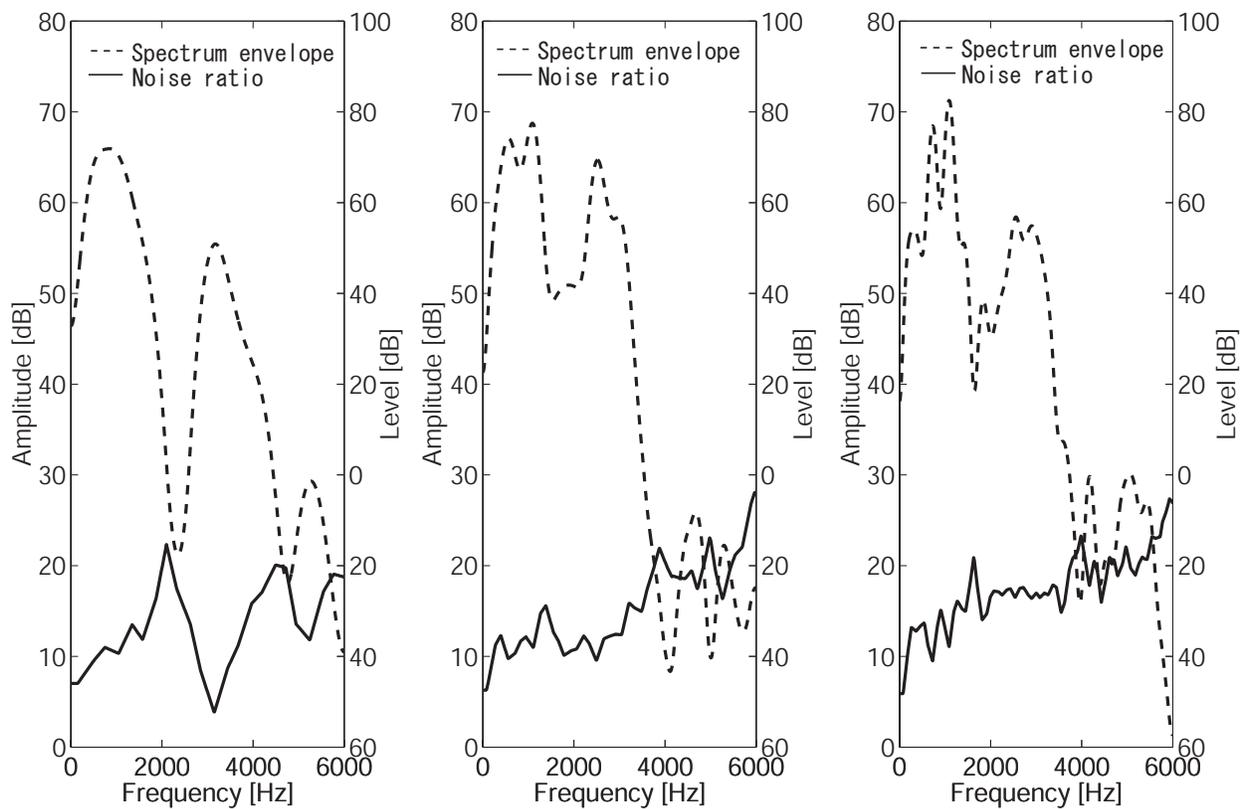
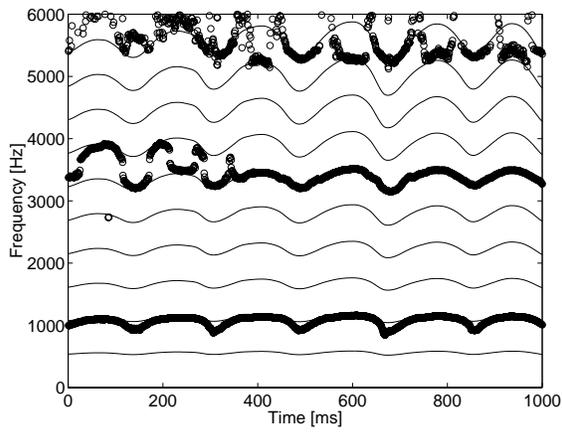
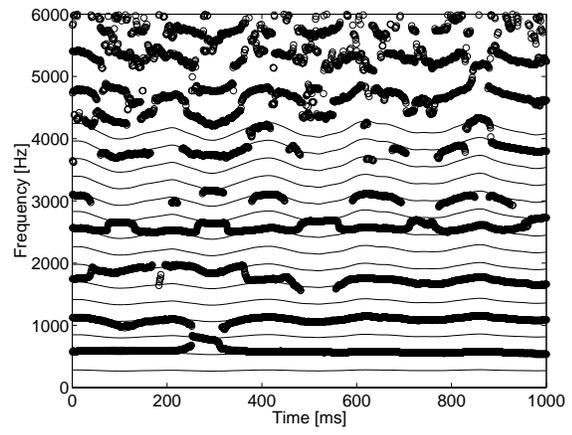


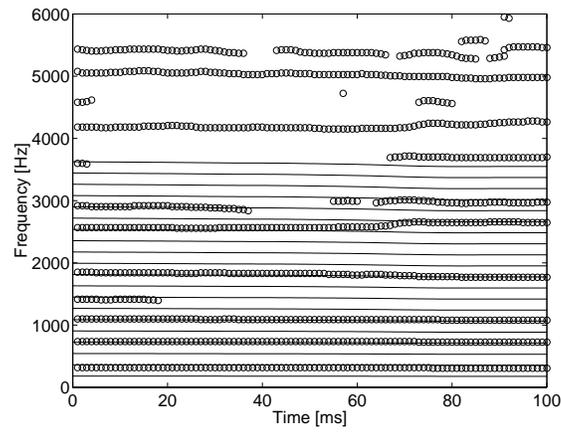
図 4.10: 高調波成分と 3 kHz 付近のピークの関係



(a) no.3



(b) no.1



(c) no.9

図 4.11: 高調波の時間変化と 3 kHz 付近のピークの時間変化

### 4.3 STRAIGHT を用いた合成法の概要

前節までは歌声らしさの要因である揺れと響きに関連する音響的特徴の分析を行った。そこで、それらの音響的特徴を付加した合成音を作成し聴取実験を行う事により、それらの音響的特徴が揺れ、又は響きに関連があるのか検証を行う。そのため、本研究では STRAIGHT を用いて合成音を作成する。そこで、図 4.12 に STRAIGHT を用いた合成法の概要を示す。合成音を作成する際の元とする音声データは、STRAIGHT の分析系により、F0、スペクトル包絡、非周期成分の割合に分解される。そして、それぞれの制御部により、本研究の分析から得られた音響的特徴を付加した後、STRAIGHT の合成系に入力され合成音が作成される。

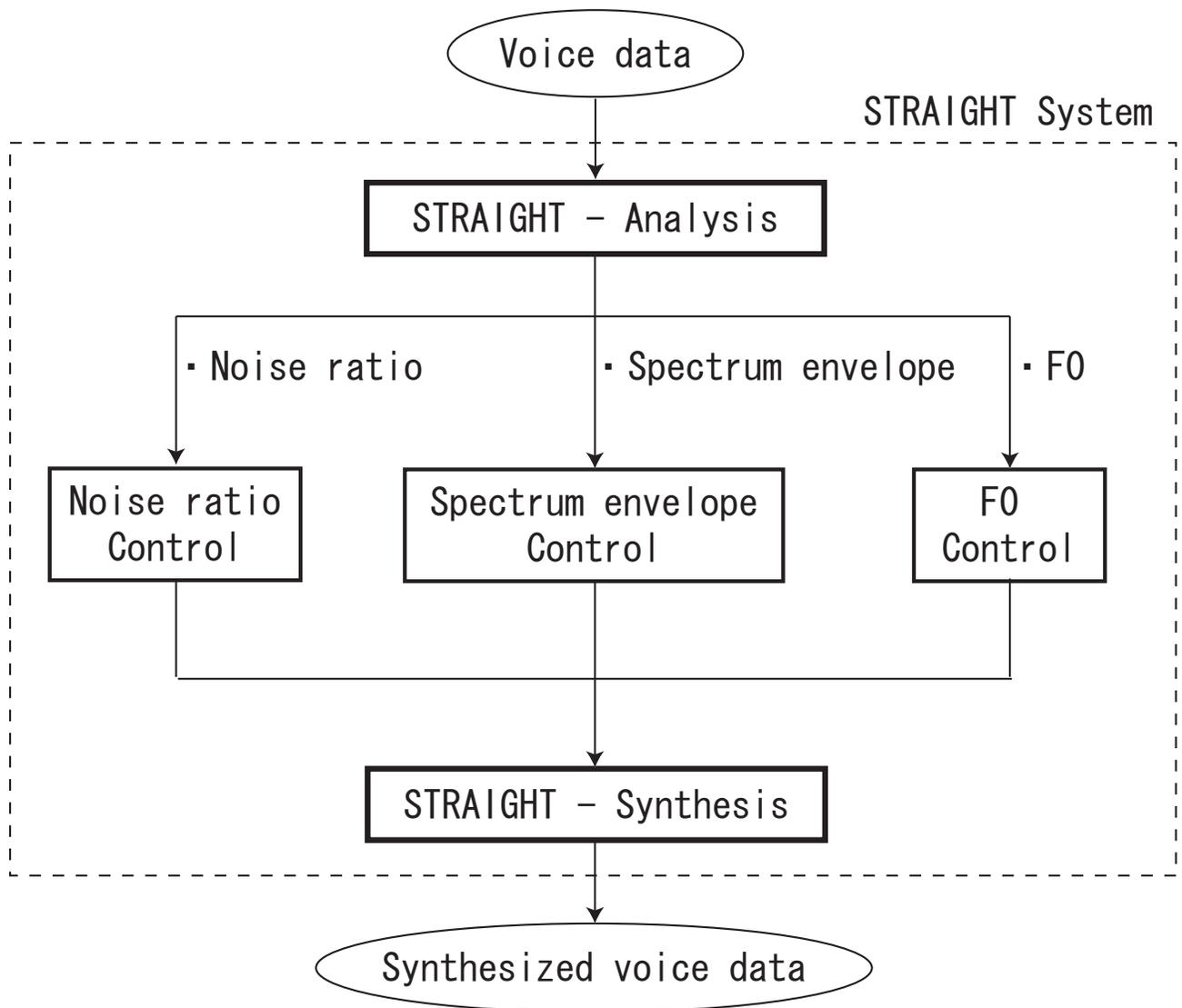
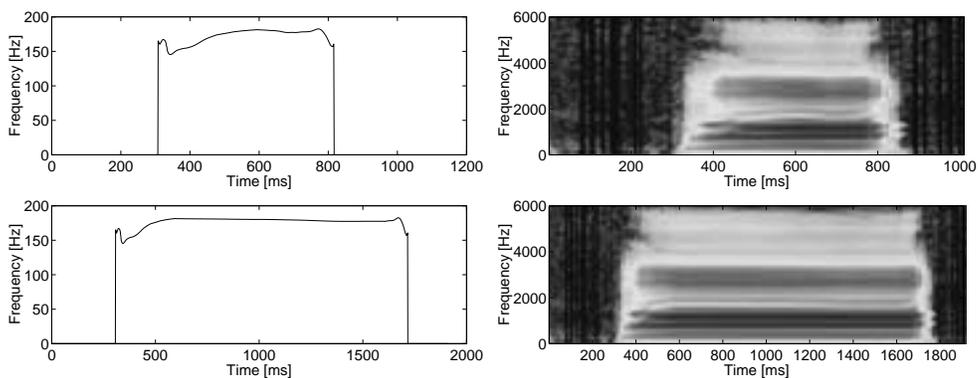


図 4.12: STRAIGHT のを用いた合成手法

## SYN-BASE について

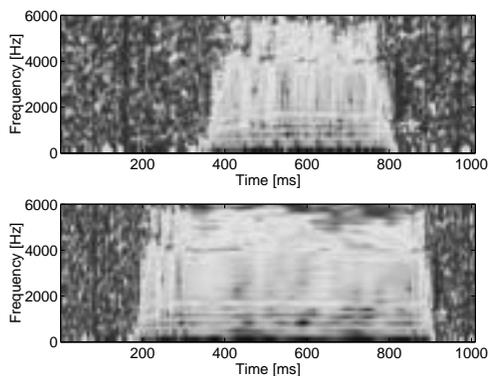
合成音による声の揺れ、響き、歌声らしさの聴取実験を行うために、分析から得られた音響的特徴を付加するためのベースとなる合成音を作成する必要がある。以下にベースとなる SYN-BASE の作成方法について示す。

SYN-BASE を作成するための元となる音声データは no.9 (テノールの男性の話声) を用いた。元音声データ (no.09) の音声の定常部 100 ms の区間における F0、スペクトル、非周期成分の割合について時間方向に spline 補間する事により、その区間を 1000 ms に伸長させた。図 4.13 に元音声データの F0、スペクトル、非周期成分の割合と、それぞれを時間方向に伸長させた SYN-BASE の F0、スペクトル、非周期成分の割合を示す。



(a) no.9 (上) と SYN-BASE (下) の F0

(b) no.9 (上) と SYN-BASE (下) のスペクトル



(c) no.9 (上) と SYN-BASE (下) の非周期成分の割合

図 4.13: SYN-BASE の作成方法

## 4.4 実験4.1 ( SYN-BASE の検証 )

### 実験目的

SYN-BASE は no.9 の音声データを時間方向に伸長させたものであり、伸長させることによって声の揺れ、響き、歌声らしさの知覚に対して与える影響を調査するために聴取実験を行う。

### 実験方法

実験は no.9 の音声データの duration を 200 ms ずつ変化させて、SYN-BASE の duration まで時間方向に伸長させたものを 6 個呈示した。被験者には、呈示した音声データが揺れているか、又は揺れていないかの 2 段階とし絶対評価により回答させた。また、響き、歌声らしさについても同様の評価尺度で行った。

### 実験結果・考察

その結果、声の揺れ、響き、歌声らしさにおいて、時間方向に伸長を行っても全ての被験者の回答には変化は見られず、声の揺れ、響き、歌声らしさの知覚は変化しなかった。また、SYN-BASE は、揺れや響きが知覚されていないことや、歌声らしくないことを確認した。以上より、伸長させることによる、声の揺れ、響き、歌声らしさの知覚に及ぼす影響は見られないことが分かった。

## 4.5 揺れに関連する音響的特徴の検証

### 4.5.1 実験4.2 ( 揺れ )

#### 実験目的

4.1 節において、分析から得られた音響的特徴が揺れに関連する特徴であるのかを検証するために聴取実験を行う。

#### 合成音の作成

分析から得られた揺れに関連する音響的特徴について、聴取実験により検証を行うために、それらの特徴を付加した合成音を作成した。以下に作成した 5 種類の合成音について示す。

## (1) SYN-BASE

no.9 の音声データの定常部を時間方向に伸長させたもの。(4.3 節に示した合成音 SYN-BASE)

## (2) SYN-F0

4.1.2 節の分析から、F0 が変動していることが分かった。そのため、SYN-BASE に対して、F0 のみ周波数変調させる事により、SYN-F0 を作成した。以下に作成方法を示す。

周波数変調を行うためには変調周波数と偏移幅を決定する必要があり、各音声データの F0 の平均値からの変位について分析を行った。その結果、揺れの上位の音声データには、5 Hz 付近に支配的な変調周波数が見られた。また、偏移幅の標準偏差は約 3 % であった。そして、式 (4.1)[10] を用いて、変調周波数  $f_s = 5$  Hz、偏移幅  $\alpha = 0.03$  (3 %) として SYN-F0 の F0  $f(t)$  を作成した。その結果を図 4.14 に示す。そして、作成した F0 と SYN-BASE のスペクトルおよび非周期成分の割合から SYN-F0 を作成した。

$$f(t) = (1 + \alpha \cdot \sin(2\pi f_s t)) \cdot f_0 \quad (4.1)$$

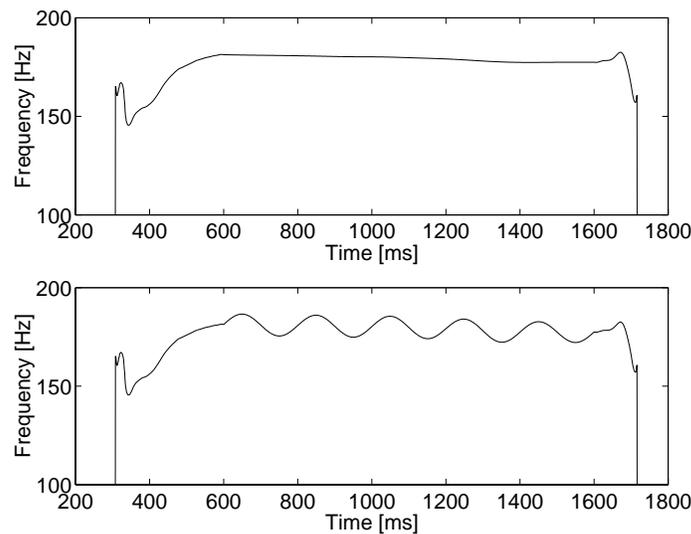


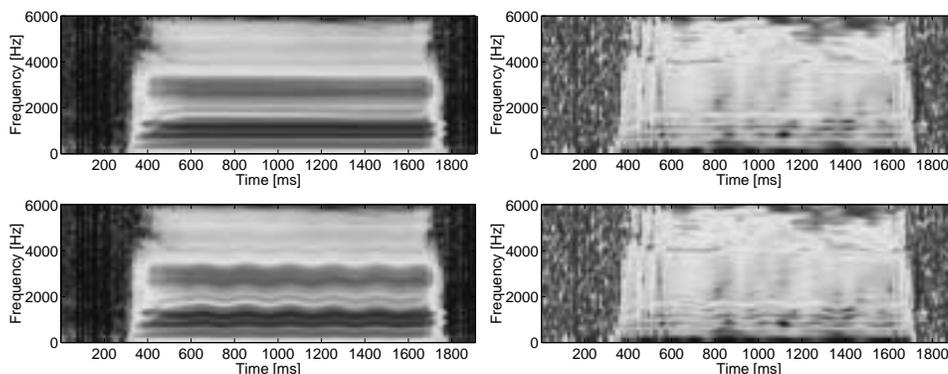
図 4.14: SYN-BASE (上) と SYN-F0 (下) の F0

## (3) SYN-SPEC

4.1.3 節の分析から、フォルマントが周波数方向に変動していることが分かった。そのため、SYN-BASE に対して、フォルマントのみ周波数変調させる事により、SYN-SPEC を

作成した。以下に作成方法を示す。

フォルマントはF0の変動と対応して周波数方向に変動していたため、周波数変調を行うための変調周波数と偏移幅は、SYN-F0の場合と同様とし変調周波数  $f_s = 5$  Hz、偏移幅を  $\alpha = 0.03$  (3%) とした。そして、スペクトルを周波数方向に伸縮させる事により、全てのフォルマントを周波数変調させた。また、非周期成分の割合に対してもスペクトルと同様の操作を行った。その結果を図 4.15 に示す。そして、作成したスペクトルおよび非周期成分の割合と SYN-BASE の F0 から SYN-SPEC を作成した。



(a) SYN-BASE(上) と SYN-SPEC(下) のスペクトル

(b) SYN-BASE(上) と SYN-SPEC(下) の非周期成分の割合

図 4.15: SYN-SPEC の作成方法

#### (4) SYN-ENV

4.1.2 節や 4.1.3 節の分析から、振幅エンベロープが変動していることやフォルマントが振幅方向に変動していることが分かった。そのため、SYN-BASE に対して振幅エンベロープのみ振幅変調させる事により、SYN-ENV を作成した。以下に作成方法を示す。

ここでの振幅エンベロープの振幅変調はフォルマントを振幅変調させる事と同義である。振幅変調を行うためには変調周波数と偏移幅を決定する必要がある、各音声データの振幅エンベロープの平均値からの変位について分析を行った。その結果、揺れの上位の音声データには、5 Hz 付近に支配的な変調周波数が見られた。また、偏移幅の標準偏差は 20% ~ 40% であった。そこで、以下に示す式 (4.2) を用いて、振幅変調させた振幅エンベロープ  $e_k(t)$  を作成した。但し、変調周波数  $f_s = 5$  Hz、偏移幅  $\alpha = 0.2$  (20%) とした。その結果を図 4.16 に示す。そして、式 (4.3) [11] を用いて、SYN-BASE のキャリア成分と作成した振幅エンベロープから SYN-ENV を作成した。但し、 $l(t)$  は SYN-BASE の時間波形、 $e_1(t)$  は SYN-BASE の振幅エンベロープとする。

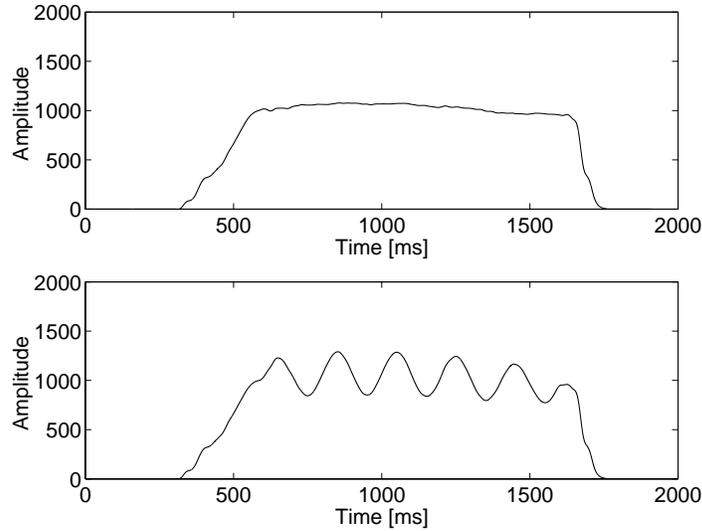


図 4.16: SYN-BASE(上) と SYN-EBV(下) の振幅エンベロープ

$$e_k(t) = (1 + \alpha \cdot \sin(2\pi f_s t)) \cdot e_1(t) \quad (4.2)$$

$$k(t) = e_k(t) \times \frac{l(t)}{e_1(t)} \quad (4.3)$$

## (5) SYN-ALL

SYN-ALL は、SYN-BASE の合成音に対して上記に示した SYN-F0、SYN-SPEC、SYN-ENV の全ての操作を行うことに、より作成した。但し、4.1.2 節の分析により、揺れの知覚には F0、スペクトル、振幅エンベロープは同位相である事が重要である可能性があり、スペクトルと振幅エンベロープの変動は F0 の変動と同位相となるように合成音を作成した。

## 実験方法

実験はシェッフェの対比較法を用いて、被験者には声の揺れについて図 4.17 に示す 5 段階の評価尺度で回答させた。また、呈示音は前節で示した 5 種類の合成音であり、刺激対の数は逆の順序も含め合計 20 個である。被験者は正常な聴力を有する大学院生 8 名であり、1 人の被験者について 1 回の実験を行った。また、被験者には実験を行う前に予備的な訓練を行い評価尺度に慣れてもらった。

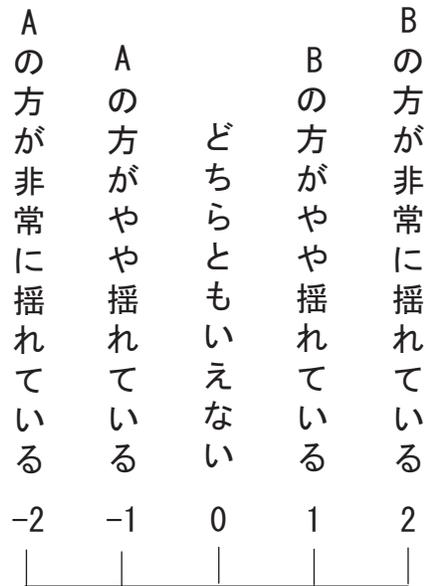


図 4.17: 評価尺度

## 実験結果・考察

実験から得られた全被験者の評価値を浦の変法 [7] によって処理を行い、揺れにおける音声データ間の心理的距離を直線上に示した結果を図 4.18 に示す。図において、数値が大きいものほど揺れが知覚されていることを表す。

実験結果より、SYN-F0、SYN-SPEC、SYN-ENV、SYN-ALL の合成音は、SYN-BASE と比較して正の側にある事から揺れが知覚されていると考えられる。また、SYN-ALL が SYN-F0、SYN-SPEC、SYN-ENV の合成音と比較して正の側にあることから、フォルマントの周波数変調および振幅エンベロープの振幅変調と F0 の変動とを同位相で変動させることにより、より揺れが知覚されたと考えられる。以上の考察から、F0 と振幅エンベロープの 4 ~ 6 Hz の変位、又、フォルマントの周波数変調や振幅変調、さらにこれらの関係が同位相であることは揺れに関連する音響的特徴である事が確認できた。

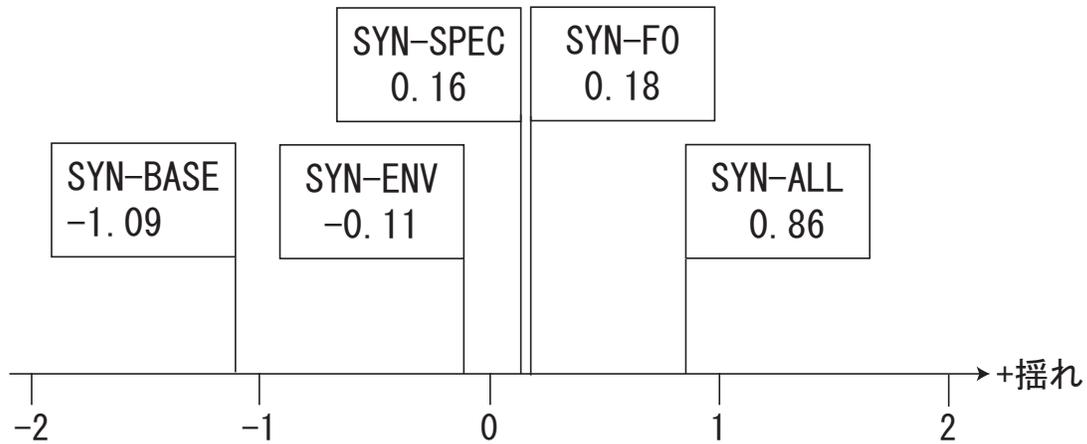


図 4.18: 合成音の揺れの関係

## 4.6 響きに関連する音響的特徴の検証

### 4.6.1 実験 4.3 (響き)

#### 実験目的

4.2 節において、分析から得られた音響的特徴が響きに関連する特徴であるのかを検証するために聴取実験を行う。

#### 合成音の作成

分析から得られた響きに関連する音響的特徴について、聴取実験により検証を行うために、それらの特徴を付加した合成音を作成した。以下に作成した4種類の合成音について示す。

#### (1) SYN-BASE

no.9 の音声データの定常部を時間方向に伸長させたもの (4.3 節に示した合成音 SYN-BASE)

## (2) SYN-UP

4.2.2 節の分析から、スペクトル包絡の 3kHz 付近に顕著なピークが見られた。そのため、SYN-BASE に対して、スペクトル包絡の 3 kHz 付近にある最も高いピークを中心に全体的に持ち上げる事により、SYN-UP を作成した。以下に作成方法を示す。

このような操作を行うために、何 dB 持ち上げるのか、又、その帯域幅を決定する必要がある。そこで、同一発声者であり、響きの順位が 2 位と 9 位の no.1 と no.9 の音声データの時間平均したスペクトル包絡について分析を行った。その結果、響きの順位が高い no.1 は、no.9 に比べスペクトル包絡の 3kHz 付近のピークがおおよそ 12dB 持ち上がっていることが分かった。また、その帯域幅は約 2000 Hz であった。この結果から、図 4.19 にスペクトル包絡と非周期成分の割合に対して行った操作を示す。図に示す窓関数（最大値 12 dB、帯域幅 2000 Hz）を用いて、3~4 kHz 付近にある最も高いピークを窓の中心とし、そのピークが 12 dB 持ち上がるように、スペクトル包絡を変形した。また、非周期成分の割合に対しても、スペクトル包絡と対応させるように変形を行った。図 4.20 に音声の定常部 1000 ms に対して上記に示した操作を行い、作成したスペクトルと非周期成分の割合を示す。そして、作成したスペクトルおよび非周期成分の割合と SYN-BASE の F0 から SYN-UP を作成した。

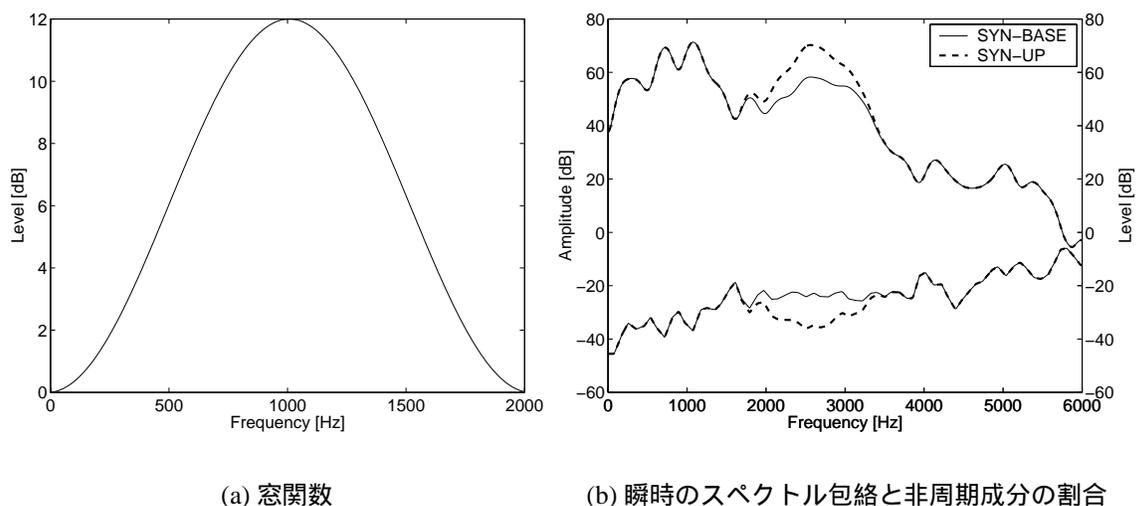
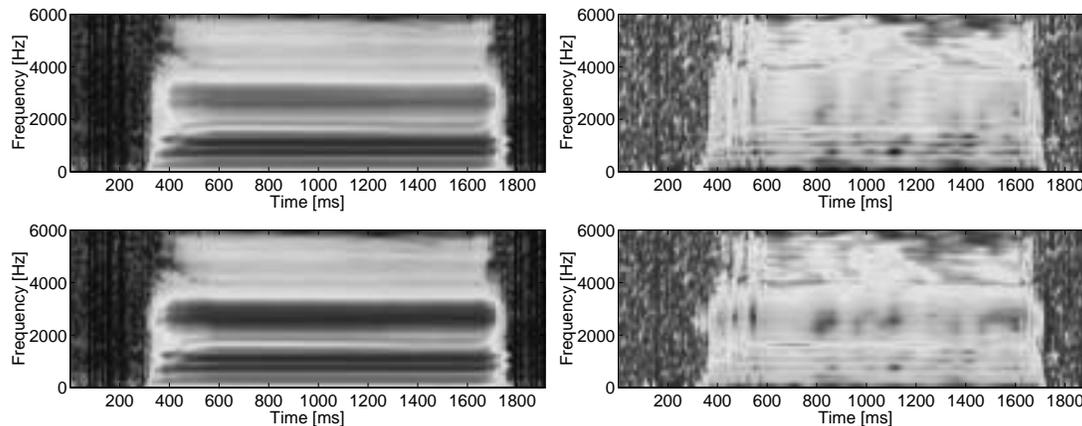


図 4.19: SYN-UP の作成方法



(a) SYN-BASE(上)とSYN-UP(下)のスペクトル

(b) SYN-BASE(上)とSYN-UP(下)の非周期成分の割合

図 4.20: SYN-UP のスペクトルと非周期成分の割合

### (3) SYN-MOVE

4.2.3 節の分析から、高調波とスペクトル包絡のピークが重なり合うことによって、3 kHz 付近の非周期成分の割合に強い高調波が見られた。そのため、SYN-BASE に対して、スペクトル包絡の 3 kHz 付近にある最も高いピークと高調波の周波数を合わせ、その周波数の高調波成分を強調させることにより、SYN-MOVE を作成した。以下に作成方法を示す。

まず、スペクトル包絡の 3 kHz 付近にある最も高いピークと高調波を合わせた結果を図 4.21 に示す。次に図 4.22 にスペクトル包絡と非周期成分の割合に対して行った操作を示す。図に示した窓関数（最大値 12 dB、帯域幅 500 Hz）を用いて、3 kHz 付近にある最も高いピークを窓の中心とし、そのピークが 12 dB 持ち上がるように、スペクトル包絡を変形させた。但しスペクトルの変形前と変形後ではパワーは同じになるようにした。また、非周期成分の割合に対しても、スペクトル包絡と対応させるように変形を行った。図 4.23 に音声の定常部 1000 ms に対して上記に示した操作を行い、作成したスペクトルと非周期成分の割合を示す。そして、作成したスペクトルおよび非周期成分の割合と SYN-BASE の F0 から SYN-MOVE を作成した。

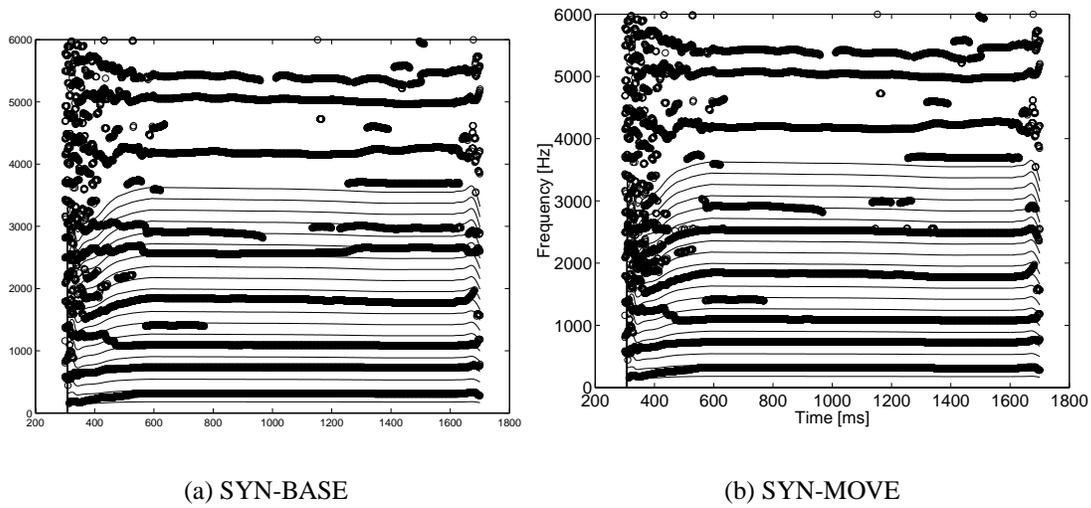


図 4.21: SYN-MOVE のスペクトル包絡のピークと高調波の操作

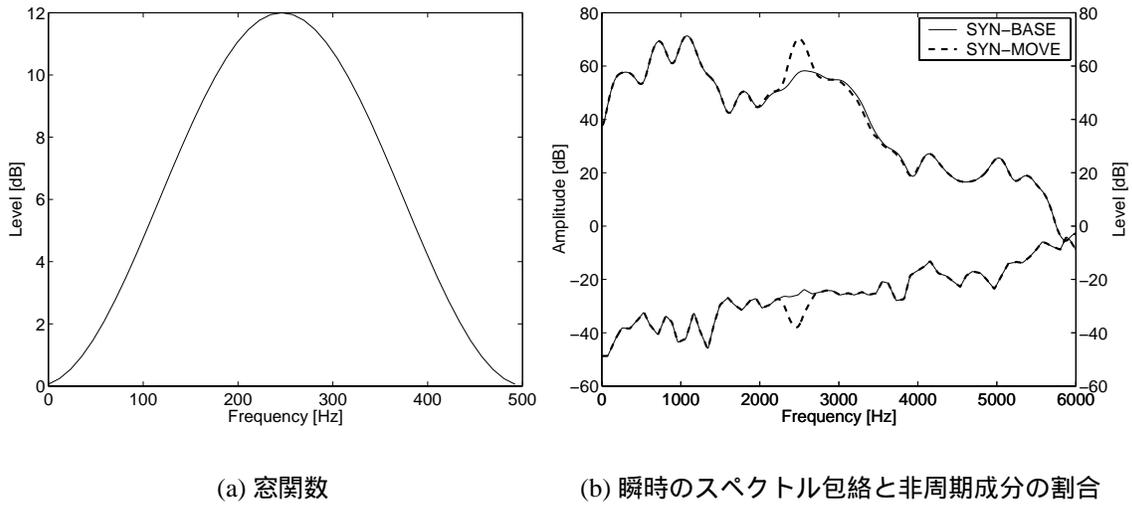
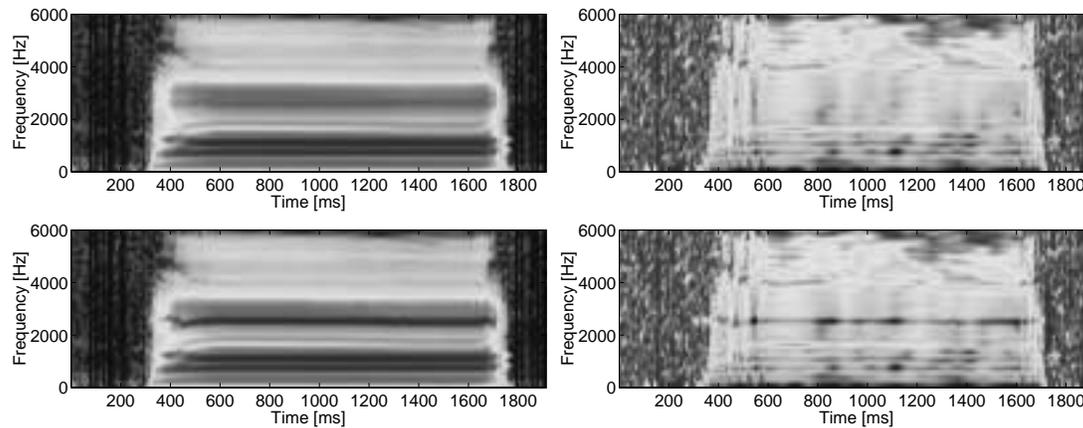


図 4.22: SYN-MOVE のスペクトルと非周期成分の割合の操作



(a) SYN-BASE(上) と SYN-MOVE(下) のスペクトル

(b) SYN-BASE(上) と SYN-MOVE(下) の非周期成分の割合

図 4.23: SYN-MOVE のスペクトルと非周期成分の割合

#### (4) SYN-ALL

SYN-ALL は、SYN-BASE の合成音に対して上記に示した SYN-UP、SYN-MOVE の全ての操作を行うことにより作成した。

#### 実験方法

実験はシェッフェの一対比較法を用いて、被験者には声の響きについて図 4.24 に示す 5 段階の評価尺度で回答させた。また、呈示音は前節で示した 4 種類の合成音であり、刺激対の数は逆の順序も含め合計 12 個である。被験者は正常な聴力を有する大学院生 8 名であり、1 人の被験者について 1 回の実験を行った。また、被験者には実験を行う前に予備的な訓練を行い評価尺度に慣れてもらった。

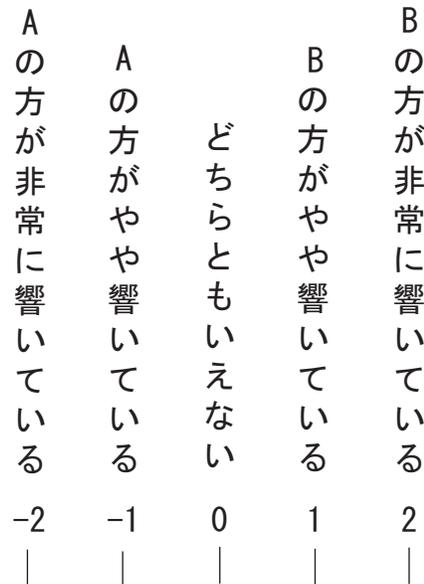


図 4.24: 評価尺度

## 実験結果・考察

実験から得られた全被験者の評価値を浦の変法 [7] によって処理を行い、響きにおける音声データ間の心理的距離を直線上に示した結果を図 4.25 に示す。図において、数値が大きいものほど響きが知覚されていることを表す。

実験結果より、SYN-UP、SYN-MOVE、SYN-ALL の合成音は、SYN-BAS と比較して正の側にある事から響きが知覚されていると考えられる。しかし、SYN-MOVE h は SYN-UP や SYN-ALL と比べると SYN-BASE に近い事から、3 kHz 付近の高調波とスペクトル包絡のピークが重なることにより見られた強い高調波と響きとの関係は検討する必要があると考えられる。例えば、今回、合成音を作る際にベースとして用いた SYN-BASE は男性の話声の音声データであり、基本周波数が低く、高調波の間隔が狭い。そのため、常に 3 kHz 付近の高調波とピークと近い、又は一致していたため、高調波とピークを重ねる事による影響があまり表れなかった可能性が考えられる。しかし、SYN-ALL は SYN-UP より正の側にあることから、SYN-MOVE は響きに及ぼす影響があると推測される。以上の考察から、3 kHz 付近のピークと、3 kHz 付近の高調波とスペクトル包絡のピークを重ね合わせる事による強い高調波成分、さらにはこれらの相互関係は響きに関連する音響的特徴である事が確認できた。

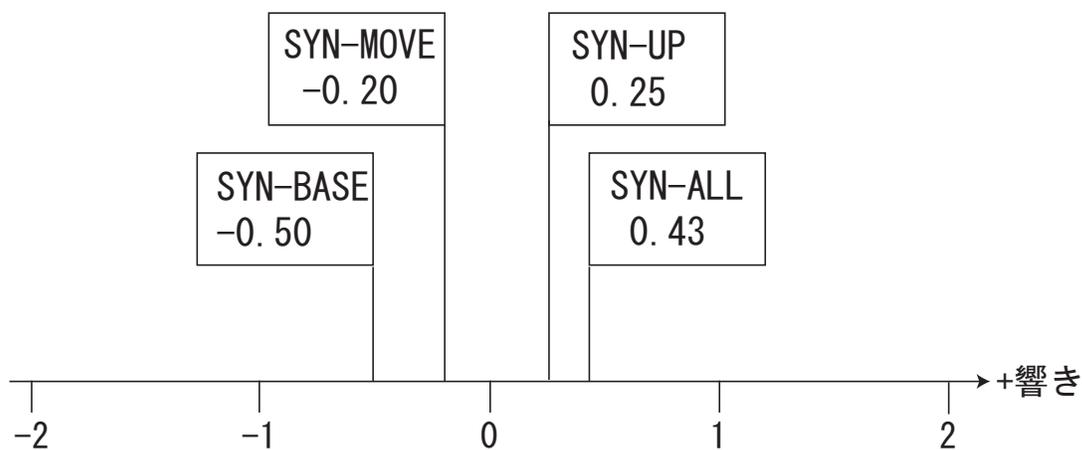


図 4.25: 合成音の響きの関係

## 4.7 まとめ

本章では、歌声らしさの要因である声の揺れと響きに関連する音響的特徴の分析を行った。また、それらの音響的特徴を付加した合成音を作成し聴取実験を行うことにより、声の揺れと響きの心理的特徴とそれらの音響的特徴との関係について検証を行った。

その結果、F0 と振幅エンベロープの 4~6 Hz の変位、又、フォルマントの周波数変調や振幅変調、さらにこれら関係が同位相であることは揺れに関連する音響的特徴である事が分かった。また、スペクトル包絡の 3~4 kHz 付近の顕著なピークと、3~4 kHz 付近の高調波とスペクトル包絡のピークを重ね合わせる事による強い高調波成分は、響きに関連する音響的特徴であることが分かった。以上の結果より、図 1.2 に示した歌声らしさの 3 層モデルのうち第 2 層目（基本的な心理的特徴）と第 3 層目（物理的特徴）の関係が明らかになった。

# 第5章 歌声らしさに影響を及ぼす音響的特徴の検討

本章では、図 1.2 に示した歌声らしさの3層モデルのうちの第1層目（歌声らしさ）と第3層目（物理的特徴）の関係について検討する。前章では、声の揺れ、響きに関連する音響的の分析を行った。また、それらの関係について合成音による検証を行い、対応関係が取れている事を確認した。そこで、第3層目の揺れや響きに関連する音響的特徴を組み合わせ2層目の揺れや響きの心理量を変化させる事により、第1層目の歌声らしさがどのように変化するのか検討を行うために聴取実験を行う。

## 5.1 歌声らしさに影響を及ぼす音響的特徴の検討

### 5.1.1 実験 5.1（歌声らしさ）

#### 実験目的

本研究の分析から得られた揺れや響きに関連する音響的特徴を組み合わせ、2層目の歌声らしさの要因である心理的特徴を変化させる事により、1層目の歌声らしさに対する知覚が変化するのか検証するために聴取実験を行う。

#### 合成音の作成

揺れや響きに関連する音響的特徴を組み合わせ、2層目の基本的な心理的特徴を変化させる事により、1層目の歌声らしさがどのように変化するのか聴取実験により検討するために、それらの特徴を付加した合成音を作成する。以下に作成した4種類の合成音について示す。

- SYN-BASE  
no.9の音声データの定常部を時間方向に伸長させたもの(4.3節に示した合成音 SYN-BASE)
- SYN-VIV  
SYN-BASE に対して、本研究で得られた揺れに関連する音響的特徴を全て付加した

もの。4.5 節における揺れの音響的特徴を全て付加した SYN-ALL と同様である。

- SYN-RIN

SYN-BASE に対して、本研究で得られた響きに関連する音響的特徴を全て付加したもの。4.6 節における響きの音響的特徴を全て付加した SYN-ALL と同様である。

- SYN-ALL

SYN-BASE に対して、SYN-VIV、SYN-RING の全ての操作を行ったもの

## 実験方法

実験はシェッフェの一対比較法を用いて、被験者には歌声らしさについて図 5.1 に示す 5 段階の評価尺度で回答させた。また、呈示音は前節で示した 4 種類の合成音であり、刺激対の数は逆の順序も含め合計 12 個である。被験者は正常な聴力を有する大学院生 8 名であり、1 人の被験者について 1 回の実験を行った。また、被験者には実験を行う前に予備的な訓練を行い評価尺度に慣れてもらった。

A	A		B	B
の	の		の	の
方	方	ど	方	方
が	が	ち	が	が
非	が	ら	が	非
常	や	と	や	常
に	や	も	や	に
歌	歌	い	歌	歌
声	声	え	声	声
ら	ら	な	ら	ら
し	し	い	し	し
い	い		い	い
-2	-1	0	1	2

図 5.1: 評価尺度

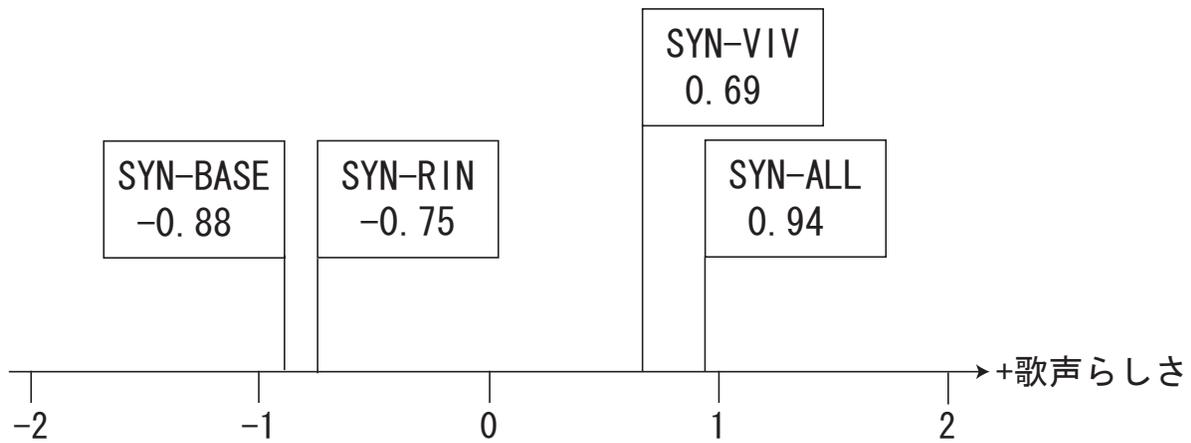


図 5.2: 合成音の歌声らしさの関係

## 実験結果・考察

実験から得られた全被験者の評価値を浦の変法 [7] によって処理を行い、歌声らしさにおける音声データ間の心理的距離を直線上に示した結果を図 5.2 に示す。図において、数値が大きいものほど、その音声データが歌声らしいことを意味する。

実験結果より、SYN-VIV、SYN-RIN、SYN-ALL の合成音は、SYN-BASE と比較して正の側にある事から歌声らしさが知覚されていると考えられる。SYN-VIV は SYN-BASE との間に大きな差があることから揺れは歌声らしさに強く影響を及ぼすと考えられる。また、SYN-RIN は SYN-BASE と距離が近いため、響きのみでは歌声らしさはあまり影響を及ぼしていない事が分かる。しかし、SYN-ALL が SYN-VIV より正の側にある事から、響きは、揺れと合わせるにより、より歌声らしくする要因の一つであると考えられる。

## 5.2 まとめ

本章では、本研究より明らかになった第 3 層目の揺れや響きに関連するの音響的特徴を組み合わせを変化させ、第 2 層目の揺れや響きの心理量を変化させる事により、第 1 層目の歌声らしさがどのように変化するか検討を行った。その結果、揺れや響き、又はそれらの音響的特徴を組み合わせる事により、歌声らしさの知覚が変化したことから、図 1.2 に示した歌声らしさの 3 層モデルのうち各層の対応付けができていたことが明らかになった。また、揺れに関連する音響的特徴は歌声らしさに強く影響を及ぼすことが分かった。そして、歌声らしさに及ぼす影響は揺れにくらべると小さいが、響きに関連する音響的特徴は、より歌声らしくする要因の一つであると考えられる。

# 第6章 結論

## 6.1 本論文のまとめ

本研究では、歌声らしさという心理的特徴と物理的特徴の関係について検討するために図 1.2 に示した歌声らしさの3層モデルを提案し、それぞれに関係について検討し対応付けを行った。

モデルの第1層目（歌声らしさ）と第2層目（基本的な心理的特徴）の関係において、MDS と重回帰分析より、声の揺れ、響き、明瞭さは歌声らしさの要因であることが分かった。

モデルの第2層目（基本的な心理的特徴）と第3層目（物理的特徴）の関係において、歌声らしさの要因である声の揺れと響きに関連する音響的特徴の抽出を行った。そして、それらの音響的特徴を付加した合成音を作成し、聴取実験を行うことにより、分析から得られた音響的特徴と歌声らしさの要因である声の揺れや響きとの対応関係について検証した。その結果、F0 と振幅エンベロープの4~6 Hz の変位、又、フォルマントの周波数変調や振幅変調、さらにこれらの関係が同位相であることは揺れに関連する音響的特徴である事が分かった。また、スペクトル包絡の3~4 kHz 付近の顕著なピークと、高調波とスペクトル包絡の3~4 kHz にあるピークとを合わせるにより表れる強い高調波成分は、響きに関連する音響的特徴であることが分かった。

モデルの第1層目（歌声らしさ）と第3層目（物理的特徴）の関係において、本研究から明らかになった第3層目の声の揺れや響きに関連する音響的特徴を組み合わせ、第2層目の揺れや響きの心理量を変化させる事により、第1層目の歌声らしさがどのように変化するのか検証した。その結果、揺れや響きに関連する音響的特徴を組み合わせる事により、第1層目の歌声らしさが変化したことから、モデルの各層の対応付けができていたことが示された。また、揺れに関連する音響的特徴は歌声らしさに強く影響を及ぼすことや、響きに関連する音響的特徴は、歌声らしさに及ぼす影響は揺れにくらべると小さいが、より歌声らしくする要因の一つであることが分かった。

## 6.2 今後の課題

以下に今後の課題を示す。

- 歌声らしさの要因は、声の揺れ、響き、明瞭さ以外にも存在する可能性がある。本研究では、聴取実験の結果から、歌声らしさの要因の候補として、声の明るさ、柔らかさ、揺れ、響き、伸び、長さ、明瞭さの7個の表現語が考えられる。しかし、今回、その中から歌声らしさに強く関連すると考えられる声の揺れ、響きについては検討を行ったが、それ以外のものについて検討を行っていない。そのため、それらの表現語についても検討を行い、歌声らしさの要因について調査を行う必要がある。
- 今回、響きに関連する音響的特徴の分析を行ったが、11個の音声データに共通した音響的特徴を明らかにすることはできなかった。そのため、響きに関連する音響的特徴については、さらに分析を行う必要がある。
- 本研究で提案した歌声らしさの3層モデルの対応関係について検討を行った。今後、歌声合成などの応用を考えると、分析から得られた音響的特徴の量的な違いがどのように歌声らしさに影響を及ぼすのか検討する必要がある。

# 謝辞

本研究を進めるにあたり、多大なる御指導ならびに御鞭撻を賜りました赤木正人教授に深く感謝の意を表します。

また、御有益な助言を賜りました党建武助教授、ならびに鵜木祐史助手に心より感謝致します。そして、日頃から多大なる討論と激励を頂きました伊藤一仁様、齋藤毅様、ならびに赤木研究室の皆様に厚くお礼申し上げます。

最後に、大学院での貴重な研究生生活を支えて頂いた、家族や友人に心から感謝致します。

## 参考文献

- [1] Sundberg, J, "The Science Of The Singing-voices,"Northern Illinois University Press. 1987 .
- [2] 齋藤毅, 鷓木祐史, 赤木正人, "歌声のF0制御モデルにおけるパラメータ決定に関する考察," 日本音響学会聴覚研究会資料, Vol.33, No.10 H-2003-111, 2003 .
- [3] 西内美登里, 大串健吾, " 専門家と非専門家の歌声の評価," 日本音響学会聴覚研究会資料, H-90-1, 1990 .
- [4] 中山一郎, " 日本語を歌・唄・謡う," 日本音響学会会誌 59 卷 11 号, 特別公演, pp.688-693, 2003 .
- [5] 林 知己夫, 飽戸 弘, "多次元尺度解析法" サイエンス社, 1976.
- [6] 上田 和夫, "音色の表現語に階層構造は存在するか," 日本音響学会誌 44 卷 2 号, pp.102-107, 1988.
- [7] 天坂 格郎, 長沢 伸也, "官能評価の基礎と応用," 日本規格協会, 2000.
- [8] Kawahara . H . . Masuda - Katsuse. I . . , and de Cheveigne , A, "Restructuring speech representations using a pitch adaptive time-frequency smoothing and an instantaneous-frequency based on F0 extraction ;Possible role of a repetitive structure in sounds," Speech Communication , Vol.27 , pp.187-207 , 1999 .
- [9] 中山 一郎, 小林 範子, " 歌の声," 日本音響学会誌 52 卷 5 号, pp.383-388 , 1996 .
- [10] 北風裕教, 赤木正人, " 基本周波数の微細変動成分に対する知覚," 日本音響学会聴覚研究会資料, H-2000-23 , 2000 .
- [11] 広林茂樹, 野村博昭, 小池恒彦, 東山三樹夫, "パワーエンベロープ伝達関数の逆フィルタ処理による残響音声の回復," 電子情報通信学会論文誌 A, Vol.J81-A, No.10 pp.1323-1330, 1998.

# 学会発表リスト

辻 直也，赤木正人，”歌声らしさの要因とそれに関連する音響特徴量の検討，” 日本音響学会聴覚研究会資料，Vol.34，No.1 H-2004-8，2004 .