

# 長唄における韻律的特徴のモデル化とその評価

松岡 文吾<sup>†</sup> 峯松 信明<sup>†</sup> 広瀬 啓吉<sup>††</sup>

<sup>†</sup> 東京大学大学院情報理工学系研究科 〒113-0033 東京都文京区本郷 7-3-1

<sup>††</sup> 東京大学大学院新領域創成科学研究科 〒113-0033 東京都文京区本郷 7-3-1

E-mail: †{matsuoka,mine,hirose}@gavo.t.u-tokyo.ac.jp

**あらまし** 情緒や雰囲気を重視する日本の伝統的歌唱法と、音の響きを重視する西洋的歌唱法「ベル・カント」では韻律的特徴に多くの違いが存在する。しかし、日本の歌唱法はベル・カントに比べてまだ研究例も少なく、音響的・生理的特徴に未知な部分が多い。本研究では日本の伝統的歌唱法「長唄」を対象とし、同一の歌い手と楽譜によるベル・カントとの比較から長唄特有の装飾音の韻律的特徴を分析し、この装飾音を新たに指令として導入した2次系の制御モデルを提案する。評価実験では、実際の長唄音声、モデルから生成した長唄的韻律遷移を用いて再合成した音声、実際の長唄から抽出した韻律遷移を用い再合成した音声などを比較し、提案モデルの妥当性を検証する。

**キーワード** 長唄, ベル・カント, 韻律的特徴, 装飾音, モデル化

## Modeling of prosodic features in NAGAUTA singing and its evaluation.

Bungo MATSUOKA<sup>†</sup>, Nobuaki MINEMATSU<sup>†</sup>, and Keikichi HIROSE<sup>††</sup>

<sup>†</sup> Graduate School of Information Science and Technology, University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033 Japan

<sup>††</sup> Graduate School of Frontier Sciences, University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033 Japan

E-mail: †{matsuoka,mine,hirose}@gavo.t.u-tokyo.ac.jp

**Abstract** In previous studies on analysis and modeling of singing speech, Western (originally Italian) style of singing, BEL CANTO, was often focused. In this paper, we investigated one of the Japanese classical singing styles, NAGAUTA, which respects atmosphere and mood in the singing while BEL CANTO does resonance of singing in the hall. These two styles also differ in their prosody and NAGAUTA's pitch patterns are remarkably unique with its grace notes. In this work, singing materials with BEL CANTO and those with NAGAUTA were firstly prepared for the same songs from a single singer. By comparing the prosodic patterns between the two, we proposed a model to generate NAGAUTA's prosody from the standard notation. The model was represented by a second-order system and its parameters were greedily estimated. Evaluation experiments were done using re-synthesized speech with the original prosody and that with the modeled prosody. Results showed high validity of the proposed model.

**Key words** NAGAUTA, BEL CANTO, prosodic features, grace note, modeling

### 1. はじめに

近年のコンピュータ性能の向上などによって音声合成技術は格段に進歩し、特に朗読文章に関しては自然に近い韻律を持った声の合成が可能となってきた。しかし、一方で歌声の合成に関してはそれほど研究がなされていない。その理由として、まず歌唱能力の高い歌い手に実験に協力してもらうことが極めて困難であり、必然的に扱える被験者やデータが限られてしまうことがあげられる。また、歌い手の表現方法として「額の真中に音を当てる」など主観的な表現がなされる場合が多く、しかも歌い手によってその意味するところが異なるため、研究者と歌い手の間で意図を理解し合える共通言語をもつことが難しいということも研究を困難にさせる理由の一つとされる [1]。

歌声に関する過去の実験例を見ても、被験者として選ばれるのがオペラ歌手の場合、コーラス経験者の場合、ボイススクー

ルの学生の場合など、歌唱法はさまざまであるにも関わらず、全てをいわゆる西洋的歌唱法「ベル・カント」として同一視している感が強く、歌声の分析やモデル化が歌唱法ごとに扱われる機会は極めて少ないのが現状である。そもそも、世界中には極めて多くの歌唱法があるにも関わらず、ほとんどの研究がこういったベル・カントのみを対象としているのも事実である。

しかし、ベル・カントは全ての歌唱法に適用出来るというわけではない。その例として「日本語で書かれている歌詞をオペラなどの洋楽の唱法で歌った場合には、響きを重要視するあまりに日本語としてのニュアンスや自然さが失われ、何を言っているのか分からない」とし、日本の歌唱法とベル・カントの歌唱法の比較を目的とした研究もある [2]。つまり、歌唱法にはそれぞれの特性があり、自然な歌声合成の実現にはこれらの歌唱法による特性の違いを見分けることが必須の要素技術であると考えられる。また、民族音楽の中には音楽文化の継承が難しく

なっているものもあり、師弟間の口伝による継承だけではなく、物理的な実態を客観的に分析しておくことが今後ますます要求されてくる。

本研究では日本の伝統的歌唱である「長唄」に焦点を絞り、長唄に含まれる「日本の歌唱法らしさ」、さらには「長唄らしさ」そのものの分析と合成を目標とし、長唄の持つ独特の韻律遷移についてのモデル化を試みる。

## 2. 歌唱法による韻律遷移の違い

### 2.1 歌声資料の録音

伝統的な日本の歌唱法には、西洋のオペラや合唱スタイルには見られない韻律的な特徴がいくつか存在する [2]。その1つの例として、音程の変化時もしくは同一音程中であっても譜面の進行に伴って発生する基本周波数 ( $F_0$ ) の突発的な振動があげられる。これらは一般的に「装飾音」と呼ばれ、その表現方法は歌手の長年の勘や経験などから状況によって様々に変化するため、楽譜に書かれたタイミングで正しい音程に沿って歌う西洋的な歌唱法とは根本的に異なっている。

本研究では装飾音として、音程が変化する場合に付加される「ふり (上げふり, 下げふり)」と、同じ高さが続く場合に付加される「あたり」についての分析を行う。分析に先立ち、韻律遷移に与えるこれらの装飾音の影響を分析するため、以下に示す3種類の楽譜を用いた歌声資料の録音を行った。なお、長唄では楽譜として五線譜が使用されることはあまり無いが、五線譜でも長唄として歌うことに問題はないというコメントを歌手からもらっている。

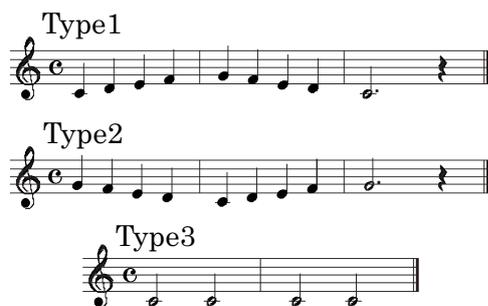


図1 3種類の楽譜

歌声資料の録音は、簡易防音室を用いて行った。歌手は長唄を15年間経験し、現在は合唱団に在籍して6年となる女性である。録音パターンについては、装飾音「ふり」についてはC4 (261.63Hz) からG4 (391.99Hz) まで上がって下がるType1とG4からC4まで下がって上がるType2を、装飾音「あたり」についてはC4を4度繰り返すType3を用意し、それぞれについて母音の/a/のみを用いて長唄スタイルとコーラススタイルの2通りで歌ったものを録音した。その後、サンプリング周波数16kHz、量子化ビット数16bitにてデジタル化を行った。

### 2.2 基本周波数遷移の抽出

基本周波数の分析は matlab 上で起動する高品質音声分析変換成法である STRAIGHT (v30kr16) [3] を用いた。装飾音「ふり」を示すものとして、STRAIGHT により抽出した Type1 の長唄スタイルでの基本周波数遷移を図2に、同じく Type1 のコーラススタイルでの基本周波数遷移を図3に示す。

同様に装飾音「あたり」を示すものとして、Type3 の長唄ス

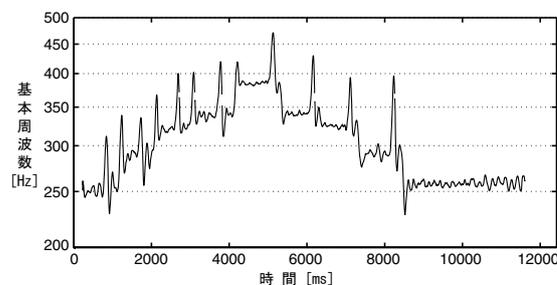


図2 Type1 の長唄スタイルでの基本周波数遷移

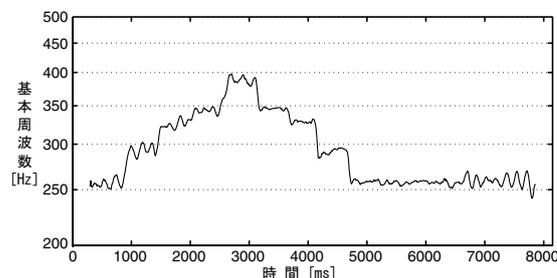


図3 Type1 のコーラススタイルでの基本周波数遷移

スタイルでの基本周波数遷移を図4に、同じく Type3 のコーラススタイルでの基本周波数遷移を図5に示す。

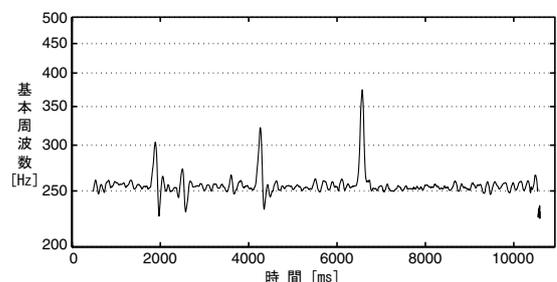


図4 Type3 の長唄スタイルでの基本周波数遷移

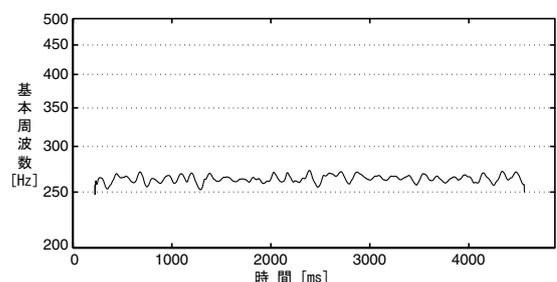


図5 Type3 のコーラススタイルでの基本周波数遷移

### 2.3 パワー遷移の抽出

過去の研究例では、装飾音がパワー遷移に与える影響として、一般的な朗読音声には見られない変動成分が確認されている [4]。この変動成分を再確認するため、パワーに関しても基本周波数の場合と同様にして STRAIGHT を使い、フレーム幅1024、フレームシフト1msecにて分析を行った。

抽出した Type1 の長唄スタイルでの分析結果を図6に、同じく Type1 のコーラススタイルでの分析結果を図7に示す。上から順にパワー、基本周波数、波形をそれぞれ表している。

朗読音声の場合には「基本周波数の上昇に伴うパワーの上昇」が見られるが [5]、長唄の装飾音ではこの現象と異なり、図6に示すような「基本周波数の突発的な上昇に伴うパワーの急激な

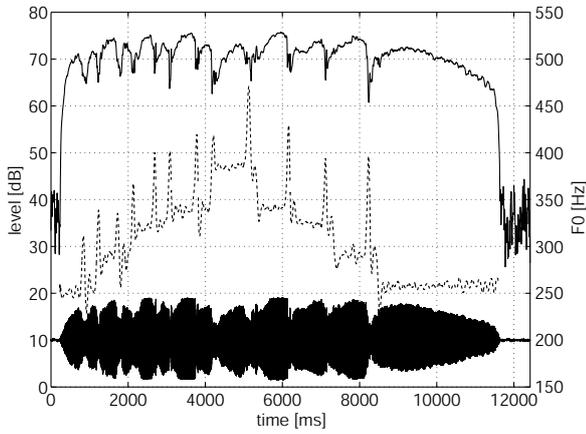


図6 Type1の長唄スタイルでの分析結果

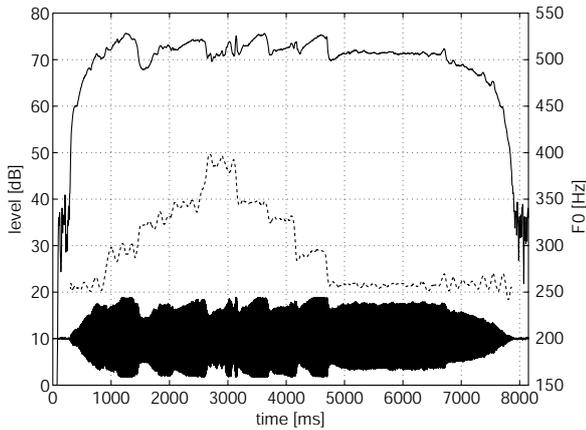


図7 Type1でのコーラススタイルでの分析結果

下降」が報告されている。この背景には、装飾音を付加する際の急激な基本周波数操作に伴う人間の生理的特性があるとされているが、詳細は分かっていない。

### 3. 長唄的韻律的特徴のモデル化

#### 3.1 基本周波数遷移のモデル化

##### 3.1.1 モデルの概要

歌声のモデル化に関しては、式1に示す2次系の伝達関数を用いて、音程（基本周波数値）を大きさとしたステップ指令に対する過渡応答に、ビブラート成分として正弦波などを直接付加するなどして表現されることが多い。

$$H(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2} \quad (1)$$

今回提案するモデルは、図8に示すようにこれら2つの成分の他に新たに装飾音を表現するための短時間ステップを「 $F_0$ 装飾音指令」として追加し、長唄の基本周波数遷移を生成しようとするものである。

##### 3.1.2 $F_0$ 装飾音指令のパラメータ抽出

図8に示す $F_0$ 装飾音成分の抽出については、長唄スタイルとコーラススタイルの両者において目視により音程遷移指令の立ち上がりを検出し、これを原点として揃えた時の対数軸上での引き算で行った。この処理によって得られた音程上昇部分の装飾音「ふり」の成分を図9に示す。

この図から1つの音程遷移個所に対して2つの突発的な振動

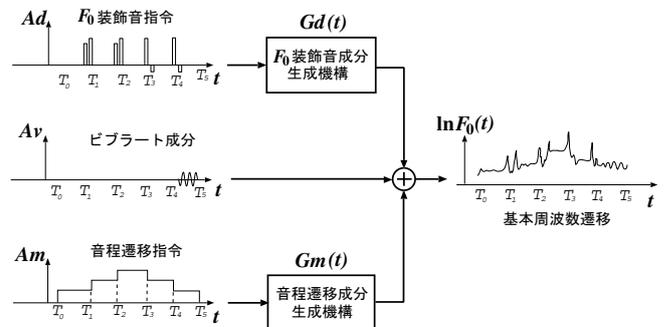


図8 基本周波数遷移生成モデルの概要

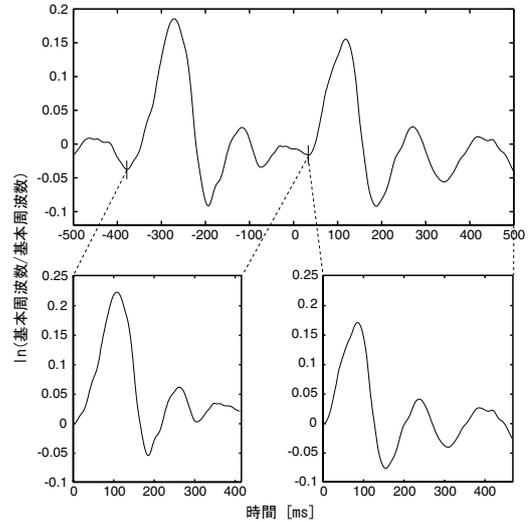


図9 対数軸上での長唄とコーラスの周波数遷移差

成分を見て取れるが、今回はこれを上向きのステップ応答として近似した。音程下降部分でも上昇部分と同様にして2つの突発的振動を検出することが出来たが、2つ目の振動成分は下向きの弱いステップ応答として検出された。

このようにして得たそれぞれの振動成分を以下の順序に従って短時間ステップ応答として近似し、モデルのパラメータを検出した。なお、ここでは1回目のダンピングの頂上の時間が立ち下がり時間となるような短時間ステップ指令に対する応答として近似を行った。

(1)  $F_0$ 装飾音成分の立ち上がり位置と1回目のダンピングの頂上、そして3回目までのダンピング時間間隔から、大まかなくと $\omega$ の関係を取得する。

(2) 同時に1回目のダンピングの大きさから、 $\zeta$ と $\omega$ の関係に対する短時間ステップ指令の大きさも取得する。

(3) その後、 $\zeta$ と $\omega$ 、短時間ステップ指令の大きさ、 $F_0$ 装飾音成分の立ち上がり位置をそれぞれ少しずつ調整しながら、短時間ステップ応答としての最適化を行う。

(4) その際の評価尺度として、調整した $F_0$ 装飾音成分の立ち上がり位置( $t_0$ )から3回目のダンピングの頂上( $t_3$ )までの区間に対し、近似した短時間ステップ応答 $y(t, \zeta, \omega)$ と実際の $F_0$ 装飾音成分 $x(t)$ との二乗誤差が最小となる $\zeta$ と $\omega$ の組、つまり二乗誤差に対する比較区間長(式2)が最も大きくなる $\zeta$ と $\omega$ の組を2次系のパラメータ、短時間ステップ指令の大きさを $F_0$ 装飾音指令の大きさとする。

$$F(\zeta, \omega) = \frac{t_3 - t_0}{\sum_{t=t_0}^{t_3} \{x(t) - y(t, \zeta, \omega)\}^2} \quad (2)$$

この近似の過程で短時間ステップ応答として最適化を行った例を図10に、そして評価尺度である $\zeta$ と $\omega$ に対する $F(\zeta, \omega)$ の値を図11に示す。

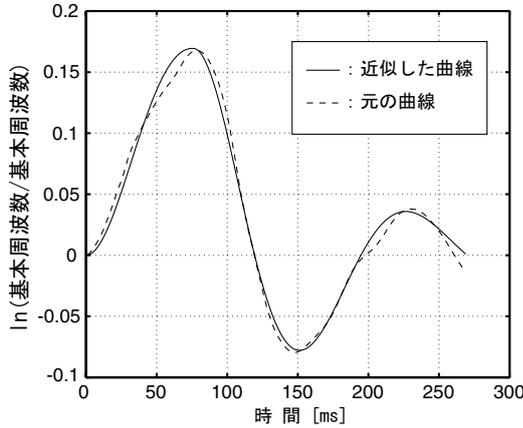


図10 実際の $F_0$ 装飾音成分と近似した短時間ステップ応答

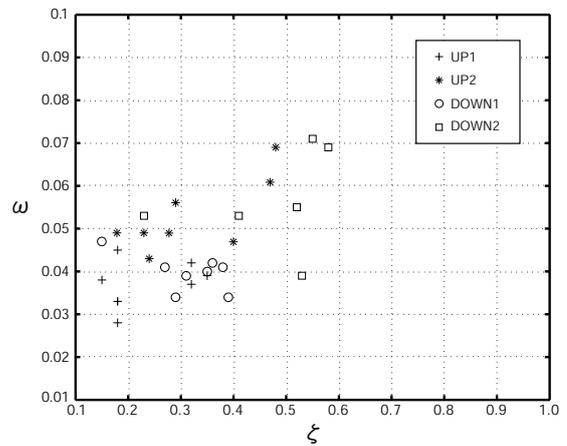


図12 抽出した装飾音「ふり」の $\zeta$ と $\omega$

表1 「ふり」のパラメータ

	$\zeta$	$\omega$	位置 [ms]	大きさ [cent]
UP1	0.23	0.036	-389	229
UP2	0.32	0.052	37	128
DOWN1	0.31	0.039	-250	287
DOWN2	0.45	0.056	46	-112

表2 「あたり」パラメータ

$\zeta$	$\omega$	位置 [ms]	大きさ [cent]
0.35	0.037	7	348

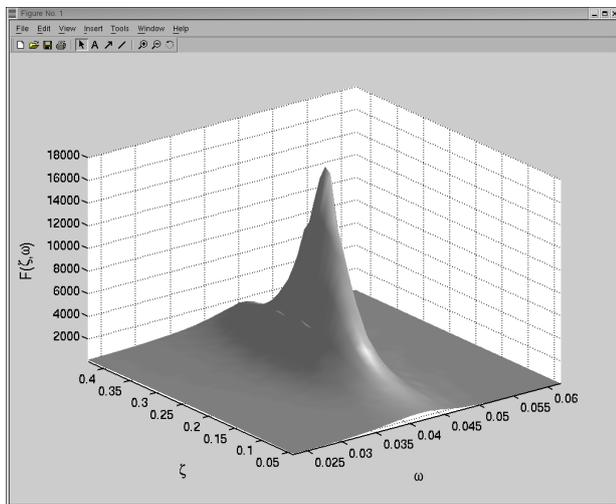


図11  $\zeta$ と $\omega$ に対する $F(\zeta, \omega)$ の値

この操作を全ての音程変化部分に対して行い、得られた振動成分を2次系のステップ応答として近似した結果、伝達関数パラメータ $\zeta$ と $\omega$ に関して図12に示す分布が得られた。ここで、UP1, UP2はそれぞれ音程上昇部分に見られる1つ目, 2つ目のステップを指し、DOWN1, DOWN2はそれぞれ音程下降部分に見られる1つ目, 2つ目のステップを指す。

この分布の平均値として、今回用いた「ふり」の $F_0$ 装飾音指令のパラメータを表1に示す。また、同様の作業から得た装飾音「あたり」のパラメータを表2に示す。「位置」は装飾音が付随している音程遷移指令ステップの立ち上がり時間との差分を、「大きさ」は音程遷移指令に対しての値を表している。

### 3.1.3 モデルを用いた基本周波数遷移の生成

基本周波数遷移を図8のモデルに従い生成した。音程遷移成分に関しては、音程上昇時のパラメータとして $\omega = 0.035$ ,  $\zeta = 0.55$ , 音程下降時のパラメータとして $\omega = 0.030$ ,  $\zeta = 0.55$ を用いて生成し[6], ビブラートに対応する微細振動成分に関し

ては、毎秒6.6回振動する正弦波と白色雑音を足し合わせたものをローパスフィルタに通し、全体として $\pm 48$  cent程度のゆれを持たせることで近似した[7]。最後に今回抽出した $F_0$ 装飾音指令パラメータから生成した $F_0$ 装飾音成分を足し合わせた。図13と図14に、それぞれType1とType2に対して提案モデルから生成した長唄の基本周波数遷移を示す。

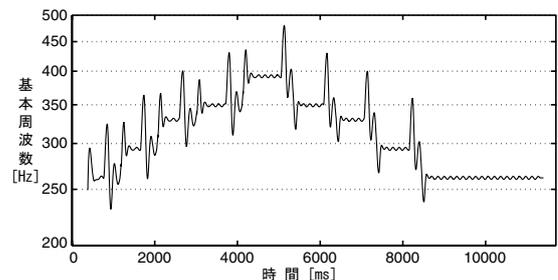


図13 モデルから生成したType1の基本周波数遷移

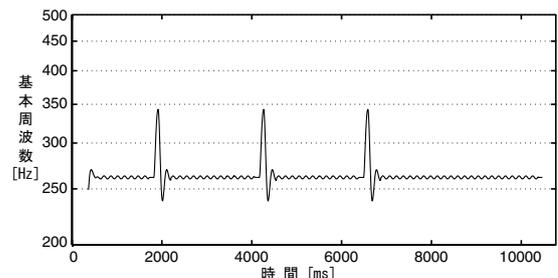


図14 モデルから生成したType3の基本周波数遷移

## 3.2 パワー遷移のモデル化

### 3.2.1 モデルの概要

パワーの遷移に関しては、生理学的な裏付けのある生成モデ

ルは得られていない。ここでは基本周波数のモデル化に際しても用いた2次系の過渡応答によってパワーの急激な低下を表現することを試みた。具体的には、時間長20ms程度の短時間ステップ状の指令（パワー装飾音指令）から突発的の下降のパワー装飾音成分を生成し、発声区間全体にかけてのなだらかなステップ応答である音源強度変化成分に加えることによって長唄のパワー遷移を生成する。以下にその概要を示す。

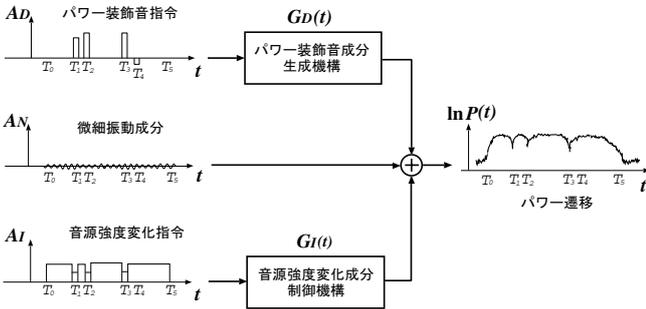


図15 パワー遷移生成モデルの概要

なお、この音源強度変化成分については、文献[8]にて提案されている音源強度変化パターンの生成過程モデルを参考とした。このモデルは式3と式4によって定式化されており、複数のステップ指令から音源強度変化成分  $\ln I(t)$  を生成している。

$$\ln I(t) = \sum_{l=1}^L A_{Il} \{G_I(t - T_{1l}) - G_I(t - T_{2l})\} \quad (3)$$

$$G_I(t) = \begin{cases} 1 - (1 + \lambda t) \exp(-\lambda t) & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (4)$$

### 3.2.2 モデルを用いたパワー遷移の生成

上記のモデルによって生成したパワー遷移を用いて評価実験用音声を合成した。モデルのパラメータの値を種々に変えて、装飾音部分のパワー変動が最も図6に近くなるようにした。その結果、パワー装飾音指令については音程の上昇時と下降時共に  $\omega = 0.045, \zeta = 1.2$  を、式4で用いられる音源強度変化指令のパラメータについては  $\lambda = 0.012$  とした。また、基本周波数遷移の生成に関してはこれまでと同じパラメータを用いて図8のモデルから生成した。図16にType1に対応させて生成した長唄のパワー遷移(上)と基本周波数遷移(下)を示す。

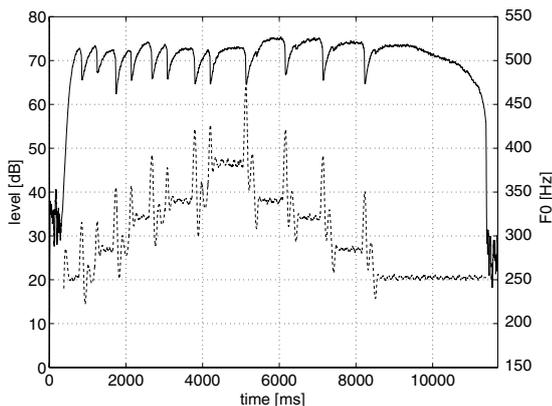


図16 生成したパワーと基本周波数の遷移

## 4. モデルの評価

### 4.1 基本周波数遷移生成モデルの評価

最終的なモデルの評価となる聴取実験は、音声データの収録と同様、簡易防音室内にてD/A内臓のUSBヘッドフォンを用いて行った。その際に使用する合成音については一定の音程で/a/の母音を長時間保っただけの音声を用意し、STRAIGHTを用いて基本周波数を入れ替えることで生成した。合成元の音声は長唄歌手とは異なる男性話者による発声であるため、評価の尺度はこの音声の高さ(C3)との相性から、Type1, Type2, Type3のそれぞれ1オクターブ下の音程遷移について、「今回の提案モデルから生成した基本周波数遷移と入れ替えて合成した歌声」と「歌い手の長唄から直接抽出した基本周波数遷移と入れ替えて合成した歌声」を比較し、どちらの音程変化があらかじめ聴取させておいた「実際の長唄音声」の持つ独特な音程遷移に近いと感じるかを5段階にて評価した。共に3を基準にし「今回のモデルから作成した合成音の方が元の長唄に近い」と感じれば1を、「実際の長唄の基本周波数から作成した合成音の方が元の長唄に近い」と感じれば5を選択させた。長唄に関する専門知識をもたない10名を対象に行った結果を図17に示す。

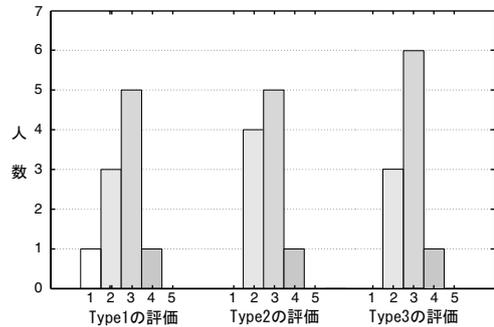


図17 基本周波数遷移生成モデルの評価結果

この結果では、「歌い手から抽出した基本周波数遷移を用いた合成音よりも、今回の提案モデルから生成した基本周波数遷移を用いた合成音の方が、より歌い手の長唄に近い音程変化である」という、矛盾にも取れる評価がなされた。これに関する歌い手の意見として以下を得ることが出来た。

「長唄において装飾音はあくまで情緒や感情を表現するために付加するもので、その状況により装飾音の程度は様々に変化するものである。しかし今回の聴取実験のように情緒や感情などが全く無い場合には、装飾音を色々変化させて味を付けた場合より、単に一貫して装飾音を大きく付けた方が長唄の特徴をより捉えていると解釈されることもある。」

確かに、図13や図14のように本モデルを用いて生成した長唄の基本周波数遷移は、図2や図4で示した実際の長唄の基本周波数遷移に比べ、やや人工的な単調さは感じられる。実際、歌い手に同様の聴取実験を行ったところ「モデルから生成した基本周波数遷移を用いて合成した長唄音声は、機械的で不自然である」という回答が得られた。しかし、長唄に関する専門知識を持たないものにとっては、音程が変わるところに付加される長唄独特の遷移が逆に安定している場合の方が「良い長唄」と解釈する傾向があり、そのことがモデルから生成した基本周

波数推移を用いた合成音を選ぶ結果につながったと考えられる。とは言え、全体を通じて「モデルから生成した基本周波数推移を用いた再合成音声と、実際の長唄から抽出した基本周波数推移を用いた再合成音声の差は感じられない」という結果が一番多かったと言うことは、今回提案したモデルやそのパラメータが、比較的妥当性が高かったことを表していると解釈できる。

#### 4.2 パワー遷移生成モデルの評価

合成音の生成については4.1と同様に、一定の音程で母音 /a/ を長時間発音した音声を用意し、この音声を元に STRAIGHT を用いて基本周波数とパワーを入れ替えることで生成した。評価尺度は Type1 と Type2 のそれぞれの音程遷移について、まず「実際の長唄音声から抽出した基本周波数推移と入れ替えた音声」を合成し、この音声に対して、「パワー遷移を平坦にして再合成した歌声 (sd1)」と「モデルから生成したパワー遷移と入れ替えて再合成した歌声 (sd2)」, 「歌い手の長唄から直接抽出したパワー遷移と入れ替えて再合成した歌声 (sd3)」の3種類を比較し、どの合成音に付加した装飾音があらかじめ聴取させておいた「実際の長唄音声」の装飾音の特徴に近いと感じるかを3択にて評価するものとした。なお、事前に何も予備知識を与えずこの実験を行ったところ、本来注目されるべきパワー変動は軽視される傾向が見られたため、インストラクションとして、本来の長唄には装飾音によるパワー変動が見られることを明示して実験を行った。長唄に関する専門知識をもなない10名を対象に、簡易防音室にてD/A内臓のUSBヘッドフォンを用いて行った実験結果を図18に示す。

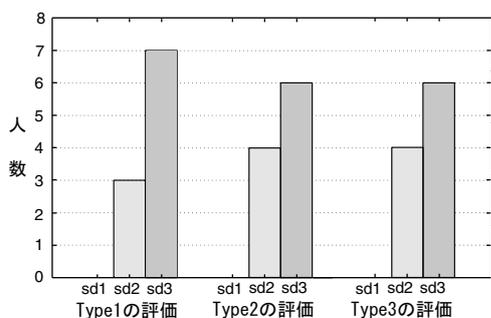


図18 パワー遷移生成モデルの評価結果

図18に示した評価結果から、「モデルから生成したパワー遷移が、最も元の長唄の装飾音の特徴に近いと感じた」という回答が4割程度を占めており、今回定義したパワー遷移の生成モデルは近似の程度こそまだ十分とは言えないものの、ある程度の特徴は表現できたとすることが出来る。

また参考実験として、基本周波数とパワーのどちらの要素が長唄らしさに強く起因するかを調べる聴取実験も行った。同じ歌い手による「コーラススタイルでの歌声」に対し、「長唄の基本周波数変動を付加して再合成した歌声」と「長唄のパワー変動を付加して再合成した歌声」を比較してどちらにより長唄らしさを感じたかを判定させた。その結果、被験者全員から『「長唄の基本周波数変動を付加して再合成した音声」により長唄らしさを感じた』という回答が得られた。このことから、我々は実際にはパワーの変動より基本周波数の変動に長唄らしさを感じていると解釈できる。

なお、パワーにおける変動成分の有無について実際の歌い手に意見を聞いてみたところ、「長唄としての上手下手とパワー変

動の有無については、一概にはどちらが上手下手と言えないが、個人的にはsd2やsd3(パワー変動が有る音声)よりもsd1(パワー変動が無い音)の方が好きだ」という回答が得られた。この判断に関しては、長唄に長く関わっている者でも「どちらが良い」とまではなかなか言えないようである。

#### 5. まとめ

今回の研究では今まであまり考えられることの無かった歌唱法に応じたモデルの必要性を述べ、日本の伝統的歌唱法である長唄を対象とした韻律遷移の生成についてのモデル化を行った。

基本周波数遷移生成のモデル化については、長唄に見られる独特の突発的振動を短時間ステップ「 $F_0$ 装飾音指令」に対する2次系の応答として近似し、ベースとなるステップ指令「音程遷移指令」に対する応答と「ビブラート成分」に付加することで長唄の基本周波数遷移を生成した。また、聴取実験によって提案したモデルについての妥当性を示した。パワー遷移生成のモデル化についても、長唄に見られる独特の突発的下降を短時間ステップ「パワー装飾音指令」に対する2次系の応答として近似し、ベースとなるステップ指令「音源強度変化指令」に対する応答と「微細振動成分」に付加することで長唄のパワー遷移を生成した。そして、聴取実験によって提案したモデルについての妥当性を示した。

これにより、情緒や感情を持たせない(歌詞を付けない)場合の長唄的韻律遷移の生成については、ある程度の特徴を引き出すことに成功したと言える。今後は情緒性を持たせた場合の韻律遷移の特徴や、今回提案したモデルについての統一性、スペクトルなど分節の特徴についても長唄らしさを抽出していくことが求められる。

#### 文献

- [1] 中山 一郎, 小林 範子, “歌の声 — 声質の魅力と問題点 —,” 音響学会誌, vol. 52, no. 5, pp. 383–388, 1996
- [2] 中山 一郎, 天野 文雄, 上畠 力, 河内 厚郎, 小島 美子, 小林 範子, 杉藤 美代子, 高木 浩志, 柳田 益造, “日本の「声の音楽」の諸相 — 共通の歌詞を用いた邦・洋楽の歌唱表現法の比較の試み —,” 信学技報, SP97-114, pp. 47–54, 1998
- [3] Hideki KAWAHARA, “Speech Representation and Transformation Using Adaptive Interpolation of Weighted Spectrum: Vocoder Revisited,” Proc. ICASSP’97, pp. 1303–1306, 1997
- [4] 小林 範子, 東倉 洋一, 天白 成一, 新美 成二, “日本音楽の歌唱における生成的特徴,” 音講論集, 2-3-11, pp. 245–246, 1988
- [5] Kenzo Itoh, “Correlation Analysis Between Speech Power and Pitch Frequency for Twenty Spoken Languages,” Proc. ICSLP’94, pp. 331–334, 1994
- [6] 斎藤 毅, 鶴木 祐史, 赤木 正人, “歌声における  $F_0$  動的変動成分の抽出と  $F_0$  制御モデル,” 研究会資料, H-2001-93, pp. 683–690, 2001
- [7] ダイアナ・ドイチュ, “音楽の心理学 (上),” 西村書店, 1987
- [8] 大野 澄雄, “韻律の表現における声帯音源強度パターンと基本周波数パターンの関係 — 生成過程のモデル化に基づくアプローチ —,” Proceedings for 2001 2nd Plenary Meeting and Symposium on Prosody and Speech Processing, pp. 1–4, Jan 2002