

# 実世界情報システムプロジェクト

## 視聴覚研究グループ峯松研究室 伝統歌唱芸能研究ユニット

峯松 信明

情報理工学系研究科電子情報学専攻

### 概要

本研究では、日本の伝統歌唱芸能の一つである「長唄」に着眼し、音声工学的にその歌唱方法を分析・モデル化する。「長唄」の特徴は、頻繁に観測される突発的な音程遷移（装飾音）とそれに付随する種々の音声現象である。本研究では「長唄」の韻律特徴に着目し、最終的には、五線譜表記された音程遷移からそれを「長唄」らしく唄いあげるための韻律パターンの自動生成を検討した。

### 1 はじめに

人と人が集まり、共存することで社会が生まれ、その社会が持続することで、固有の規則、習慣、文化が生まれる。当然我が国にも固有の文化がある。多くの場合文化の作り手は「若者」であることが多く、この場合、新しいもの、初めて見る・聞くもの、がその発端となることが多い。しかし、世の中を知り、社会を知り、世界を知るとなると、本当の意味で他者と自分とを区別するアイデンティティの拠り所が「日本古来の」ものの考え方・感じ方、であることに気付くことが多いのも事実である。本研究ユニットでは歌唱芸能を対象にするが、上記した、本当の意味での我々のアイデンティティの拠り所の一つとなる伝統歌唱芸能に着眼し、その歌唱を音声工学的に分析・モデル化することで、その文化を保存すると共に、その文化の新たな側面の発掘、更には、「日本の心」をより深く理解することを目的とする。

従来歌唱を工学的な分析対象とする場合、西洋

的（イタリア）歌唱法であるベル・カントを題材とすることが多い。これはベル・カントが全ての歌唱の基盤であることを意味するものではない。例えばベル・カントに沿って「日本語で書かれた歌詞」を歌う場合「音の響きを重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」というのは誰しも経験があるであろう。一方日本の歌唱法は一般的に、情緒や雰囲気重視するものが多く、本研究で題材とする長唄もその流れを汲む。但し、分析、モデル化に際してはベル・カント法に対して行なわれた先行研究を基盤とし、同一の「歌い手」と「楽譜」によるベル・カントと長唄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次系の韻律制御モデルを提案する。評価実験では、モデルから生成した長唄的韻律遷移を用いた再合成音声、実際の長唄から抽出した韻律遷移を用いた再合成音声などを比較し、提案モデルの妥当性を検証する。

### 2 歌唱法による韻律遷移の違い

#### 2.1 歌声資料の録音

長唄には、ベル・カントには見られない韻律的な特徴がいくつか存在する。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る。これらは「装飾音」と呼ばれ、その表現方法は歌い手に任されているため、楽譜に書かれたタイミングで正しい音程に沿って歌う

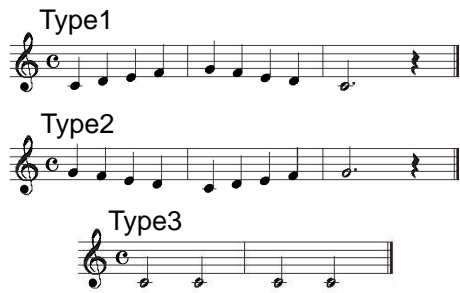


図 1: 3 種類の楽譜

ベル・カント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、同じ音程が続く場合に付加される「あたり」についての分析を行う。図 1 に示す楽譜を用いて歌声資料の録音を行なった。歌い手は長唄を 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。

## 2.2 $F_0$ パターンとパワーパターン

収録した各歌声から  $F_0$  を抽出した。図 2 に Type1 のパターン（上：コーラス調，下：長唄調）を示す。この突発性の変動が「ふり」である。図 3 は Type3 のものであるが、ここに観測されている突発性の変動が「あたり」である。

図 4 に Type1 におけるパワーパターンを波形、 $F_0$  パターンと同時に示す。上がコーラス調，下が長唄調である。一般に  $F_0$  とパワーは正の相関を持つことが知られているが、長唄の場合、急激な  $F_0$  の上昇に伴うパワーの急激な“下降”が観測されている。この背景には、人間の生理的特性があると考えられるが、詳細は分っていない。

## 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  $F_0$  モデリングに関しては、式 1 に示す 2 次系の伝達関数を用いて、音程（ベースとなる  $F_0$  からの差分）を大きさとしたステップ指令に対す

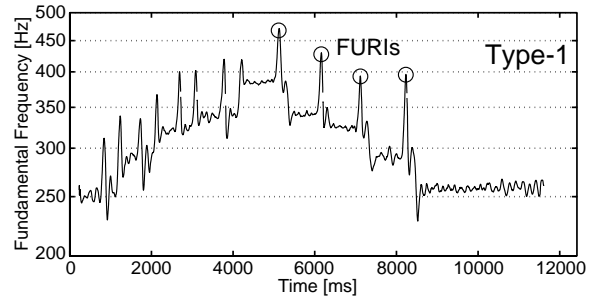
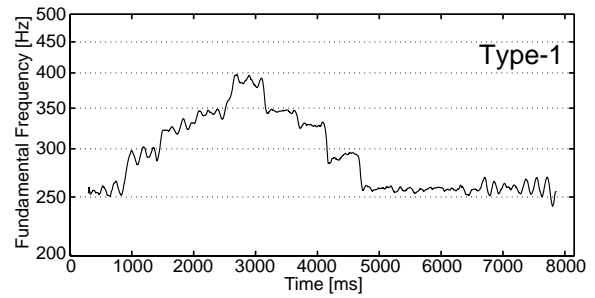


図 2: Type1 の  $F_0$  パターン

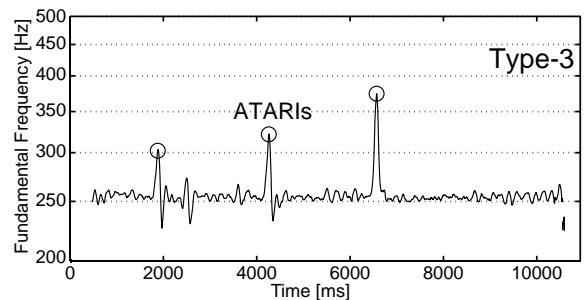
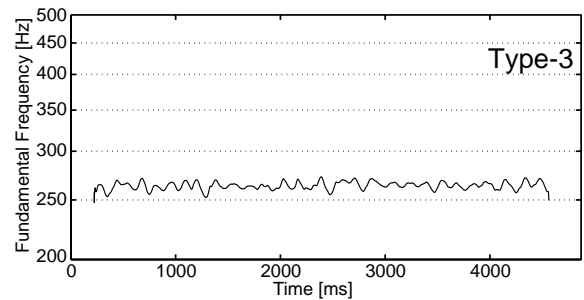


図 3: Type3 の  $F_0$  パターン

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。

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

提案するモデルは、図 5 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  $F_0$  遷移を生成しようとするものである。

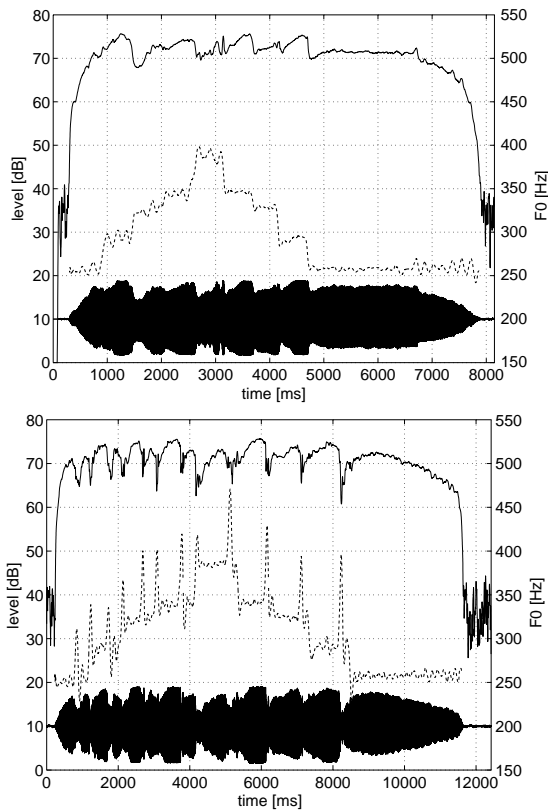


図 4: Type1 のパワーパターン

### 3.1.2 $F_0$ 装飾音指令のパラメータ抽出

図5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を行った。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6に示す。視察により得られる情報を元に、パラメータ推定は基本的に最小二乗誤差となるパラメータを、初期値を適切に定め、実験的に探索することで求めた。

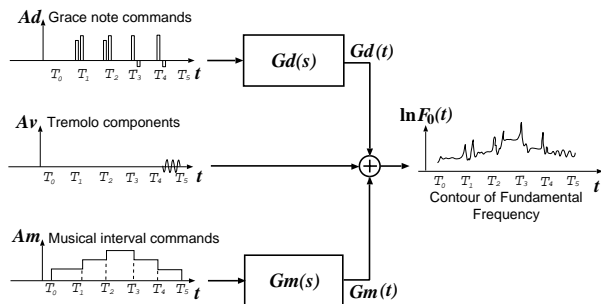


図 5: 基本周波数遷移生成モデルの概要

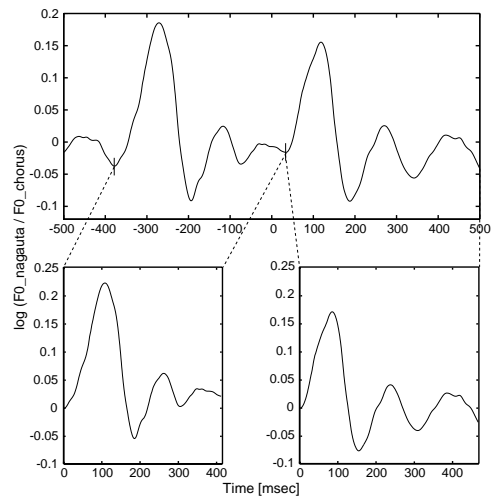


図 6: 対数軸上での長唄とコーラスの周波数遷移差

### 3.1.3 モデルを用いた韻律パターンの生成

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、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  $\pm 48\text{cent}$  程度のゆれを持たせることで近似した。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成分を足し合わせた。図7と図8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  $F_0$  遷移を示す。

なお、パワー遷移についても  $F_0$  と同様、2次系の伝達関数で表現し、急激なパワー降下も短時間ステップ応答として近似した。図9にモデルより生成したパワーと  $F_0$  遷移を合わせて示す。

## 4 モデルの評価

### 4.1 基本周波数遷移生成モデルの評価

一定の音程で /a/ の母音を長時間保つただけの音声を用意し、基本周波数を入れ替えた再合成音声を作成して評価実験を行なった。Type1~3の音程遷移について、「提案モデルから生成した  $F_0$  パターン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  $F_0$  パターンと入れ替え

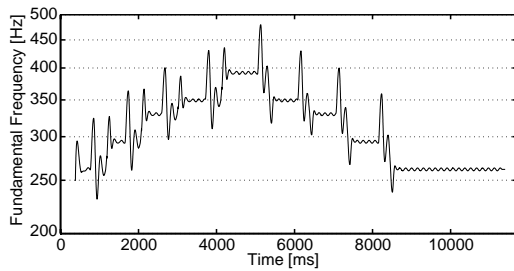


図 7: モデルから生成した Type1 の  $F_0$  遷移

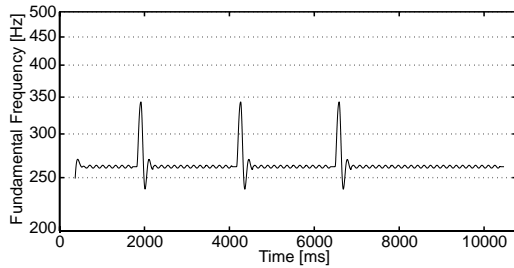


図 8: モデルから生成した Type3 の  $F_0$  遷移

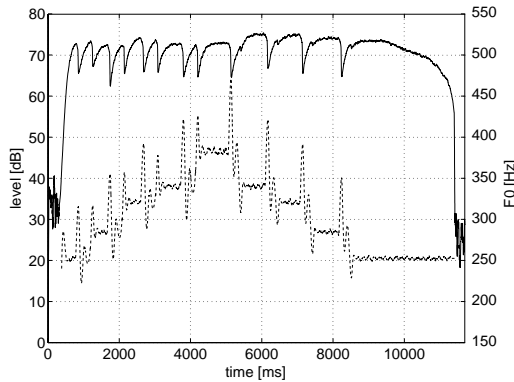


図 9: モデル生成したパワーと  $F_0$  遷移

て合成した歌声」を比較し、どちらが予め聴取させた「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段階で評価させた。なお、数字が低い方がモデルベースの韻律パターンを支持していることを示す。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10に示す。

この結果は「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、より長唄に近い音程変化である」という矛盾とも言える結果となっているが、これに対して、長唄歌手から以下のコメントを得ることができた。

「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

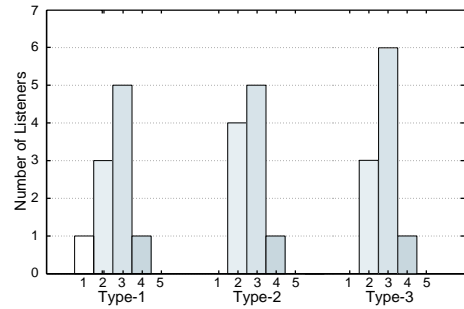


図 10: 基本周波数遷移生成モデルの評価結果

飾音の程度は様々に変化するものである。しかし今回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。」

パワー遷移モデルについても同様の実験を行った。パワー遷移に関しては実際の歌手の遷移に基づくパターンが支持されたが、近似精度としては十分なモデルが構築できたと考えている。

## 5 まとめ

日本の伝統的歌唱法である長唄を対象としてその韻律モデリングと評価を行ない、良好な近似精度を得た。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、その独特な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的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。

## 参考文献

- [1] 松岡文吾, 峯松信明, 広瀬啓吉, “長唄における音程遷移のモデル化について”, 音響学会秋季講演論文集, 3-10-3, pp.333-334 (2002)
- [2] 松岡文吾, 峯松信明, 広瀬啓吉, “長唄における韻律的特徴の分析とそのモデル化”, 音響学会春季講演論文集, 2-6-6, pp.309-310 (2003)
- [3] 松岡文吾, 峯松信明, 広瀬啓吉, “長唄における韻律的特徴のモデル化とその評価”, 電子情報通信学会音声研究会, SP2002-183, pp.25-30 (2003)