

# 日本民謡の大規模音楽コーパスを用いた旋律の構造抽出

著者	河瀬 彰宏
雑誌名	国立国語研究所論集
号	7
ページ	121-150
発行年	2014-05
URL	<a href="http://doi.org/10.15084/00000528">http://doi.org/10.15084/00000528</a>

## 日本民謡の大規模音楽コーパスを用いた旋律の構造抽出

河瀬 彰宏

国立国語研究所 コーパス開発センター 非常勤研究員

### 要旨

本論文では、(1) 日本民謡の音楽的特徴——旋律に内在する法則——を科学的に捉え、(2) 抽出した特徴に基づき日本民謡の地域性を客観的に判断する指標を示す。はじめに『日本民謡大観』(1944–1993) に収録されている種目のうち全国的に網羅的に存在する日本民謡 1,794 曲と、Web 上に公開されている大規模音楽データベースに収録されている中国民謡 1,984 曲から、それぞれ音楽コーパスを構築する。日本民謡の比較対象として、中国民謡を用いる理由は、中国音楽が日本音楽の形成に多大な影響を与えてきたにもかかわらず、音楽文化や旋律のもつ雰囲気などの点で違いが見受けられるためである。分析の手順としては、各コーパスに対して VLMC モデルを用い、旋律中に繰り返し出現する音程推移パターンを抽出する。そして地域・種目などの背景要因に応じて日本民謡の楽曲データをグループに分割し、計量的な手法を用いてその特徴を比較検討する。その結果、日本音楽の特徴は、民俗学や歴史学において提唱されている日本列島の東西二分論（社会組織論）のほか、方言研究やアクセントの分布図とも一致することがわかった。

**キーワード：**音楽コーパス、計量分析、可変長マルコフ連鎖モデル、音組織

### 1. はじめに

#### 1.1 音楽コーパスを用いた計量分析

近年、デジタルコンテンツの普及に伴い、音楽に関連したインタフェースの需要が高まっている。とくに、人はどのように音楽を認知しているのか、その仕組み——人の音楽認知メカニズム——を把握し、インタフェースを実現する重要性が社会的にも学術的にも求められている。しかしながら、なぜ人は音楽を聴くと様々な印象を受けたり、旋律の類似性を認識できたりするのか。なぜ人は異文化の音楽を受け容れたり、受け容れにくかったりするのか。人と音楽に関するあらゆる疑問は、未だに解決されていないものが多い (Temperley 2001)。これらの問題を解明するにあたり、従来の音楽学や心理学の方法論では、最終的な解釈や判断が分析者の主観に委ねられることがあるため、旋律の特徴を精確に捉えられず限界があった。

近年は音楽研究の研究理念として、異分野の方法論を横断的に融合しながら問題解決に取り組む傾向が高まっている (Purwins et al. 2008)。例えば、2007 年に欧州委員会が立ち上げた S2S2 Coordination Action や、その後続プロジェクト Sound and Music Computing (以下、SMC)<sup>1</sup> がある。SMC では、人の音楽認知メカニズムを解明するために、既存の音楽学や心理学の方法論だけでなく、音響物理学、情報工学、電子工学、計算機科学、神経脳科学、言語学など、音声・音響と密接な関係にある諸分野の方法論を積極的に融合させて問題解決にあたることをアジェンダに掲げている。

<sup>1</sup> <http://smcnetwork.org/> (2010 年発表。2013 年 12 月 12 日参照)

音楽学では、旋律に内在する法則のことを音組織（おんそしき, Tonesystem）という用語で表す。この抽象的な概念は、Sachs (1943) の古典研究に基づき「ある音楽の音の成立, 楽音の高さと間隔」と定義でき、そこから転じて「ある音楽の時間的要素を除いた体系」「ある音楽における音の相互関係」といった意味をもつ（藤井 1978）。

音楽には、音組織以外にも拍子構造, 対位法の構造, 和声進行, 調性など, 重要な様相をいくつも挙げることができる。しかし、それらの中で音組織がとくに重視される理由は、人が音楽を聴取する際に旋律をフレーズに分割したり, 結合したりする作用に関わることが実証されているためである（Snyder 2000）。

例えば、音楽情報処理（Music Information Retrieval）の分野では、旋律をフレーズに分割し、旋律のクエリー検索（Query-by-Singing/Humming, QBSH）システム（McNab 1996, Cao et al. 2009）に応用され、音楽ジャンルに特化した大規模音楽コーパスが公開されている。Essen Folksong Collection（Schaffrath 1995）、“The Dictionary of Musical Themes”<sup>2</sup>、“Répertoire International des Sources Musicales Collection”（RISM Collection）<sup>3</sup>などが代表的な音楽コーパスとして挙げられる。そして、大規模音楽コーパスの構築と並行して、旋律のフレーズ分割やグルーピング（Lerdahl and Jackendoff 1983）を計算機上で自動処理するアルゴリズムが提案されてきた背景がある（Cambouropoulos 2001）。

本研究では、人の音楽認知メカニズムを解明することを究極の目標とし、これを達成するための基礎研究として、(1) 日本民謡の音組織——旋律に内在する法則——を科学的に捉え、その上で(2) 日本民謡の地域性を客観的に判断する指標を示す。ある文化の音組織を精確に捉えることは、その文化の音楽を形成する音の相互関係、伝播と変容、普遍性の解明につながり、ひいては人の音楽認知メカニズムの解明につながると考えられるためである。具体的には、次の3点を実施する：

1. 日本民謡の楽曲データを集めて音楽コーパスを構築し、旋律中に繰り返し出現するパターンを確率・統計的に抽出する。
2. 古来、日本音楽に影響を及ぼした中国音楽からも同様にパターンを抽出し、日本民謡のパターンと比較することで、日本民謡に固有の特徴（音組織）を決定する。
3. 決定した音組織に基づき、日本民謡の地域性を統計的に明らかにする。

## 1.2 先行研究と本研究の意義

音楽学や心理学における楽曲分析では、とりわけ近代西洋音楽における特定の作品の聴取、音楽体験、解釈が主な対象である。しかし、非西洋音楽においては、後述する『日本民謡大観』から音組織の抽出を試みた Kawase and Tokosumi (2008, 2010a) や、日本民謡と中国民謡との比較を実施した Kawase and Tokosumi (2010b)、明治維新以降から戦後までの日本音楽を対象とした

<sup>2</sup> <http://www.multimedialibrary.com/barlow/index.asp> (2008年発表。2013年12月12日参照)

<sup>3</sup> <http://www.rism.info/> (2006年発表。2013年12月12日参照)

高際ほか(2013)などの僅かな事例しかなく、音楽人類学の視点まで踏み込んだ分析はほとんど成し遂げられていない。また、抽出した結果と文化的な因果関係を検証した研究は将来の課題とされてきた。

本論文では、日本民謡や中国民謡を対象を絞り、国際比較を計量的に行っている点、民謡の楽曲データの計量分析を確立している点で先駆的である。これは客観性の担保が困難な従来の人文的手法とは対極に位置するものであり、より客観的な分析が可能となるだけでなく、計量的見地から民謡の音組織に迫ることが可能である。本論文で確立した方法論は、計算機による非西洋音楽の分析や発見・支援の発展に寄与するものであり、究極的には、上述の人の音楽認知メカニズムの解明に迫る基礎研究である。

### 1.3 日本の伝統音楽

日本の伝統音楽は、奈良時代から平安時代にかけて仏教や儒学とともに中国から輸入した音楽文化を背景にもち、それを独自に昇華したものである。大きく芸術音楽、仏教音楽、民俗音楽の3つに分類でき、一般に、芸術音楽(雅楽や能楽)と仏教音楽(声明)は、明確な演奏様式・流儀を追究する姿勢をとるため、伝統音楽の主流と見なされる。一方、民俗音楽(わらべうた、民謡、民俗芸能)は、主流から外れているものの、伝統音楽の本質を十分に備えていると考えられている。

音楽学者の樋口昭の言葉を借りれば、わらべうたは、遊びの中で創作・伝承した子供たちの音楽であり、民謡は、歌うことを主眼として個人差、地域差、世代差を認めた大人たちの音楽である。そして民俗芸能は、祭礼で神仏へ奉納する形式で地域社会の中に伝承された音楽である(小島1982)。つまり、民俗音楽の旋律は、昔の人々の自然なままの感情・情緒の表れであり、プリミティブな特徴が継承されている。日本の民俗音楽は、芸術音楽と仏教音楽が時代の経過とともに失った特徴を大いにもつ、世界的に特異な音楽である。

さらに、日本の伝統音楽を器楽と声楽に大別した場合、器楽の種目は、雅楽の管絃、箏曲の段物、尺八楽、歌舞伎の下座音楽など、僅かしかなく、少なく見積もっても声楽が伝統音楽全体の約85%を占める。このことから、日本の伝統音楽は器楽よりも声楽を主体とした音楽といえる。

したがって、民俗音楽を構成するジャンルのうち声楽に特化した民謡の旋律には、日本音楽の変遷過程や普遍的な要素を見出すヒントが埋められていると考えられる。以下では、日本民謡<sup>4</sup>の旋律を分析し、その背後にある法則を取り出すことで、日本音楽の音組織を捉える。

### 1.4 全体の構成

本論文の構成は次の通りである：2節では、分析に用いる楽曲データを概説する。3節と4節

<sup>4</sup>「民謡」とは、ドイツ語の *Volkslied* の訳語として明治期に日本国内へ持ち込まれた用語である。学術的な定義では、自然性・伝承性・移動性・集団性・素朴性・郷土性の6条件を備えた音楽を指す。平たくいえば、無名のアマチュアたちが口ずさんだところから始まり、個人差や地域差を認めつつ、代々歌い継がれてきた音楽のことである。

では、音楽コーパスの旋律中に繰り返し出現するパターンをどのように抽出するか、青森県の子守唄を事例に挙げて説明する。5節と6節では、Kawase and Tokosumi (2010b)を基礎に、音楽コーパスから抽出したパターンから音組織を捉える。7節から10節は、日本民謡のデータを地域ごとに分け、5節で抽出したパターンが地域ごとにどのように異なるのかを検証する。ここではKawase and Tokosumi (2010a)を概説し、音楽学者の小泉文夫が述べたように、民謡を文芸学、民俗学、音楽学の3つの側面から総合的に把握する立場(小泉1958)から議論を展開する。最後に11節では、全体を総括する。

## 2. 分析データの概要

### 2.1 日本民謡の音楽コーパス

『日本民謡大観』(以下、『大観』)に収録された楽譜資料を電子データ化して、日本民謡の音楽コーパスを構築する。『大観』は、音楽学者の町田佳聲が日本放送協会(NHK)の協力のもと、1944年から半世紀にわたって津々浦々の民謡を調査・記録した資料である。日本列島の地域ごとに刊行されており、全9巻から構成される。

これまでに、用途や目的によっていくつもの日本民謡の楽譜集が出版されているが、『大観』は全国的な規模で民謡を採録しており、高度な知識をもった採譜者や校正者が資料を丹念に選別している点で、他の楽譜集と一線を画する。そのため、『大観』は、計量分析を試みる上で、質と量を兼ね備えた、現状で最も信頼のおける日本民謡の資料といえる(小島1992)。

本研究では、『大観』に採録された楽曲のうち、全国的な規模で掲載されている上位5種目1,794曲——盆踊唄546曲、田植唄383曲、地形唄279曲、田草取唄200曲、子守唄139曲、および、それらの変型247曲——を電子データ化・集積し、日本民謡の音楽コーパスを構築する。ただし、北海道の楽曲は曲数が圧倒的に少なく、沖縄県の楽曲は本土とあまりにも異なり、分類実験の際にこれらの特徴は全体に影響が及ぶため、今回の分析対象から除外した。

### 2.2 中国民謡の音楽コーパス

日本民謡の比較対象として、スタンフォード大学音楽学部CCARH(Center for Computer Assisted Research in the Humanities)が所有する楽譜資料*KernScores*<sup>5</sup>を電子データ化して、中国民謡の音楽コーパスを構築する。*KernScores*は、CCARHが楽曲データの学術利用を目的に構築している大規模音楽データベースの一部であり、J.S. Bach, W.A. Mozart, L. v. Beethovenなどの近代西洋音楽の楽曲が大半を占める。僅かにS. JoplinやG. Gershwinらのジャズの古典や、“Essen Folksong Collection”(以下、『EFSC』)の楽曲も含まれている<sup>6</sup>。

本研究では、*KernScores*に採録されている楽曲データのうち、『EFSC』の中国民謡1,984曲を

<sup>5</sup> <http://www.musiccog.ohio-state.edu/Humdrum/> (2001年発表。2013年12月12日参照)

<sup>6</sup> 『EFSC』の由来については不明な点が多く、*KernScores*のメタデータにも詳細が記述されていない。CCARHのディレクターを務めるE. Selfridge-Fieldによれば、『EFSC』の多くは楽譜として出版されたことがあるという(Temperley 2007)。

電子データ化・集積し、中国民謡の音楽コーパスを構築する。ただし、『EFSC』のメタデータには採譜された中国国内の地域が明確でない楽曲が含まれるため、中国本土の地域差をみる上で必ずしも十分なデータとはいえない。したがって、5.2節以降の中国民謡の分析結果は、あくまでも中国民謡に平均的なパターンであることを予め断わっておく。

### 2.3 音楽コーパスの基本統計量

図1は、1曲に含まれる音符の出現回数の分布、図2は、音価（音長）を考慮した音符の持続時間の分布である。ただし、図2では四分音符を48 tick という長さに統一して集計している<sup>7</sup>。

図1の平均値 $\mu$ は、日本民謡112.78音、中国民謡62.84音であり、変動係数 $C.V.$ は、日本民謡7.61、中国民謡6.85であった。また、図2の平均値 $\mu$ は、日本民謡6485.67 tick、中国民謡4684.28 tickであり、変動係数 $C.V.$ は、日本民謡7.97、中国民謡6.81であった。つまり、日本民謡よりも中国民謡の方が1曲あたりの発音数は少なく、歌唱時間も短い傾向があること、分布のばらつき具合にほとんど違いがないことがわかる<sup>8</sup>。日本民謡の音楽コーパスのデータのうち、音符の出現回

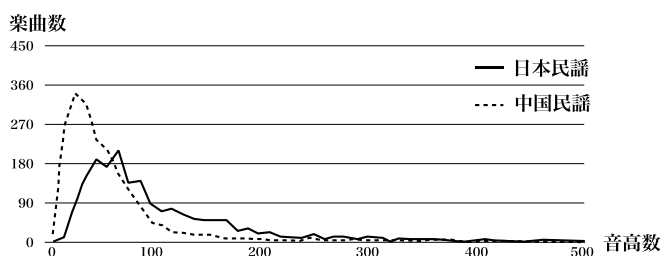


図1 1曲あたりの音符の出現音数

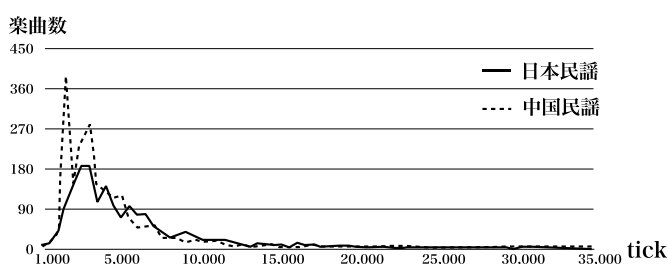


図2 1曲あたりの音符の持続時間

<sup>7</sup> “tick”はMIDIの用語で1拍を分割する分解能の単位である。したがって、ここでは四分音符1拍の長さを48分割した長さを基本単位としている。また、日本民謡は1曲全体を通してテンポが一定ではなく、拍節も固定されていないため、現代のポップスのように1曲を1つのテンポで記述することが難しい。『大観』には、テンポが具体的に記録されている場合もあるが、あくまでも採譜者の判断による大まかな目安でしかない。しかし、四分音符あたり60というテンポが標準であると仮定すると、日本民謡1曲あたりの歌唱時間は平均して $6485.67 \div 48 = 135.11$ 秒の長さと考えられる。

<sup>8</sup> 変動係数 (coefficient of variation) は、データの相対的なばらつきを表す指標である。標準偏差を平均値 ( $\mu \neq 0$ ) で割ることで求められる。

数・持続時間がともに最大の楽曲は、高知県の盆踊唄（1,153 音，65,184 tick）であった。また、最小の楽曲は、新潟県の田植唄（13 音，720 tick）であった。

### 3. パターンの抽出方法

ここでは、音楽コーパスから音組織を抽出する方法として、可変長マルコフ連鎖（variable length Markov chain, 以下 VLMC）モデルを用いる。VLMC モデルは、状態系列の中から繰り返し出現するパターンを確率・統計的に抽出し、後述する文脈木とよばれる木構造（tree structure）で表現する（Ron et al. 1996）。その起源は、情報理論の符号化（Rissanen 1983）にあり、現在は、DNA の塩基配列のパターン特定、音声認識技術、文書のスペルチェック機能などに応用されている。楽曲分析に確率的な情報を用いる利点は、複数の候補（仮説）の中から、最適な（尤もな）候補を選択でき、候補間の順位付けが可能となる点にある。

ここでの狙いは、音楽コーパスの旋律を状態系列として表現し、VLMC モデルに当てはめることで、旋律の中に繰り返し出現するフレーズを効率的に抽出することである。抽出したフレーズに共通する性質を読み解くことで、日本民謡の音組織を決定する。

#### 3.1 可変長マルコフ連鎖

確率過程  $\{X_t\}(t \in \mathbf{Z})$  を考え、任意の確率変数  $X_t$  の実現値  $x_t \in \mathcal{X}$  を時刻  $t$  の状態とよぶ。そして  $X_t$  が直前の  $N$  個の状態のみに依存する確率過程を  $N$  重マルコフ過程とよぶ。とくに、人は先の展開が不確かな旋律を聴取する際に、聴き手が過去に蓄積してきた知識（スキーマ, schema）に基づいて後続音を予測する（Temperley 2007）。このように過去の有限個の情報に依存して、未来の挙動を決定する仕組みは、マルコフ過程に他ならない。直前の  $N$  個の状態から次の状態（ $N+1$  個目）が決まることは、次のように定式化できる：

$$\begin{aligned} & P(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_1 = x_1) \\ &= P(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_{t-N} = x_{t-N}). \end{aligned} \quad \text{———} (*)$$

時刻が離散的であるマルコフ過程のことをマルコフ連鎖とよぶ。マルコフ連鎖は確率・統計的言語モデルにおいて最も一般的なモデルである（Manning and Schütze 1999）。ただし、 $N$  の値が小さいと、状態系列の構造上の特徴を十分に反映したモデルを構築することが難しくなる。一方で、 $N$  の値が大きいと、パラメータ数が指数関数的に増大して統計的に信頼性におけるモデルを構築することが難しくなる（Pearce et al. 2008）。以下では、 $N$  を固定長ではなく可変長にすることで高い精度と信頼性を保ったモデルを扱う。

VLMC は、状態の履歴に応じて動的に  $N$  の値が変化するマルコフ連鎖である。例えば DNA の塩基配列  $\{X_t\}$  があり、確率変数  $X_t$  の実現値が  $\mathcal{X} = \{A, C, G, T\}$  であるとする。このとき通常の 2 重マルコフ連鎖は次のように最大 16 通りの推移確率ベクトルをもった確率過程として表される：

$$\mathbf{P}(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_1 = x_1) = \begin{cases} \mathbf{P}(X_t = x_t | X_{t-1} = A, X_{t-2} = A) \\ \mathbf{P}(X_t = x_t | X_{t-1} = A, X_{t-2} = C) \\ \mathbf{P}(X_t = x_t | X_{t-1} = A, X_{t-2} = G) \\ \mathbf{P}(X_t = x_t | X_{t-1} = A, X_{t-2} = T) \\ \mathbf{P}(X_t = x_t | X_{t-1} = C, X_{t-2} = A) \\ \vdots \\ \mathbf{P}(X_t = x_t | X_{t-1} = T, X_{t-2} = G) \\ \mathbf{P}(X_t = x_t | X_{t-1} = T, X_{t-2} = T). \end{cases}$$

これに対して VLMC モデルでは、次のように  $\{X\}$  の履歴に応じていくつかのベクトルを凝縮して表現する：

$$\mathbf{P}(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_1 = x_1) = \begin{cases} \mathbf{P}(X_t = x_t | X_{t-1} = A) \\ \mathbf{P}(X_t = x_t | X_{t-1} = C) \\ \mathbf{P}(X_t = x_t | X_{t-1} = G, X_{t-2} = C) \\ \mathbf{P}(X_t = x_t | X_{t-1} = G, X_{t-2} \in \{A, G, T\}) \\ \mathbf{P}(X_t = x_t | X_{t-1} = T). \end{cases}$$

この事例では、 $\{X_t\}$  の履歴を再帰的に参照し、 $X_{t-1} = A, C, T$  に該当するパターン直前の振る舞い  $X_{t-2} \in \{A, C, G, T\}$  はモデルに組み込まないことを判断している。また、 $X_{t-1} = G$  を含むパターンでは、 $X_{t-2} = C$  および  $X_{t-2} \in \{A, G, T\}$  に場合分けして表現している。以降の節では、状態系列は

$$x_i^j = x_j, x_{j-1}, x_{j-2}, \dots, x_{i-1}, x_i \quad (i < j; i, j \in \mathbf{Z})$$

として、 $x_t$  の時刻が遡るように系列を記述する。この表記法を導入することで式 (\*) の  $N$  重マルコフ過程は次のように短く表現できる：

$$P(X_t = x_t | X_1^{t-1} = x_1^{t-1}) = P(X_t = x_t | X_{t-N}^{t-1} = x_{t-N}^{t-1}).$$

### 3.2 文脈木の構築

VLMC モデルは文脈木 (context tree) とよばれる木構造として表現できる。図3は、文脈木の例である。頂上に根をもち、上から下に向かって枝葉を広げた階層構造である。各階層は、上から順に1次、2次、3次、…と数え、そこに配置されている枝葉の分岐のことを節点とよぶ。各節点には、実現値  $\mathcal{X}$  のラベルが割り当てられている。ただし、根のラベルは空列  $\varepsilon$  である。

節点から根に至るまでの経路は、文脈 (context) とよばれ、状態系列に存在する推移パターンに対応する。つまり、文脈をたどることで元の状態系列の中に出現したパターンを復元することができる。各節点には、その文脈の条件付き生起確率値が割り当てられている。例えば、図3ではある節点  $n$  のラベルが  $\alpha \in \mathcal{X}$ 、その子節点  $n'$  のラベルが  $\beta \in \mathcal{X}$  であるとき、 $n'$  から  $n$  に至



る経路は  $\alpha\beta$  (ただし,  $x_0 = \alpha, x_1 = \beta$ ) と表現する。そして, 任意の節点  $n$  における確率分布  $P_n(x)$  は  $\sum_{x \in \mathcal{X}} P_n(x) = 1$  を満たす。

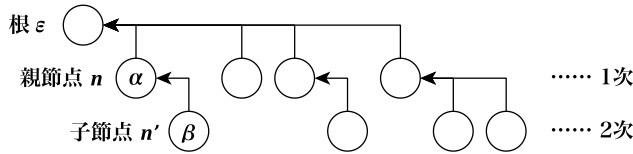


図3 文脈木の例

3.4節で述べるように, 節点から根に至るまでの確率値を掛け合わせることで, その文脈の生成確率を求めることができる。したがって, 生成確率の値が高い文脈ほど, 状態系列の中に頻出するパターンであると判断することができる。

VLMC モデルは, 根  $\varepsilon$  のみ存在する文脈木を初期状態とし, ここに状態系列の履歴を有効に反映する節点を順次追加していき, 最終的に目的の文脈木を完成させていく。図3のように, 任意の節点  $n$  に新しく子節点  $n'$  を追加するには, 状態系列の条件付き生起確率を有効に計算できるものを採用しなければならない。ここでは Kullback-Leibler 距離<sup>9</sup>を用いて, 2つの確率分布  $P(\cdot|x_0 = \alpha)$  と  $P(\cdot|x_0 = \alpha, x_1 = \beta)$  を比較し, 両者が類似している節点を追加する。与えられた状態系列のラベルを先頭から読み取り, 以上の確率分布の比較を再帰的に実行し, 根  $\varepsilon$  に節点を順次追加していく。

しかし, すべての履歴を考慮してしまうと, 生起確率の計算量が増大するため, 出現頻度の低いパターンは, 何らかの閾値を設けて考慮しない方が経済的である。この操作は, 文脈木を整えることにあわせて枝刈り (pruning) とよばれる。ここでは context algorithm (Rissanen 1983, Bühlmann 2002) を用いて, 出現頻度の低いパターンを枝刈りしながら文脈木を構築する。

### 3.3 閾値 $\hat{K}$ の決定

ここでは文脈木の枝刈りに使う閾値の決定方法について概説する。文脈  $x$  の項数を  $|x|$  とすれば, context algorithm では, 原理的に文脈  $x_i^j$  を  $wu = w_{|w|}, w_{|w|-1}, \dots, w_2, w_1, u_{|u|}, u_{|u|-1}, \dots, u_2, u_1$  のように任意の時刻で文脈  $w$  と  $u$  に分割し, Kullback-Leibler 距離に基づく指標  $\Delta_{wu}$  を用いて,  $w$  と  $u$  の確率分布を比較する:

$$\Delta_{wu} = \sum_{x \in \mathcal{X}} \hat{P}(x|wu) \log \left\{ \frac{\hat{P}(x|wu)}{\hat{P}(x|w)} \right\} \cdot N(wu) = D(\hat{P}(x|wu) || \hat{P}(x|w)) \cdot N(wu) > K.$$

<sup>9</sup> 真のモデルと近似モデルの比較 (モデル選択) では一般的な尺度であり, 相対エントロピーともよばれる。2つの確率分布が  $P(x)$  と  $Q(x)$  であるとき, Kullback-Leibler 距離は次のように定義される:

$$D(P(x) || Q(x)) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}.$$

厳密には三角不等式を満たさないため, 数学的な意味での距離ではない。

文脈木の枝刈りは、 $\Delta_{wu}$  が閾値  $K$  を越えるか否かによって決断されるため、閾値  $K$  の決定はモデルの構築を左右する重要な役割を担う。恣意性を限りなく排除し客観的に閾値を決定するために、ここでは赤池情報量基準 (Akaike's information criterion, 以下 AIC) が最小となる  $K$  を閾値  $\hat{K}$  として用いる<sup>10</sup>。

以上の方針に従って、音楽コーパスから生成した状態系列に対して、閾値  $\hat{K}$  を用いて文脈木の枝刈りを実施する。

### 3.4 確率計算

#### 3.4.1 節点の推移確率

文脈木の任意の節点には  $P(X_t = x_t | X_{t-N}^{t-1} = x_{t-N}^{t-1})$  が割り当てられる。これを節点の推移確率とよぶ。状態系列における文脈  $x_t^t$  の出現回数を  $\#(x_t^t)$  とすれば、節点の推移確率は状態系列に出現する文脈の相対頻度から次のように算出できる：

$$P(X_t = x_t | X_{t-N}^{t-1} = x_{t-N}^{t-1}) = \frac{\#(x_{t-N+1}^t)}{\#(x_{t-N+1}^{t-1})}.$$

#### 3.4.2 文脈の生成確率

文脈  $x_1^t$  の生成確率  $P(X_1^t = x_1^t)$  は、乗法定理 (chain rule) によって次のように算出する：

$$P(X_1^t = x_1^t) = P(X_t = x_t) \prod_{i=2}^t P(X_{t-i+1} = x_{t-i+1} | X_{t-i+2}^t = x_{t-i+2}^t).$$

この計算を文脈木の各文脈について実行し、生成確率の高い文脈をもって音楽コーパスに頻出する特徴的なパターンとして抽出する。以降では、 $x_0$  からみた  $x_{N+1}$  を第  $N$  推移とよぶ。

## 4. モデルの実現方法

### 4.1 記号列の生成手順

音楽コーパスの旋律から記号列 (状態系列) を生成する方法を述べる。ここでは、旋律の音程情報をもつ記号列を生成する。音程とは、音高と音高の差分のことであり、半音の数や表 1 の名称を使って表現する。人は旋律を記憶したり、違いを聴き分けたり、歌うときに音程を頼りにしていることが心理実験を通して確認されている (Eerola et al. 2001)。そのため、旋律の音高では

<sup>10</sup> Kullback-Leibler 距離の定義より、 $D(P(x)||Q(x)) = 0 \Leftrightarrow P(x) = Q(x)$  が成り立つ。このことから、 $D(P(x)||Q(x))$  の値が小さいほど  $P(x)$  と  $Q(x)$  が類似していると判断できる。同様に、 $\Delta_{wu}$  の値が小さいほど文脈  $w$  に  $u$  を連結する必要性はないと判断し、 $wu$  の後半  $u$  部分を枝刈りする。したがって、ここでは真のモデルと VLMC へ近似するモデルを相互に比較する Kullback-Leibler 距離が最小となる  $K$  の値を枝刈りに用いることが望ましい (Shibata 1997)。

3.1 節で述べたように、パラメータ数が多いほど精度の高いモデルが得られる一方で、統計的に信頼性におけるモデルを構築することが難しいというトレードオフの問題があった。AIC は、客観的にモデルのパラメータ数を抑制する方針であり、AIC が最小となるモデルを選択することで、良いモデルが選択される (坂元ほか 1983)。そして、AIC が最小となる  $K$  が Kullback-Leibler 距離を最小とする (Tong 1975)。

よって、ここでは  $\hat{K} = \arg \min_K \{AIC(K)\}$  を探索し、これを閾値として文脈木の枝刈りに用いることで、問題を回避しながら最適な文脈木を構築する。

なく音程を  $x_i$  に当てはめることは、民謡を創作した人々の音楽認知に近い表現方法といえる。

表1 半音の数と音程名称の対応関係 (抜粋)

半音	音程名称	半音	音程名称
0	完全1度	7	完全5度
1	短2度	8	増5度, 短6度
2	長2度	9	長6度
3	短3度	10	短7度
4	長3度	11	長7度, 減8度
5	完全4度	12	完全8度
6	増4度, 減5度	13	増8度

音程が上行する場合、音高の差分は半音の数を正の整数 (+1, +2, +3, …) で表し、音程が下行する場合、音高の差分は半音の数を負の整数 (-1, -2, -3, …) で表す。例えば、音程  $x_i = +3$  とは、A4 → C5 (ラ→ド) のように音高間が半音3つ分 (短3度) 上行していることを示し、 $x_i = -7$  とは、C5 → F4 (ド→ファ) のように音高間が半音7つ分 (完全5度) 下行していることを示す (図4)。

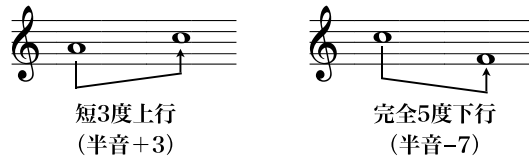


図4 旋律の数値化の例

また、音価 (音の持続時間) と音程0 (同じ音高間の移動, 完全1度音程) の情報をどのように処理するかという問題がある。ここでは、音価も音程0も考慮しないことにした。音価を考慮しない場合、音組織を精確に捉えられないという懸念がある。しかし、音価は必ずしも重要な情報でないことが報告されている。Chai and Vercoe は民謡の旋律を複数の記号列——音価を考慮しない音程の記号列；音価を重み付けした記号列；音程を数段階の高低に近似した記号列——で表現し、マルコフ連鎖モデルに基づく分類実験を行ったところ、音価を考慮しない記号列が最も高い分類精度を与えることを示した (Chai and Vercoe 2001)。人は、音の反復を結びつけて1つのパターンとして知覚し、音高の段階的な変化がない部分を音楽の進行と知覚しないことが心理実験を通して確認されている (Snyder 2000)。このため、音楽の進行に関わらない音程0については考慮しないことにする。

以上の先行研究の知見に基づき、旋律の音程情報から音価と音程0を除いた記号列を次の手順で作成する：

- Step 1. 楽曲データの旋律の音高情報を状態系列  $\{s_i\}$  として取り出す。
- Step 2.  $\{s_i\}$  の任意の音高  $s_k$  を次式で整数値に変換し、 $\{\sigma_i\}$  を生成する：

$$\sigma_k = 69 + 12 \times \left\lfloor \log_2 \frac{s_k}{440} \right\rfloor, \quad \text{ただし音高A4} \equiv 440[\text{Hz}].$$

Step 3.  $\{\sigma_k\}$  の前後の音高の差分を取って音程情報の状態系列  $\{x_k\}$  を生成する。ただし、任意の音程が  $\sigma_{k+1} - \sigma_k = 0 (k = 1, 2, 3, \dots, |s|-1)$  となる項は取り除く。

## 4.2 青森県の子守唄を使ったパターン抽出の事例

### 4.2.1 記号列の生成

音楽コーパス全体の解析に入る前に、青森県に古くから伝わる子守唄（図5）の旋律から文脈木を構築し、楽曲に頻出するフレーズを抽出する。

はじめに、旋律から音程の記号列を生成する。この楽曲には、50個の音高が採譜されており、 $\{s_k\}$  は次のように取り出せる：

$$(s_1, s_2, s_3, \dots, s_b, \dots, s_{|s|}) = (E^b 5, D5, E^b 5, C5, \dots, D5, D5, D5) \quad \text{——(Step 1)}$$

ただし、 $|s| = 50$  は  $\{s_k\}$  の項の総数を表す。続いて、 $\{s_k\}$  の各項を整数値に変換する：

$$(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_b, \dots, \sigma_{|s|}) = (75, 74, 75, 72, \dots, 74, 74, 74) \quad \text{——(Step 2)}$$

そして、 $\{\sigma_k\}$  の前後の項の差分を取り、 $\sigma_{k+1} - \sigma_k = 0$  となる項をすべて除いた記号列  $\{x_k\}$  を生成する：

$$(x_1, x_2, x_3, \dots, x_p, \dots, x_n) = (-1, +1, -3, +3, \dots, -2, -5, +7) \quad \text{——(Step 3)}$$

ただし、 $n(\leq |s|-1)$  は  $\{x_k\}$  の項数を表す。



図5 青森県の子守唄

### 4.2.2 文脈木の構築

この記号列を context algorithm に当てはめることで文脈木を構築する（図6）。頂上に根  $\varepsilon$  があり、各節点には音程情報のラベル  $\mathcal{X}$  とともに、その音程が文脈中に出現する確率値が与えられている。

図6では、1次の文脈が3つ—— $x_0 = -2, -1, +3$ ——；2次の文脈が5つ—— $(x_0, x_1) = (-3, -1), (+1, +2), (+1, +4), (+2, +4), (+4, -1)$ ——；3次の文脈が1つ—— $(x_0, x_1, x_2) = (+1, -1, +3)$ ——；以上9つの文脈が抽出されており、それぞれ図5の旋律中に繰り返し出現するフレーズに対応する。例えば、 $(x_0, x_1, x_2) = (+1, -1, +3)$  は文脈全体の音程推移が +3（短3度上行）を形成するフレーズで

あり、この楽曲中に 4 回出現している。

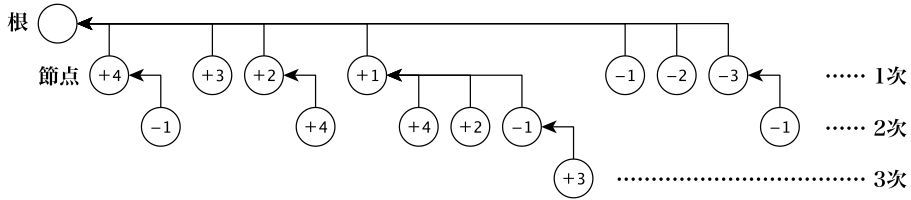


図6 青森県の子守唄から構築した文脈木

## 5. パターン抽出の結果

### 5.1 音楽コーパスの文脈木

#### 5.1.1 閾値の推定

音楽コーパスの楽曲データを用いて  $K$  の値を 0.01 刻みの間隔で変化させながら  $AIC(K)$  の値を求めた。その結果、各コーパスにおける  $AIC$  の最小値を与える  $K$  は次の通りであった：

$$\begin{aligned} \text{日本民謡} : \hat{K} &= 28.92, \min\{AIC\} = 452738, \\ \text{中国民謡} : \hat{K} &= 24.82, \min\{AIC\} = 363319. \end{aligned}$$

これらの値を context algorithm における閾値  $\hat{K}$  として設定し、音楽コーパスから生成した記号列  $\{x_i\}$  から文脈木 (VLMC モデル) を構築する。

#### 5.1.2 文脈木の基本統計量

日本民謡と中国民謡の音楽コーパスから構築した文脈木をそれぞれ図 7, 図 8 に示す。日本民謡の音楽コーパスから構築したモデルでは、総計 412 の文脈をもつ巨大な文脈木を得た。このうち末端の節点から根  $\varepsilon$  に至る文脈は 289 種類存在した。マルコフ連鎖の最大次数 (最長の文脈) は、 $N = 7$  次であり、文脈あたりの平均次数は  $\bar{N} = 3.132$  次であった。このことから、日本民謡はおよそ 3 音程、長くても最大 7 つの音程を連結した動きがフレーズの単位であり、それらを相互に組み合わせて旋律を作り上げていると考えることができる。

また、中国民謡の音楽コーパスから構築したモデルでは、総計 316 の文脈をもつ文脈木を得た。このうち末端の節点から根  $\varepsilon$  に至る文脈は 234 種類存在した。マルコフ連鎖の最大次数は、 $N = 5$  次であり、文脈あたりの平均次数は  $\bar{N} = 2.594$  次であった。

このことから、中国民謡もおよそ 3 音程、長くても最大 5 つの音程を連結した動きがフレーズの単位であると考えられる。したがって、2 カ国の文脈数を比較すると、中国民謡よりも日本民謡の方が推移パターンに多様性があることがわかる。

次節では、2 つの文脈木の推移パターンをより詳細に比較していく。ただし、文脈の総数が多いため、紙面の都合上、ここでは文脈の生成確率が 1.00% 以上のパターンに着目する。そして、

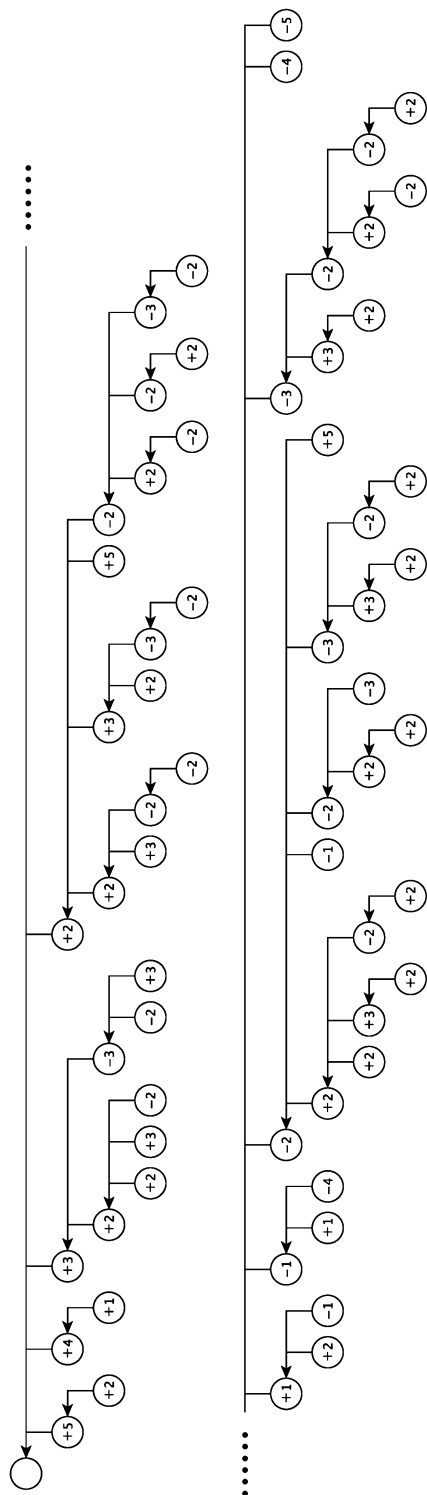


図7 日本民謡の音楽コーパス全体から構築した文脈木 (生成確率 1.00% 以上)

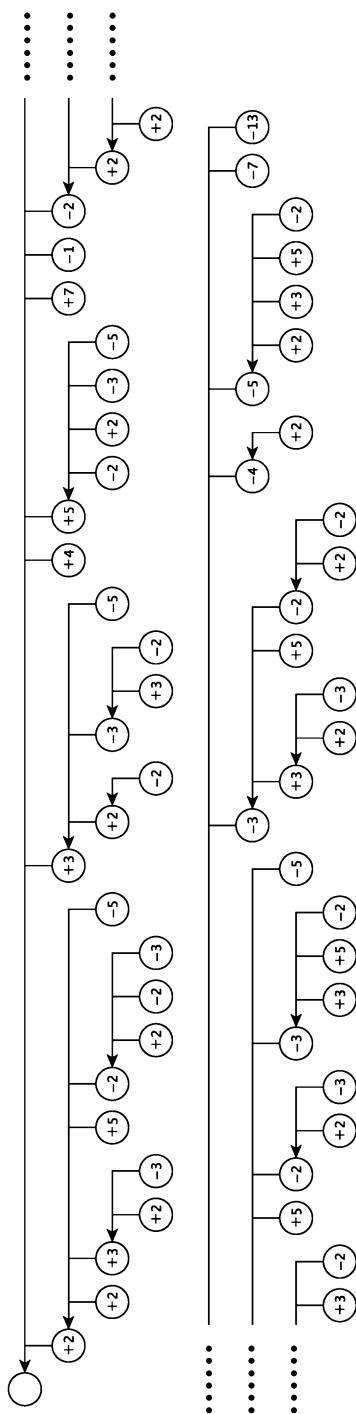


図8 中国民謡の音楽コーパス全体から構築した文脈木 (生成確率 1.00% 以上)

$N=1$  次, 2 次, 3 次, …の順にパターンの特徴を考察し, 日本民謡の音組織を捉える。

## 5.2 音程情報の分布

日本民謡と中国民謡の音楽コーパスから生成した記号列  $\{x_i\}$  に含まれる項の集計結果——(a) 音程  $x$  のヒストグラム (確率分布) と (b) 音程  $x$  の累積度数分布——をそれぞれ図 9, 図 10 に示す。また, 分布の基本統計量を表 2 (括弧内は音程 0 を考慮した場合の値) に示す。

表 2 より, 2 カ国の平均値を比較すると, どちらも  $0 < x < 1$  の範囲内にあることから, 音程の上行と下行には, ほとんど違いがないことがわかる。また, 日本民謡の変動係数の値が中国民謡よりも高いことから, 日本民謡の 1 音 1 音の唄われ方は中国民謡よりも偏りがあるといえる。

図 9 より, 日本民謡の第 1 推移  $x_0$  は,  $x = \pm 2, \pm 3$  を使用する傾向が極めて高く,  $|x| \leq 5$  の区間に全体の 97.04% が収まることから,  $|x| > 5$  よりも広い音程をほとんど使用しない傾向が読み取れる。これに対して, 図 10 より, 中国民謡の第 1 推移  $x_0$  は,  $|x| \leq 5$  の区間に全体の 90.31%,  $|x| \leq 7$  の区間に全体の 95.78% が収まる。

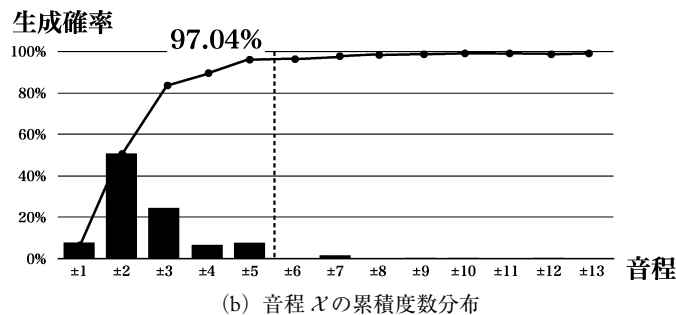
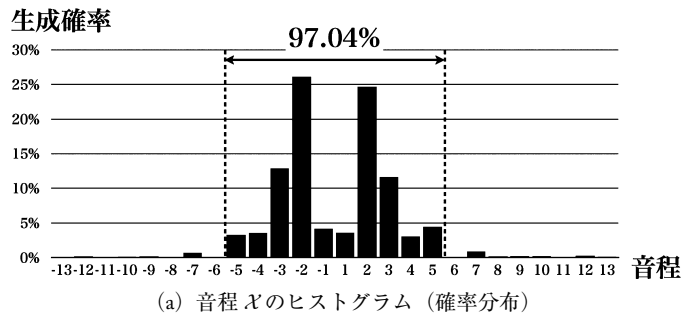


図 9 日本民謡の音程情報  $x$  の度数分布

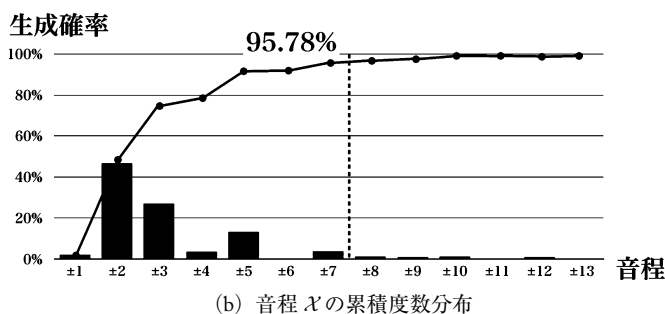
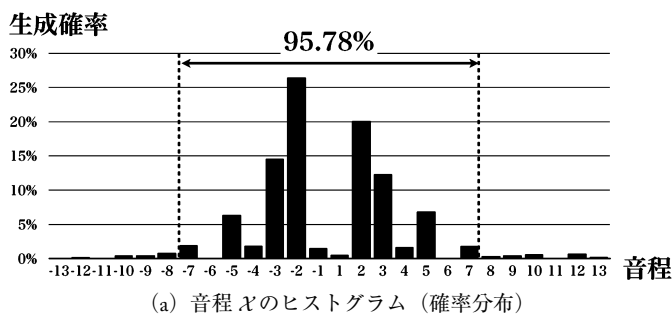
図 10 中国民謡の音程情報  $\mathcal{X}$  の度数分布

表 2 音程情報に関する基本統計量

	日本民謡		中国民謡	
音程の総数	161,177	(202,425)	106,013	(124,677)
音程の平均値	0.01	(0.51)	-0.13	(0.39)
音程の標準偏差	3.09	(7.59)	3.77	(12.07)
変動係数 $C.V.$	400.29	(14.88)	-29.60	(30.95)

したがって、中国民謡の音程情報と比較して、日本民謡の旋律を構成する 1 音 1 音の動きは、 $\mathcal{X} = \pm 5$ 、すなわち完全 4 度音程の狭い音程の範囲内に、ほぼ収まる特徴があることがわかる。

### 5.3 第 2 次 (3 音高間) の推移パターン

表 3 と表 4 は、2 次の文脈 ( $x_0, x_1$ ) の生成確率を算出し、節点ごとに上位 3 位までの文脈を集計した結果である。

表 3 より、日本民謡では、文脈全体の音程推移 ( $\sum x_i$ ) が 0 または  $\pm 5$  を形成するパターンが上位に現れることがわかる。文脈全体の音程推移が 0 となるパターンは、3 音の並びのうち、最初と最後の音高の位置が変化しないフレーズ——例えば、図 11 (a) の  $A4 \rightarrow C5 \rightarrow A4$  (ラ→ド→ラ) や  $A4 \rightarrow E4 \rightarrow A4$  (ラ→ミ→ラ) ——を指す。合計が  $\pm 5$  となるパターンは、3 音の並びのうち、最初と最後の関係が完全 4 度音程を形成するフレーズ——例えば、図 11 (b) の  $A4 \rightarrow C5 \rightarrow D5$



(ラ→ド→レ) や A4 → G4 → E4 (ラ→ソ→ミ) ——を指す。とくに、このような3音の並びのうち、中間の音高が完全4度音程を越えないパターン——例えば、図 11 (b)。図 11 (c) の A4 → E5 → D5 (ラ→ミ→レ) や A4 → C4 → E4 (ラ→ド→ミ) では、中間の音高が完全4度音程を越えてしまう——は、小泉文夫のテトラコルド理論<sup>11</sup> とよばれる音型 (フレーズを構成する単位) そのものである。

表3 日本民謡における2次の文脈の生成確率 (上位3位/節点)

$x_0$	$x_{-1}$	合計	生成確率		$x_0$	$x_{-1}$	合計	生成確率		$x_0$	$x_{-1}$	合計	生成確率		$x_0$	$x_{-1}$	合計	生成確率	
-1	-2	-3	34.87	都節	+1	-1	0	57.23	琉球	-2	+2	0	38.61	律	+2	-2	0	44.58	民謡
-1	+1	0	29.34		+1	+4	+5	29.08		-2	-3	-5	26.41		+2	+3	+5	20.16	
-1	-4	-5	19.32		+1	-5	-4	6.95		-2	-2	-4	19.38		+2	+2	+4	17.38	
-3	-2	-5	44.38	民謡	+3	+2	+5	51.39	律	-4	-1	-5	42.79	琉球	+4	-4	0	30.74	都節
-3	+3	0	37.06		+3	-3	0	34.55		-4	+4	0	25.05		+4	+1	+5	22.17	
-3	+5	+2	7.39		+3	-5	-2	5.36		-4	+2	-2	16.71		+4	-2	+2	20.11	
-5	+2	-3	21.97	律	+5	+2	+7	31.02	律	-7	+2	-5	49.57		+7	-2	+5	39.04	
-5	-2	-7	21.41		+5	-2	+3	24.78											
-5	+3	-2	21.20		+5	-3	+2	20.03											

表4 中国民謡における2次の文脈の生成確率 (上位3位/節点)

$x_0$	$x_{-1}$	合計	生成確率		$x_0$	$x_{-1}$	合計	生成確率		$x_0$	$x_{-1}$	合計	生成確率		$x_0$	$x_{-1}$	合計	生成確率	
-1	-2	-3	67.94		+1	+2	+3	34.40		-2	+2	0	29.82	律	+2	-2	0	35.85	民謡
-1	+1	0	16.79		+1	-1	0	24.96		-2	-3	-5	23.36		+2	+3	+5	22.02	
-1	+3	+2	8.08		+1	-3	-2	10.34		-2	-2	-4	20.33		+2	+2	+4	13.54	
-3	-2	-5	48.79	民謡	+3	-3	0	33.90	律	-4	+2	-2	45.79		+4	-2	+2	58.25	
-3	+3	0	30.50		+3	+2	+5	32.13		-4	-3	-7	23.13		+4	+3	+7	12.57	
-3	+5	+2	7.29		+3	-5	-2	13.78		-4	+4	0	12.36		+4	-4	0	11.19	
-5	-2	-7	24.32	律	+5	-3	+2	25.75	民謡	-7	+2	-5	30.57	律	+7	-2	+5	34.11	
-5	+2	-3	20.86		+5	+2	+7	18.73		-7	+3	-4	15.73		+7	+2	+9	13.81	
-5	+3	-2	18.27		+5	-2	+3	17.31		-7	+5	-2	9.79		+7	-3	+4	12.31	

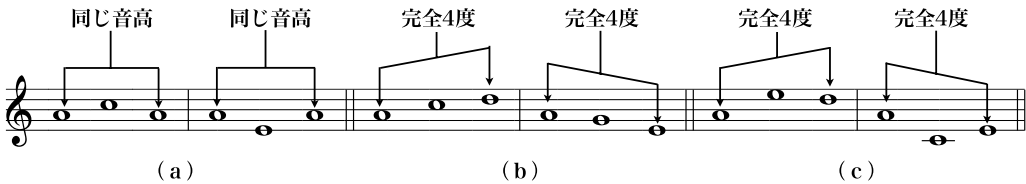


図 11 2次の音程推移が0と±5となるパターンの譜例

<sup>11</sup> 小泉文夫のテトラコルド理論 (小泉 1958) については、稿末の付録を参照。

これに対して、表4より、中国民謡においても、文脈全体の音程推移が0または±5を形成するパターンが上位に現れることがわかる。しかし、その割合は日本民謡よりも低い。そして日本民謡との決定的な違いは、文脈全体の音程推移が±7、すなわち完全5度音程を形成するパターンが上位に現れる点にある。このことから、中国民謡は日本民謡よりも完全5度音程を形成し易い傾向があるといえる。

したがって、中国民謡と比較して、日本民謡の特徴は、2次の音程推移では、フレーズの最初の音に回帰するパターンと、フレーズ全体が小泉のテトラコルドを形成するパターンが極めて多く出現することとまとめることができる。

#### 5.4 第3次（4音高間）以上の推移パターン

日本民謡について、3次の文脈( $x_0, x_{-1}, x_{-2}$ )の生成確率を算出し、降順に並べてみると、これまでの結果と同様に、文脈全体の音程推移が0または±5を形成するパターンが上位を占める。表5は、生成確率の上位10位の結果——(a)文脈全体で0を形成するパターンと(b)文脈全体で±5を形成するパターン——であり、図12は、これらのパターンを具体的な音高に変換して再現した譜例である。表5および図12から明らかのように、文脈全体の音程推移が0を形成するパターンでは、4つの音高の内部に小泉のテトラコルド——民謡、律、都節——を維持しつつ、最初の音高に戻っている。また、音程の合計が±5を形成するパターンにおいても、やみくもに音高が動いて完全4度音程を形成するのではなく、小泉のテトラコルドを維持しながら完全4度

表5 日本民謡の3次の文脈全体が0または±5を形成するパターン（上位10位）

	$x_0$	$x_{-1}$	$x_{-2}$	合計	生成確率	
(イ)	+5	-3	-2	0	0.65	民謡
(ロ)	-2	-3	+5	0	0.46	律
(ハ)	-3	-2	+5	0	0.43	民謡
(ニ)	+2	+3	-5	0	0.39	民謡
	-2	-2	+4	0	0.36	
(ホ)	-5	+3	+2	0	0.27	律
(ヘ)	+5	-1	-4	0	0.17	都節
	-2	-1	+3	0	0.11	
	-3	+1	+2	0	0.08	
(ト)	+4	+1	-5	0	0.06	都節

	$x_0$	$x_{-1}$	$x_{-2}$	合計	生成確率	
	-2	+2	+5	+5	0.70	
	+5	+2	-2	+5	0.51	
(チ)	+2	-2	+5	+5	0.49	民謡
	-5	-2	+2	-5	0.42	
(リ)	+2	+5	-2	+5	0.30	律
(ヌ)	+3	-3	+5	+5	0.29	律
(ル)	-2	-5	+2	-5	0.27	民謡
(ヲ)	-2	+2	-5	-5	0.24	律
	+2	-2	-5	-5	0.23	
	-3	-4	+2	-5	0.22	

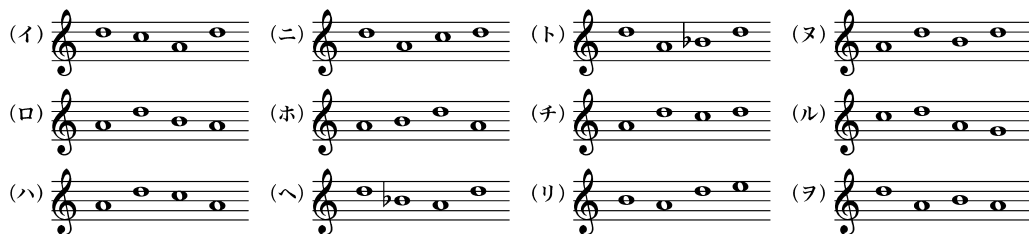


図12 表5の文脈に対応するフレーズの具体例

音程を形成していることが確認できる。これに対して、中国民謡では、日本民謡と異なる傾向を示すことがわかる（表6および図13）。

表6 中国民謡の3次の文脈全体が0または±5を形成するパターン（上位10位）

	$x_0$	$x_{-1}$	$x_{-2}$	合計	生成確率	
(ワ)	-2	-3	+5	0	1.17	律
(カ)	-5	+3	+2	0	0.81	律
(ヨ)	+5	-3	-2	0	0.63	民謡
(タ)	-3	-2	+5	0	0.59	民謡
	-2	-2	+4	0	0.49	
(レ)	+2	+3	-5	0	0.45	民謡
	+2	+2	-4	0	0.36	
	-2	-1	+3	0	0.11	
	+4	+3	-7	0	0.05	

	$x_0$	$x_{-1}$	$x_{-2}$	合計	生成確率	
	+5	+2	-2	+5	0.61	
(ソ)	-2	-5	+2	-5	0.56	民謡
	-2	+2	+5	+5	0.53	
(ツ)	-2	+2	-5	-5	0.48	律
	-2	-2	-1	-5	0.48	
	-5	-2	+2	-5	0.47	
	+2	-2	-5	-5	0.37	
(ネ)	+3	-3	+5	+5	0.37	律
(ナ)	+2	-2	+5	+5	0.33	民謡
(ラ)	+2	+5	-2	+5	0.33	律

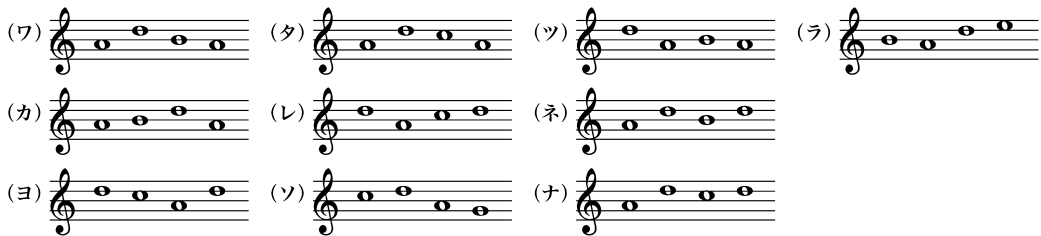


図13 表6の文脈に対応するフレーズの具体例

さらに、表7は、文脈全体の音程推移が0または±5を形成するパターンの種類を次数ごとに集計した結果である。表7より、0または±5を形成するパターンの違いは顕著に異なるといえる。これらのパターンは、日本民謡では偶数次で計62個、奇数次で計30個であるのに対して、中国民謡では偶数次で計35個、奇数次で計28個である。このことから、日本民謡では、偶数次と奇数次で形成するパターンに違いがあるといえる。

以上の結果から、日本民謡の旋律を形成する最も重要な単位は、フレーズの最初の音に戻るパターンと、フレーズ全体で完全4度音程を形成するパターンの2つであることが明白である。

表7 文脈の音程が0または±5を形成するパターンの個数

(a)	日本民謡	次数	2次	3次	4次	5次	6次
		音程推移0	11	11	15	0	7
		音程推移±5	15	19	13	0	1

(b)	中国民謡	次数	2次	3次	4次	5次	6次
		音程推移0	10	9	5	0	0
		音程推移±5	13	18	7	1	0

## 6. 日本民謡の音組織

日本民謡と中国民謡の音楽コーパスを構築し、VLMC モデルとして表現することにより、楽曲の特徴的なフレーズを抽出した。両コーパスから抽出したフレーズを相互に比較することにより、日本民謡の音組織は次のように整理することができる：

- ・日本民謡の旋律を構成する1音1音の動きは、中国民謡よりも狭く、完全4度を越えない狭い音程の範囲内にほぼ収まる
- ・日本民謡の旋律を形成する最も重要な単位は、フレーズの最初の音に戻るパターンと、フレーズ全体で完全4度音程を形成するパターンの2つである
- ・日本民謡と中国民謡には、共通して完全4度音程を形成するパターンが頻出するが、日本民謡は、完全5度音程を経由するパターンや、フレーズ全体で完全5度音程を形成するパターンを作る傾向が中国民謡よりも遥かに弱い
- ・日本民謡のフレーズの根底には、小泉文夫のテトラコルドを形成しようとする力が働く

## 7. 日本民謡の地域性

### 7.1 地域性の研究

日本民謡は、地域ごとに傾向や雰囲気が異なるといわれる。たしかに、北海道のアイヌ・ギリヤーク族の唄や、奄美大島の唄を一聴すれば、多くの人がその旋律が本土の民謡と異なることを感じるであろう。ところが、本土の民謡の地域差について考えてみると、解明できていないことが多い。例えば、東北の唄と九州の唄には、どのような旋律の違いがあるのだろうか。また、東北の唄や南西の唄は、カムチャッカ半島や朝鮮半島の音楽の影響を受けているのだろうか、といった疑問が挙げられる。そもそも本土の民謡には、明確な地域差は存在するのだろうか。

プロの民謡歌手や民謡界では、こぶしを効かせる「東物」、こぶしを効かせない「西物」という認識を共有している。また、東日本では、プロの民謡歌手が民謡を舞台歌謡にまで昇華させた一方で、西日本では、芸妓が民謡を酒宴の座敷歌として洗練させたという歴史的経緯がある。しかし、ここでいう東日本と西日本の区別は、舞台歌謡と座敷歌というスタイルの違いであり、日本民謡の音楽的特徴を捉えた上での地域性の説明として不十分である。いずれにしても、おぼろげながら日本民謡の地域性は認識されているものの、地図上に境界線を引いてその違いを明確にできているわけではない。

学術研究では、次の2つの説がある。音楽学者の柿木吾郎は、独自に考案した構造式に基づく楽曲分析から日本民謡を北方様式・中央様式・南方様式に分類した（柿木 1969, 1983）。北方様式は、東物の旋律を含み、中央様式と南方様式は、それぞれ西物の旋律を含むことから、柿木の見解は民謡界の認識とほぼ一致する。また、音楽学者の小島美子は、5つの音階を基準に『大観』に掲載されたおよそ百曲を分類し、従来の認識と異なる日本海側と瀬戸内海側の地域差を指摘した（小島 1992）。

しかし、柿木の分析は、楽曲データが少ないことや、構造式に基づく分類を誰もが同じように再現できないことなど、実証的見地から問題がある。また、小島の分析も、楽曲データが少ない

ことに加え、データの選出方法や分類の基準に5つの音階を使った論拠が示されないことなど、若干の恣意性が残る。以下では、特定の構造式や音階に基づく分類を試みるのではなく、前節において決定した日本民謡の音組織を基準に地域差を調べていく。

## 7.2 日本民謡の地域性の解析方法

音組織を決定したときと同様に、ここでも大規模な楽曲データを準備し、計量分析によって解決する。はじめに、日本民謡の音楽コーパスのデータを採録した地域ごとに整理し、地域コーパスを構築しなおす。そして、計量分析の結果から明らかにした日本民謡の音組織——小泉のテトラコルドを形成するパターン——の出現確率に基づく「地域」の分類実験を行う。ここでは、出現確率値を要素にもつ特徴ベクトルを地域ごとに算出し、それらを階層的クラスタリング (hierarchical clustering)<sup>12</sup> によってグループに分類していく。

なお、分類実験のための入力情報 (パラメータ) の取り方は数多く考えられるが、ここで小泉のテトラコルドを用いる論拠は、VLMC モデルによって明らかにされたように、日本民謡の旋律を形成する基本単位としてテトラコルドが機能しているためである。

## 8. 日本民謡の地域コーパス

### 8.1 日本列島の地域区分

日本列島の区分には、行政区画、交通事情、気候、地質学的特徴、選挙区など、いくつも考えられるが、意外なことに統一の見解がない。ここでは、地理学で一般的な「八地方区分」を基準に楽曲データの地域区分を考えていく。

八地方区分——北海道、東北、関東、中部、近畿 (関西)、中国、四国、九州・沖縄——のうち、今回は北海道と沖縄の楽曲を扱わないため、それらを除く7つの地方を軸として、最終的に図14のように、日本列島を11の地域へ分割した。

なお、中部地方にあたる9県は、山梨、長野、新潟を甲信越地方、富山、石川、福井を北陸地方、岐阜、愛知、静岡を東海地方とした。中国地方は、鳥取、島根を山陰地方、岡山、広島、山口を山陽地方とした。九州地方は、大分、宮崎、鹿児島を東九州地方、福岡、佐賀、長崎、熊本を西九州地方とした。また、東北地方、関東地方、近畿地方、四国地方は、収録曲数の割合から、八地方区分のままにしている。

音楽コーパスの楽曲データをこれら11地域に分割しなおし、これを新たに日本民謡の地域コーパスとよぶ。

<sup>12</sup> 階層的クラスタリングを用いる最大の利点は、入力データの比較対象同士を定量的な距離の尺度に落とし込み、類似するものを段階的にグループにできる点にある。

基本原理は、対象 (個体やグループ) 間の距離を定義し、その距離の近さによって段階的にグループにしていくものであり、結果をデンドログラム (dendrogram) とよばれる樹形図で表現する。デンドログラムを特定の階層 (高さ) で切断することで、グループの個体とそれに含まれる要素が確定する。ここでは実用性の高いウォード法 (Ward's method) によってデンドログラムを描く。

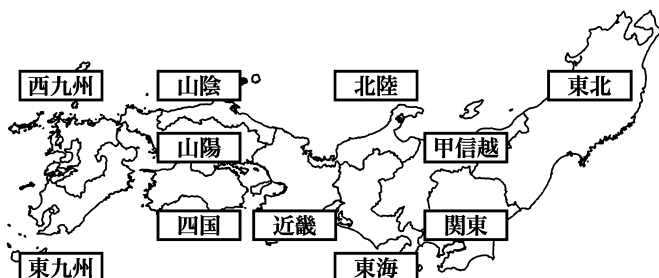


図 14 11 の地域区分 (地図上の区分)

## 8.2 地域コーパスの基本統計量

表 8 は、11 地域のコーパスに含まれる楽曲数と音程に関する基本統計量の一覧である。音程の平均値  $\bar{x}$  を比較すると、近畿を境界に北部は  $\bar{x} > 1$  であり、南部は  $\bar{x} < 1$  という特徴がみられる。ところで、表 2 では日本列島全体の楽曲データを対象として基本統計量を示し、極端に大きい変動係数の値から、1 音 1 音の使用傾向の偏りについて考察した。これに対して、表 8 では、各地域のコーパスの変動係数はいずれも低く、ばらつきが小さい。このことから、日本列島全域では、音の使用傾向に偏りはあるが、各々の地域内では音の使用傾向にほとんど偏りが生じていないことがわかる。

表 8 地域コーパスの基本統計量

	東北	関東	甲信越	北陸	東海	近畿	四国	山陰	山陽	東九州	西九州	全体
楽曲数	105	84	131	105	151	299	208	135	240	146	190	1,794
音程の平均値	1.59	1.40	2.05	1.60	1.15	0.12	0.65	0.33	0.03	0.97	0.61	0.69
音程の標準偏差	4.03	4.38	4.51	4.10	3.90	4.04	3.59	3.93	3.84	3.97	3.74	3.98
変動係数 C.V.	2.53	3.12	2.19	2.57	3.38	34.41	5.56	11.92	145.7	4.09	6.09	5.78

## 9. 階層的クラスタリングによる分類結果

### 9.1 テトラコルドの出現確率

表 9 は、日本民謡の楽曲コーパスを分割しなおした 11 地域について、小泉のテトラコルドの出現確率を求めた結果である。テトラコルドの出現確率の計算は、河瀬・往住 (2007) および Kawase and Tokosumi (2007) に基づく。表中の先頭行の項目はテトラコルドを形成する各パターンを示す<sup>13</sup>。これらの値を基準に階層的クラスタリングを適用する。

<sup>13</sup> テトラコルドは、下方核音、中間音、上方核音の 3 音から構成されており、1 種のテトラコルドにつき、その並び替え (転回形, inversion) は、 $3! = 6$  通り存在する。したがって、小泉のテトラコルドは 4 種存在することから、1 地域につき  $4 \times 6 = 24$  通りのパターンが考えられる。

表9 地域コーパスにおけるテトラコルドの出現確率

	民-p <sub>1</sub>	民-p <sub>2</sub>	民-p <sub>3</sub>	民-p <sub>4</sub>	民-p <sub>5</sub>	民-p <sub>6</sub>	都-p <sub>1</sub>	都-p <sub>2</sub>	都-p <sub>3</sub>	都-p <sub>4</sub>	都-p <sub>5</sub>	都-p <sub>6</sub>
東北	39.02	14.51	6.46	11.95	41.21	5.37	7.56	2.07	2.68	1.10	10.97	0.49
関東	44.09	11.71	5.08	9.95	46.08	4.42	13.15	3.20	3.65	1.77	21.00	1.44
甲信越	40.65	11.02	6.50	10.57	49.59	5.69	11.65	3.07	3.61	1.99	12.46	1.63
北陸	34.99	11.16	4.92	10.31	48.70	5.39	8.70	1.61	2.08	0.76	13.15	0.76
東海	38.64	10.43	4.58	14.07	40.66	5.59	11.04	2.15	1.55	2.15	17.64	1.41
近畿	43.38	10.35	4.94	9.63	48.47	2.75	7.34	1.88	2.19	0.69	12.63	0.91
四国	38.93	9.41	5.81	9.34	50.94	3.46	6.58	1.25	1.62	1.03	10.26	0.62
山陰	43.32	14.47	6.93	11.45	55.84	5.18	2.51	0.41	1.03	0.26	3.08	0.21
山陽	45.06	9.63	4.94	9.93	51.34	4.85	5.33	0.99	1.26	1.17	7.78	0.72
東九州	39.49	10.74	7.48	12.17	52.54	4.69	6.46	1.50	1.43	0.88	8.90	0.61
西九州	47.76	7.98	5.68	10.84	51.27	4.11	10.01	2.21	1.80	1.15	12.55	0.97

	律-p <sub>1</sub>	律-p <sub>2</sub>	律-p <sub>3</sub>	律-p <sub>4</sub>	律-p <sub>5</sub>	律-p <sub>6</sub>	琉-p <sub>1</sub>	琉-p <sub>2</sub>	琉-p <sub>3</sub>	琉-p <sub>4</sub>	琉-p <sub>5</sub>	琉-p <sub>6</sub>
東北	39.02	8.05	3.54	6.58	32.68	6.34	4.39	0.12	0.12	0.24	4.63	0.37
関東	39.23	5.75	2.32	5.64	39.45	6.96	3.76	0.33	0.22	1.77	8.62	1.44
甲信越	39.47	10.03	4.15	5.78	43.72	5.51	4.97	0.18	0.00	1.08	6.68	0.72
北陸	33.01	5.96	5.11	8.80	36.88	4.16	3.31	0.38	0.09	0.28	5.49	0.09
東海	33.72	5.86	5.05	5.45	31.97	5.25	5.32	0.81	0.34	0.47	7.81	0.34
近畿	35.13	6.13	5.03	4.81	37.16	2.84	4.47	0.53	0.13	0.47	6.63	0.16
四国	33.04	6.51	5.99	3.49	38.78	3.82	3.12	0.37	0.07	0.51	5.22	0.07
山陰	38.49	5.95	5.03	5.95	44.40	5.49	1.59	0.10	0.05	0.05	2.1	0.05
山陽	37.19	5.18	3.80	5.95	41.68	5.59	2.21	0.39	0.42	0.48	4.22	0.06
東九州	28.68	5.84	4.62	4.42	39.96	4.49	4.28	0.34	0.61	0.61	7.00	0.14
西九州	38.76	5.12	3.00	4.34	40.75	3.05	4.80	0.65	0.32	0.37	6.78	0.00

## 9.2 地域のクラスタリング

図15は、表9の数値を入力情報として用いて構築したデンドログラム（樹形図）である。デンドログラムの低い階層に着目すると、7つのグループ——「東北、北陸」、「東海」、「関東、甲信越」、「近畿、四国」、「山陽、西九州」、「東九州」、「山陰」——が確認できる。そして、階層を高めていくにつれて、地図上の隣接した地域同士が次第に階層的にグループを形成していく様子がわかる。

さらに、グループが二分されるようにデンドログラムを高い階層で切断すると、「東北、関東、甲信越、北陸、東海」と「近畿、四国、山陰、山陽、東九州、西九州」を得る。これを地図上にプロットすると、日本が東西に二分されることがわかる（図16）。以降、東北を含むグループを  $C_{East}$ 、近畿を含むグループを  $C_{West}$  とよぶ。

## 9.3 コサイン類似度を用いた東西の境界線の判定

$C_{East}$  と  $C_{West}$  の音楽的特徴の違いを考察する前に、東西の境界線を精査する。東西の境界線には、3地域——北陸、東海、近畿——が接している。ここではコサイン類似度（cosine similarity）を使って3地域の特徴が東日本と西日本のどちらに近いかを確認する。

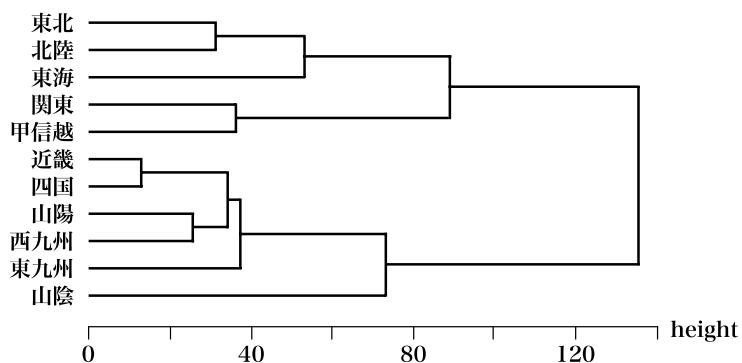
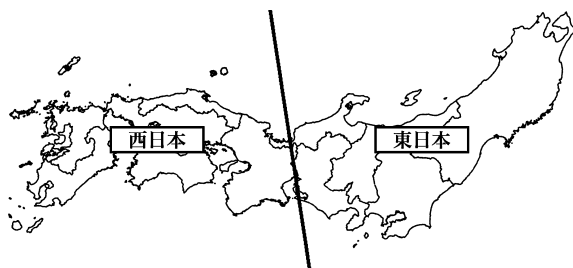


図 15 11 地域のデンドログラム

図 16 地図上にプロットした東側グループ  $C_{East}^-$  と西側  $C_{West}^-$  グループ

まず、 $C_{East}$  から北陸と東海の楽曲データを除いた  $C_{East}^-$  グループと、 $C_{West}$  から近畿の楽曲データを除いた  $C_{West}^-$  グループを準備し、両グループの特徴ベクトル  $V_{East}^-$  と  $V_{West}^-$  を求める：

$$V_{East}^- := (41.25, 12.41, 6.01, 10.82, 45.63, 5.16, 10.79, 2.78, 3.31, 1.62, 14.81, 1.19, 39.24, 7.94, 3.34, 6.00, 38.62, 6.27, 4.37, 0.21, 0.11, 1.03, 6.64, 0.84),$$

$$V_{West}^- := (42.91, 10.45, 6.17, 10.75, 52.39, 4.46, 6.18, 1.27, 1.43, 0.90, 8.51, 0.63, 35.23, 5.72, 4.49, 4.83, 41.11, 4.49, 3.20, 0.37, 0.29, 0.40, 5.06, 0.06).$$

そして、コサイン類似度を用いて北陸、東海、近畿の3地域の特徴ベクトルがベクトル空間上で  $V_{East}^-$  または  $V_{West}^-$  のどちらに近いかを測り、近い方のグループに分類する。その結果、北陸と東海は  $C_{East}^-$  に、近畿は  $C_{West}^-$  に分類され、前述のクラスタリング結果に変化は生じなかった。

なお、県単位でも同様の分類実験を試みたが、地図上の区分通りに分類されなかった。これは県単位で検証するための基礎データが不足しているためである。県や令制国（旧国）単位で分類実験を行うためには、日本民謡の音楽コーパスをさらに増強する必要がある。



#### 9.4 種目のクラスタリング

地域の違いだけでなく、本研究の分析に使用した日本民謡の5種目に対して階層的クラスタリングを適用した。結果は図17に示す。

グループが二分されるようにデンドログラムを高い階層で切断すると、2つのグループ  $C_1 = \{\text{田植唄, 盆踊唄, 田草取唄}\}$ ,  $C_2 = \{\text{子守唄, 地形唄}\}$  を得る。

$C_1$  では、デンドログラムの低い位置で  $\{\text{田植唄, 盆踊唄}\}$  というグループを形成している。この結果は、兩種目のテトラコルドの使用傾向が類似していることを示す。実際に、田植えのない時季に田植唄の旋律を盆踊唄に借用し、盆踊りのない時季に盆踊唄の旋律を田植唄に借用しており、これを計量的に裏付けていると解釈することができる。そして高い位置で  $\{\text{田草取唄}\}$  が組み込まれる。田植唄と田草取唄は、労働の観点で同じ種目に属しており、田植唄と盆踊唄の関係のように兩種目間で旋律の共有・借用が生じていると考えることができる。

$C_2$  では、 $\{\text{子守唄, 地形唄}\}$  というグループを形成している。これは労働の観点から共通点を説明することが難しい。ここでは、 $C_1$  以外の2種目が1つのグループを形成したと考えることができる。本研究では、分析対象を5種目に限定しているが、今後データを整備し、種目数を増やして同様の手法を適用することによって、日本民謡の種目間の共通点が明らかになる。

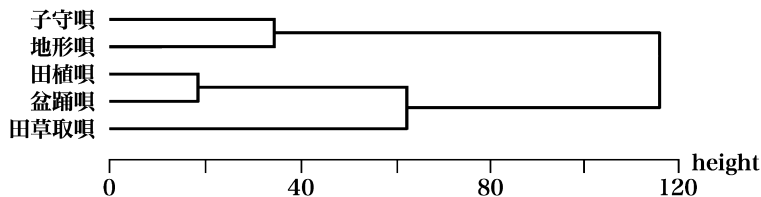


図17 5種目のデンドログラム

### 10. 日本民謡の地域性に関する考察

#### 10.1 民俗学的側面

民俗学者の宮本常一は、田楽（豊穡祈願や魔事退散祈願を目的とする伝統芸能）を調査し、東北日本と西南日本の習俗の違いを説明している（宮本1967, 2005）。田楽は、今回使用した楽曲データの大半を占める盆踊唄や田植唄と密接に関係する音楽である。このことから、階層的クラスタリングの結果は、宮本の学説を間接的に裏付ける意味で重要である。

#### 10.2 日本民謡の音楽的特徴の伝播と変遷

テトラコルドを形成するパターンの傾向を $\chi^2$ -検定の残差分析を使って比較する。検定の結果、 $C_{\text{East}}$  と  $C_{\text{West}}$  の偏りは1%の水準で有意であった。 $C_{\text{East}}$  は  $C_{\text{West}}$  よりも、都節のテトラコルドを多用する傾向が強く、 $C_{\text{West}}$  は  $C_{\text{East}}$  よりも、民謡のテトラコルドと律のテトラコルドを多用する傾向が著しく強いことがわかる。ここで、東日本と西日本の地理区分とテトラコルドを使用する傾向の関連について、従来の音楽学の知見を参照しながら考察する。

西日本の旋律に多く出現する民謡のテトラコルドは、4種のうち最も古い時代から存在した音型であり、ハンガリー、中央アジア、トルコ、朝鮮半島の旋律の中にも見出すことができるという。律のテトラコルドも、中国、朝鮮半島、東南アジアの旋律の中にも見出すことができるという。一方、都節のテトラコルドが西日本ではなく東日本一帯に定着している理由を明確にする根拠は、残念ながら過去の資料から得ることができなかった。しかし、その音型は、民衆の感性・感情に訴える節回しとして浸透し、律のテトラコルドが陰旋化した種類であったとされる（小泉1977）。

あくまでも推測の域を出ないが、西日本の音組織は、近隣諸国——具体的には、中国、朝鮮半島——の影響下で形成され、東日本の音組織は、西日本の特徴が伝播し、時代の経過とともに音型が陰旋化して形成されたと考察できる。近隣諸国からの伝播と影響をより詳細に検証するためには、各国の民謡の音楽コーパスを設計し、本論文で展開してきた科学的分析を進めていく必要がある。

### 10.3 方言区画との関係

クラスタリング結果は、国立国語研究所の『日本言語地図』<sup>14</sup>や、方言地理学で言及されている方言区画（柴田1988）と一致する点が多い。柴田によれば、言語地図の境界線（言語境界線・等語線・異語線）は、地域差を明確にする手段として用いられているが、境界線と行政区画は必ずしも一致するものではなく、情報伝達の疎密によって生じた地域差に対して引かれたものであるという。方言区画では、一般に沖縄方言（琉球方言）と本土方言（内地方言）に大きく分割し、本土方言の分割については諸説あるが、いずれもクラスタリング結果が示すように日本列島を横断する区画である。

しかし、細かいレベルでの区画では一致しない点もあり、方言区画との類似点を厳密に検証するためには、楽曲データを増強し、県や令制国（旧国）単位で分類実験を試みる必要がある。これにより、音楽学者の兼常清佐が提唱した「日本民謡は日本語のアクセントが変化した日本語の一種である」という仮説（兼常1938）の検証が期待される。

## 11. まとめ

### 11.1 日本伝統音楽の音組織と地域性

本論文では、音楽コーパスを構築し、民謡の旋律に計量分析を施すことで、過去の議論・文献と比較しながら、日本民謡に内在する音楽的特徴と地域性を実証した。その成果は、次の2点にまとめられる：

1. これまで近代西洋音楽を対象とする楽曲分析・特徴抽出が主流であり、必ずしも実施されてこなかった非西洋音楽の日本民謡を分析対象として、計量的な分析方法を提案し、実現していること。

<sup>14</sup> [http://www.ninjal.ac.jp/publication/catalogue/laj\\_map/](http://www.ninjal.ac.jp/publication/catalogue/laj_map/)（第1集～第6集刊行1966-1974。PDF形式 2013年12月12日参照）

2. これまで不明瞭だった日本民謡の地域性を科学的な手順で解析し、音組織の違いから日本本土の民謡の特徴を捉えていること。

具体的には、(i) 日本民謡の音組織は小泉文夫のテトラコルドを基本とすること；(ii) 日本本土の民謡は、地理的に隣接する地域ほど小泉のテトラコルドの使用傾向が似ていること；(iii) 地域差を巨視的に捉えると、近畿地方を境界に日本列島は東西に二分され、民俗学における文化や方言地理学における方言の差異に関する区画と概ね一致すること；(iv) 西日本の民謡の特徴には、大陸や半島からの影響が考えられ、東日本の民謡の特徴には、西日本の特徴が流入し、時代の経過とともに旋律の性格が一部変化（陰旋化）したと考えられること、である。

## 11.2 今後の課題と展望

本研究で得られた知見を活かし、検討すべき問題についてまとめる。

### 11.2.1 音楽認知メカニズムの解明に向けた分析基盤の構築

大規模な楽曲データベースを構築することにより、従来の人文学研究が網羅的に実現することが期待できる。隆盛が予測される音楽認知メカニズムの解明に向けた楽曲分析を実現していくために、大規模かつ良質な楽曲データの蓄積と、それらを音楽学に適したかたちでデータベース化する技術基盤を整える必要がある。例えば、楽曲データを MusicXML<sup>15</sup> で表現し、旋律以外にも歌詞情報やコンテンツ情報をタグ付けしていく方針が挙げられる。

### 11.2.2 様々な背景要因を合わせた音楽コーパスの解析

本研究では、分析対象を日本民謡の楽曲データに限定したが、取り扱った手法や考え方は、あらゆる楽曲データに対して汎用的に実施できる。例えば、大規模な楽曲データをジャンル、時代、地域、作者など、様々な要因によって分割し、分類実験を行うことで、ジャンルに普遍的な要素、楽曲の年代的な特徴などの解明に貢献することが期待できる。

### 11.2.3 日本民謡の言語的側面の解明

民謡の発生には、言語的側面が関係しており、言語の韻律と民謡の節回しの関連性を実証することが期待される。例えば、町田佳聲が提唱した「曲型」「詞型」の観点から楽曲の計量分析を実施する方針が挙げられる。町田は『日本民謡の詞曲形態から見た時代性と地域性』の中で、民謡の曲節と歌詞の関係を同時に論じることで、民謡の形態を「曲型」「詞型」という2つの形式に分類した（上西 2009）。

国立国語研究所では、「通時コーパスの設計」プロジェクト（プロジェクトリーダー：近藤泰弘）の一環として古典資料の形態素解析を実施している。形態素解析を行うためには、基礎資料となる古典テキストの電子化が必要となる。これまでに平安時代を中心に和文（近藤 2012）、漢

<sup>15</sup> <http://www.musicxml.com/for-developers/>（2013年発表。2013年12月12日参照）

文の要素が含まれる和漢混濁文（田中 2013）、近世口語テキスト（市村ほか 2012, 2013, 河瀬ほか 2013）などの電子化および形態素解析を進めている（Ogiso et al. 2012）。

したがって、11.2.1 節で提案した歌詞情報のタグ付けにおいて、日本語コーパスとの親和性を高める形式を採用することで、歌詞の計量分析を実現させることが考えられる。そして、民謡の旋律と詞型の一致・不一致を読み解くことで、10.3 節で言及したように、昭和初期に兼常清佐が提唱した仮説を検証することが期待される。

## 参考文献

- Bühlmann, Peter (2002) Sieve bootstrap with variable length Markov chains for stationary categorical time series. *Journal of the American Statistical Association* 97 (458): 443-471.
- Cambouropoulos, Emiliios (2001) The local boundary detection model (LBDM) and its application in the study of expressive timing. *Proceedings of the International Computer Music Conference*, 290-293.
- Cao, Wenxiao, Danning Jiang, Jue Hou, Yong Qin, T. F. Zheng, and Yi Liu (2009) A phrase-level piecewise linear scaling algorithm for melody match in query-by-humming systems. *Proceedings of IEEE International Conference on Multimedia and Expo 2009*, 942-945.
- Chai, Wei and Barry Vercoe (2001) Folk music classification using hidden Markov models. *Proceedings of International Conference on Artificial Intelligence 2001*.
- Eerola, Tuomas, Topi Järvinen, Jukka Louhivuori and Petri Toivainen (2001) Statistical features and perceived similarity of folk melodies. *Music Perception* 18: 275-296.
- 藤井知昭 (1978) 『音楽』以前』(NHK ブックス 325). 東京：NHK 出版.
- 市村太郎・河瀬彰宏・小木曾智信 (2012) 「近世口語テキストの構造化とその課題」『人文科学とコンピュータ研究報告』CH96: 1-8.
- 市村太郎・河瀬彰宏・小木曾智信 (2013) 「洒落本コーパスの構造化」『第 3 回日本語学コーパスワークショップ予稿集』249-258.
- 柿木吾郎 (1969) 「構化による日本民謡旋律の比較分析法」『宮崎大学教育学部紀要, 芸能』26: 7-39.
- 柿木吾郎 (1983) 「日本民謡の音楽様式的分析」『東洋音楽研究』49: 94-108.
- 兼常清佐 (1938) 『日本の言葉と唄の構造』東京：岩波書店.
- 河瀬彰宏・市村太郎・小木曾智信 (2013) 「TEI: P5 に基づく近世口語資料の構造化とその問題点」『人文科学とコンピュータシンポジウム論文集』4: 7-12.
- 河瀬彰宏・往住彰文 (2007) 「音高推移から見た日本民謡の地方的偏差と分類」『日本感性工学会大会予稿集』(CD-ROM) 9: G31.
- Kawase, Akihiro and Akifumi Tokosumi (2007) A regional classification of Japanese folk songs: Classification based on transition probabilities of tetrachords. *Proceedings of International Conference on KANSEI Engineering and Emotion Research 2007* (CD-ROM).
- Kawase, Akihiro and Akifumi Tokosumi (2008) Estimating the structures of Japanese folk songs using a VLMC model. *Proceedings of the 6th International Conference of Cognitive Science 2008*, 87-90.
- Kawase, Akihiro and Akifumi Tokosumi (2010a) Regional classification of Japanese folk songs: Classification using cluster analysis. *Kansei Engineering International Journal* 10 (1): 19-27.
- Kawase, Akihiro and Akifumi Tokosumi (2010b) Estimating the minimum structures of musical schemas from traditional Japanese and Chinese folk songs, *Proceedings of the Kansei Engineering and Emotion Research International Conference 2010*, 1626-1634.
- 小泉文夫 (1958) 『日本伝統音楽の研究』東京：音楽之友社.
- 小泉文夫 (1977) 『日本の音』東京：青土社.
- 小島美子 (1982) 『日本音楽の古層』東京：春秋社.
- 小島美子 (1992) 「日本民謡の地域性研究に向けての試論 (2) —日本民謡の日本海側と瀬戸内海側—」『民俗音楽研究』12: 2-12.
- 近藤泰弘 (2012) 「日本語通時コーパスの設計について」『国語研プロジェクトレビュー』3(2): 84-92.
- Lerdahl, Fred and Ray S. Jackendoff (1983) *A generative theory of tonal music*. Cambridge, MA: The MIT Press.

- Manning, Christopher D. and Hinrich Schütze (1999) *Foundations of statistical natural language processing*. Cambridge, MA: The MIT Press.
- McNab, Rodger J. (1996) *Interactive applications of music transcription*. M.S. Thesis, Computer Science at the University of Waikato.
- 宮本常一 (1967) 「民俗から見た日本の東と西」『宮本常一著作集第3巻 風土と文化』81-103. 東京: 未来社.
- 宮本常一 (2005) 『日本文化の形成』東京: 講談社.
- Ogiso, Toshinobu, Mamoru Komachi, Yasuharu Den and Yuji Matsumoto (2012) UniDic for early middle Japanese. *Proceedings of the 8th International Conference on Language Resource and Evaluation (LREC)*, 911-915.
- Pearce, Marcus T., Daniel Müllensiefen and Geraint Wiggins (2008) A comparison of statistical and rule-based models of melodic segmentation. *Proceedings of the 9th International Symposium on Music Information Retrieval (ISMIR2008)*, 89-94.
- Purwins, Hendrik, Perfecto Herrera, Maarten Grachten, Amaury Hazan, Ricard Marxer and Xavier Serra (2008) Computational models of music perception and cognition I. *Physics of Life Reviews* 5: 151-168.
- Rissanen, Jorma (1983) A universal data compression system. *IEEE Transactions on Information Theory* 29 (5): 656-664.
- Ron, Dana, Yoram Singer and Naftali Tishby (1996) The power of amnesia: Learning probabilistic automata with variable memory length. *Machine Learning* 25: 117-149.
- Sachs, Curt (1943) *The rise of music in the ancient world: East and west*. New York City, NY: W.W. Norton and Company.
- 坂元慶行・石黒真木夫・北川源四郎 (1983) 『情報量統計学』東京: 共立出版.
- Schaffrath, Helmut (1995) The Essen folksong collection in Kern format. [computer database]. Menlo Park, CA: CCARH.
- Shibata, Ritei (1997) Bootstrap estimate of Kullback-Leibler information for model selection. *Statistica Sinica* 7: 375-394.
- 柴田武 (1988) 『方言論』東京: 平凡社.
- Snyder, Robert (2000) *Music and memory*. Cambridge, MA: The MIT Press.
- 高際辰弥・河瀬彰宏・村井源・徃住彰文 (2013) 「明治維新以降における日本での流行歌の構造分析」『情報処理学会研究報告 音楽情報科学』2013-MUS-100, 31: 1-5.
- 田中牧郎 (2013) 「説話のパラレルコーバスの設計」『第3回日本語学コーパスワークショップ予稿集』259-268.
- Temperley, David (2001) *The cognition of basic musical structures*. Cambridge, MA: The MIT Press.
- Temperley, David (2007) *Music and probability*. Cambridge, MA: The MIT Press.
- Tong, Howell (1975) Determination of the order of a Markov chain by Akaike's information criterion. *Journal of Applied Probability* 12: 488-497.
- 上西律子 (2009) 「旋律分析の方法論(2)一民謡の節まわしと詞型の見かた一」『民俗音楽研究』34: 73-74.

## 資料

- 日本放送協会編 (1946) 『日本民謡大観 関東篇』東京: 日本放送出版協会.
- 日本放送協会編 (1952) 『日本民謡大観 東北篇』東京: 日本放送出版協会.
- 日本放送協会編 (1955) 『日本民謡大観 中部篇 (北陸地方)』東京: 日本放送出版協会.
- 日本放送協会編 (1960) 『日本民謡大観 中部篇 (中央高地・東海地方)』東京: 日本放送出版協会.
- 日本放送協会編 (1966) 『日本民謡大観 近畿篇』東京: 日本放送出版協会.
- 日本放送協会編 (1969) 『日本民謡大観 中国篇』東京: 日本放送出版協会.
- 日本放送協会編 (1973) 『日本民謡大観 四国篇』東京: 日本放送出版協会.
- 日本放送協会編 (1977) 『日本民謡大観 九州篇 (北部)』東京: 日本放送出版協会.
- 日本放送協会編 (1980) 『日本民謡大観 九州篇 (南部)・北海道篇』東京: 日本放送出版協会.

## 付録：小泉文夫のテトラコルド理論

テトラコルド理論とは、音楽学者の小泉文夫がフィールドワークを通して提唱した日本音楽を説明するための考え方である。小泉のテトラコルドは、完全4度音程の枠をもつ2音と、その隙間に1音を置いた3音から構成される単位のことである。本来、テトラコルドとは、古代ギリシャの豎琴リラの4弦がつくる4つの音の並びを意味する。この音の並びの両端が完全4度の音程関係をもつことから、2つの核音が作り出す枠についてもテトラコルドの用語が当てられた。完全4度音程の間隔をつくる2つの音は、核音と名付けられている。例えば、五線譜上にラとレを核音とする完全4度音程を取る。その隙間には、 $\flat$ シ、シ、ド、 $\sharp$ ドの4通りの中間音を配置する候補があることから、小泉は4種のテトラコルドを提唱し、それぞれに名称を与えた（表10）。また、図18は、ラとレの2音を核音（完全4度音程）とする場合に作られる4種のテトラコルドの譜例である。

表10 小泉の4種のテトラコルドとその音程関係

種類	名称	音程関係
I	民謡のテトラコルド	短3度+長2度
II	都節のテトラコルド	短2度+長3度
III	律のテトラコルド	長2度+短3度
IV	琉球のテトラコルド	長3度+短2度

完全4度      完全4度      完全4度      完全4度

短3度 長2度      短2度 長3度      長2度 短3度      長3度 短2度

民謡のテトラコルド      都節のテトラコルド      律のテトラコルド      琉球のテトラコルド

図18 小泉の4種のテトラコルドの譜例

小泉文夫の調査によれば、民謡のテトラコルドと琉球のテトラコルドは、近い傾向をもつ。沖縄の音楽には、民謡のテトラコルドの中間音を半音高めて琉球のテトラコルドに歌いなおす事例が報告されている。また、都節のテトラコルドと律のテトラコルドも、近い傾向をもつ。律のテトラコルドの中間音を半音低めて都節のテトラコルドに変化しているといった報告がある。この現象は「陰旋化」とよばれる。

## Structural Extraction from Large Music Corpora of Japanese Folk Songs

KAWASE Akihiro

Adjunct Researcher, Center for Corpus Studies, NINJAL

### Abstract

In this study, we focus on the melodies of endangered Japanese folk songs, which were created by anonymous nonprofessional musicians or poets, and have been orally handed down from ancient times. The goal of the study is (1) to extract the basic structures from both traditional Japanese and Chinese folk songs that represent the characteristics found within the vast majority of the music corpora, and (2) to form regional classifications based on transition probabilities of the extracted musical patterns. Chinese folk songs were chosen as a target for comparison to Japanese folk songs because, despite the considerable influence of Chinese culture on Japan, the musical characteristics of both countries differ in many aspects. Specifically, we sample 1,794 works from *Nihon Min-yo Taikan* (Anthology of Japanese Folk Songs, 1944–1993) to investigate Japanese folk songs and 1,984 folk songs from a website providing virtual musical scores for Chinese folk songs. With this data, we probabilistically created a tree structure modeling a variable length Markov chain (VLMC) in order to compare minimum transition patterns occurring with high probabilities in terms of pitch intervals. As a result, we successively classify the regions into two basic groups, Eastern and Western Japan, which correspond to geographical factors and cultural backgrounds, and also matches accent distributions observed in the Japanese language.

**Key words:** music corpora, numerical analysis, variable length Markov chain model, Tonesystem