

# 日本語歌詞からの自動作曲

嗟峨山茂樹, 中妻 啓, 深山 覚  
酒向 慎司, 西本 卓也

本稿では、任意の日本語テキストの持つ韻律に基づき、歌唱曲を自動作曲する手法について解説する。文学作品や自作の詩、ニュースやメールなど、あらゆる日本語テキストをそのまま歌詞として旋律を生成し、歌唱曲として出力する自動作曲システムは、手軽な作曲のツール、音楽の専門知識を持たない人のための作曲補助ツールとして有用であろう。さらに著作権問題の回避としても用途があろう。歌唱曲は歌詞との関連性が求められる。特に高低アクセントを持つ日本語では、発話音声にピッチの高低が付くため、歌詞を朗読する際の韻律と旋律が一致することが重要とされる。筆者らはこの点に着目し、ユーザが選択した和声、リズム、伴奏音形を拘束条件として、旋律を音高間を遷移する経路とし、韻律の上下動の制限の下で最適経路となる旋律を動的計画法により探索する問題として旋律設計を捉えた。このモデルに基づき、任意の日本語歌詞に、その韻律に一致した旋律を付ける自動作曲手法により自動作曲システム *Orpheus* を作成したので紹介する。

キーワード：自動歌唱作曲，日本語韻律，最適化問題，和声進行，リズム木

## 1. はじめに

コンピュータに人間のように歌の作曲をさせることは可能か？ もし、それができれば、彼女の誕生日に詩を作って作曲して贈り物にしたり、学校で脚本を作って作曲してミュージカルを作ったり、記憶したいことを歌にして記憶したり、社歌の歌詞を募集して作曲は自動で行ったり、著作権の問題のないオリジナル曲をビデオやホームページで挿入したり、様々な利用が考えられよう。

実は、コンピュータを用いた作曲は、現在まで様々な手法によるものがなされてきた[1]。これらは、新たな音楽の創造という「芸術」的な目的を持つものと、人間がある形式や理論に基づいて行う作曲の自動化という「工学」的目的を持つものに分類できる。

前者は、「アルゴリズム作曲」と呼ばれることが多い。これらは、新たな作曲手法（アルゴリズム）の提案により、既存の曲とはまったく異なる音楽の創造を目指している。こうした研究の例としては、初めてコンピュータを用いて作曲を行った「イリアック組曲」

[2]や、確率論に基づく作曲プログラム SMP (Stochastic Music Program) を用いて作曲された“Eonta” (I. Xenakis, 1964) などがある。コンピュータのない時代にも、モーツァルト (1756~1791) はいくつかの旋律をさいころを振って組み合わせるとい手法を用いて「音楽のさいころ遊び」という曲を作っているが、これは、手法の一部に確率による偶然性を用いたアルゴリズム作曲である。現在では、IRCAM (Institut de Recherche et Coordination Acoustique/Musique) で開発された OpenMusic, MAX など、作曲のためのアルゴリズムをグラフィカルにプログラミングできるソフトウェアもあり、プログラミング技術を持たない作曲家でも簡単にアルゴリズムを実装できる環境が整備されている。こうしたツールを利用することで乱数、カオスなどによるランダム性を用いた作曲や、センサを用いて人間の動きを取り込み、この情報をもとにリアルタイムに音楽を生成するシステムなど[3][4]、様々なアプローチの作曲システムが提案されている。

一方、工学的目的の研究では、既存の音楽理論や楽曲に基づき、人間の作曲過程を自動化する。ジャズのアドリブソロを遺伝的アルゴリズムを用いて自動生成したもの[5]や、主旋律と対旋律が協和するように厳格な規則が決められている「対位法」に基づいた楽曲生成[6]や、メロディと和声進行の関係についての理論である「和声法」を用いた自動和声づけシステム[7][8][9]やコラールを生成するシステム[10]、ジャ

さがやま しげき, なかつま けい, ふかやま さとる,  
にしもと たくや

東京大学 大学院情報理工学系研究科  
〒113-8656 文京区本郷 7-3-1

さこう しんじ

名古屋工業大学 大学院工学研究科  
〒466-8555 名古屋市御器所町

ズならばインプロヴィゼーションをするシステム[11]などが提案されている。指定したジャンルの形式に基づく曲を自動生成するシステム[12]も商業化されている。

これらに対し、本稿では、歌詞に基づく歌唱曲作曲をモデル化し、これに基づき任意の日本語テキストを歌詞とする歌唱曲を自動作曲する手法について解説し、著者らが開発した自動作曲システム *Orpheus* を紹介する。本稿は工学の立場から歌唱曲作曲を捉えるので、議論する自動作曲システムの狙いは、名曲やオリジナルティのある曲を作るのではなく、工学的に捉えることが可能な音楽の要素により、人間が聴いて不自然でない曲、音楽理論的に妥当な曲を作曲することにある。

## 2. 歌唱曲旋律への制約条件

### 2.1 音楽の要素

音楽は一般的に、和声、リズム、旋律の3要素から成り立つとされる。本稿で扱う歌唱曲では通常、歌に伴奏が付くので、歌唱曲の構成要素を、和声、リズム、旋律、およびこれを支える伴奏の4要素とする。これら4つの要素は互いに関連があり、旋律とリズム、旋律と和声、和声（と旋律）と伴奏はそれぞれ密接に関連するが、本稿では旋律を生成する以前の和声進行とリズムと伴奏音型は独立に設計可能と仮定する。

この仮定は音楽理論でも暗黙のうちに認められているようである。和声の理論である和声学では、基本的にリズムを扱わず、すなわちリズムの任意性を認めている。また、和声進行と、それを分散和音などの伴奏音型への展開とに関して、相互の制約の議論は基本的にされない。

以下では、歌唱の自動作曲に必要な要素について論じる。

### 2.2 歌詞からの制約：音符数

まず、旋律には音符はいくつ必要か？

日本語はモーラ言語である。俳句の五・七・五や短歌の五・七・五・七・七として数えるのがモーラである。モーラ (mora) とは音韻論の用語で、一定の時間的長さをもった音の分節単位とされる。英語はモーラ言語ではない。だから英語しか話さない人には「おおさか」と「おさか」の区別がなく、/osaka/と綴る。「来た」/kita/、「聞いた」/kiita/、「切った」/kiQta/の区別も難しく、真似れば物理的に真似ができるだけで、別の語であるという意識は持てない。実は、モーラ言

語はかなり特異で、諸説あるようだが、日本語以外にはタミル語やフィン語やアメリカ先住民の言語などくらいである。

似た概念としては音節という概念はどの言語にもあるが、これは音素の配列の問題であって時間の単位にはなっていない。だから、英語で俳句を作っても日本語で味わう感じは作れない。英語の場合は、歌詞に曲をつける場合の基本的な音符数は音節数である。したがって“ice cream”は/aɪs kri:m/のように2音節であり、音符は2個あれば足りる。しかし、日本語では子音で終わる閉音節は許さず、/aɪs/はまず/aɪsu/になり、さらに二重母音が存在しないから/a-i-su/と3つの音節に分かれ、それが各モーラに対応して3音符が必要になる。/kri:m/も同様で、/i:/は2モーラの長音と理解され、合計で「アイスクリーム」は7モーラの語となる。俳句や短歌では7文字と数えられることになり、音符で言えば7つの音符か、「リー」を長い音符にして6つの音符にすることになる。もし、日本語の歌で「アイスクリーム」を2音符で歌ったら、“ice cream”と聴こえて英語が混ざったような感じになるだろう。

### 2.3 リズム設計：リズム木仮説

原則的に、各モーラには少なくとも1つずつ音符が必要である。では、各音符の音価は任意であるか？

もちろんそんなことはない。音価とは、二分音符、付点四分音符、四分音符、八分音符のような楽譜上の音符の長さである。これらの連鎖はリズムを形成する。リズムは曲想を強く表す要素である一方、自由度も高い。特に、日本語の場合は、強弱アクセントを持たず、各モーラがほぼ等時的に発話されるので、旋律の上下動は韻律により強い制約を受けるが、リズムの制約は強くない。

このことは英語の場合を対比すれば、分かりやすい。例えば、“The sun was shining.”という歌詞に曲をつけるとすれば、“sun”に強拍を、“the”には弱拍を割り当てるのが自然であろう。このことは、本稿で扱う日本語とは対照的に、強弱アクセントを持つ多くの言語の自動作曲の原理となり得るが、ここでは論じない。

このため、例えば、歌の1番と2番の同じ場所で、音符数が違うためにリズムが異なっても同じ曲想を感じることがある。この2つのリズムは音符数によらずに曲想に与える同一の特徴を持っていると考えられる。本稿では、この音符数に依存しない特徴を「リズムパターン」、同じリズムパターンを持つ音符数の異なるリズムの集合を「リズム木」と呼ぶ。そして、一定のリズ

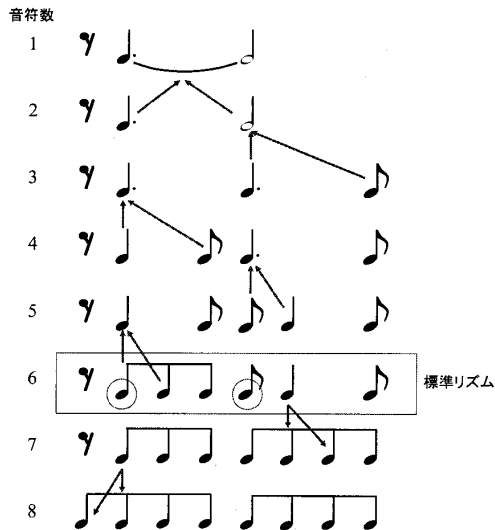


図1 リズム木構造による標準リズムからのリズム木の形成。標準リズム中で丸で示した位置がリズムパターンを表現する。標準リズムがリズム木構造に従って音符数に応じて展開されることで、この位置のリズムが保持され、一定のリズムパターンを持つリズム木が形成される。

ムパタンの下で、あるリズムが異なる音符数に展開される過程を、以下のように仮定する。

まず、あるリズムパターンから、そのリズムパターンを最もよく実現するリズムが生成される。このリズムを、「標準リズム」と呼ぶ。また、一定のリズムパタンの下で、あるリズムが異なる音符数に展開できる構造として「リズム木構造」を仮定する。リズム木構造に従い、図1のように標準リズムから音価の分割、統合が行われることで、一定のリズムパターンを持つ異なる音符数のリズムが形成される。こうして生成された異なる音符数のリズムの集合が、リズム木となる。

以上の仮説に確率モデルを導入すると、歌唱曲作曲におけるリズム設計は、曲想を基に確率的に生成されたリズムパターンからリズム木が作られ、歌詞により決まる音符数に合わせて、リズム木から使用するリズムが決定されるというモデルとなる。

リズムが決定することで、既に決定している和声と合わせ、旋律の音符が属する和音が決定する。これにより、旋律設計で用いられる各音符における音程の出現確率、および音符間における音程の遷移確率が決定する。

#### 2.4 歌詞からの制約：抑揚

音符を並べれば旋律になる。そこに言語的制約はないか？

本研究では歌詞の韻律を用いて自動作曲を行うが、音韻論で韻律とは、あるテキストを発話するときの音

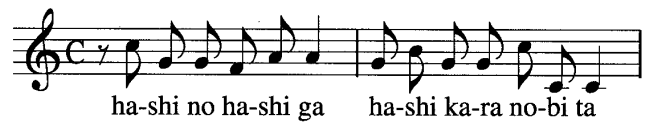


図2 イントネーションと旋律の一致の例：「箸の端が橋から伸びた」

声の、音の高さ（ピッチ）、強さ、長さ、また区切り、息継ぎなど、音色以外の幅広い特徴を指す。しかし、これらの特徴のうち韻律的特徴を強く伝えるものは、主にピッチの変化である。そこで、本稿では単に韻律と言った場合、テキストのアクセント句、アクセント型情報を指すものとする。

日本語はピッチアクセントの言語である。アクセントは語のピッチの変化であるが、長いテキストを発話するときは、テキストをいくつかのグループに分割し、そのグループにイントネーションをつけて発話する。このグループをアクセント句と呼ぶ。

関東方言（いわゆる標準語）では、一般に  $m$  モーラの単語あるいはアクセント句には 0 型から  $m$  型までの  $m+1$  種類のアクセント型があり得る。0 型はピッチの下降のないものである。他は、ピッチの下降の直前のモーラ（アクセント核）が何番目であるかにより、 $m$  型と称する。

例えば「端」「箸」「橋」はいずれも 2 モーラ語で発音は「はし」であるが、アクセント型はそれぞれ 0 型、1 型、2 型である。「端」と「橋」の相違が分かりにくいかもしれないが、それぞれに助詞「を」をつけてアクセントを視覚的に表すと、 $\underline{\text{ハ}}|\underline{\text{シ}}\text{ヲ}$ 、 $\underline{\text{ハ}}|\underline{\text{シ}}\text{ヲ}$ 、 $\underline{\text{ハ}}|\underline{\text{シ}}|\underline{\text{ヲ}}$ となる。

例えば「箸の端が橋から伸びた」というテキストは、 $\underline{\text{ハ}}|\underline{\text{シ}}\text{ノ}$   $\underline{\text{ハ}}|\underline{\text{シ}}\text{ガ}$   $\underline{\text{ハ}}|\underline{\text{シ}}|\underline{\text{カ}}\text{ラ}$   $\text{フ}|\underline{\text{ビ}}\text{タ}$ とアクセント句に分割され、イントネーションがつけられる。図2に示すような曲があるとすれば、イントネーションと旋律が一致していることが感じられるだろう。

歌唱曲の作曲では、日本語歌詞のイントネーションと旋律が一致することが要求されてきた。山田耕筰らの作曲スタイルとしても良く知られている。想像の域を出ないが、ラジオの普及し始めのころは放送局は少なく出力も低くラジオの性能が低く、録音技術がなくて繰り返し聴くこともできない状況では、一度聴いても歌詞が分かる了解性が重要で、音韻性が明瞭な歌唱法とともに、韻律旋律一致の作曲法がラジオ歌謡では重要だったのではないか。

ところが、恐らく通信技術と録音再生技術の発達で、以上の原則は必ずしも重要でなくなっているのが現代かもしれない。実際、歌詞の韻律にかなり反する作曲も行われ、それがかえって魅力になったりする。しかし、本稿で述べるような一度きりの作曲、一度きりの聴取を想定する場合等では、韻律旋律一致の原則は有効であろう。本稿ではこの原則を守ることにする。

音高（ピッチ）の上下動と、アクセントによる歌詞発話時のピッチの上下動を一致させる必要が生じる。旋律設計は、韻律の上下動を守りながら、曲想に応じてどの音域を使うか、どの程度の跳躍をするか、などを考慮して進められる。この際、音程だけでなく音価も同時に考えられる。強弱アクセント言語を歌詞とする場合と比べて、歌詞が日本語の場合ハリズムの制約は弱い。歌詞のモーラ数により音符数は決まるので、器楽曲に比べればリズム設計の自由度はかなり小さくなる。こうした制約の中で、曲想に合わせて静かなリズム、躍動的なリズム、などを決定する。

### 2.5 和声構造からの制約

では、音符の上下の動きは、韻律の上下動に従うだけで良いのか？

単旋律を作るならばまだしも、伴奏を伴った曲を作る上では、旋律の背後の和声進行が重要である。和音と、その連鎖としての和声進行に関する理論は「和声学」と呼ばれ、西洋音楽の規範であり、作曲の基本原則である。西洋音楽において多声音楽が対位法から出発し発達して、同時に鳴る音の成す響きに注意が移り、17世紀前後に次第に確立したのが和声学である。

本稿では和声学[13][14]を詳しく述べることはできないが、要点は以下の通りである。和音は、根音の上に三度音程で積み重ねた3つの音で構成される三和音や4つの音で構成される七の和音を基本とする。図3

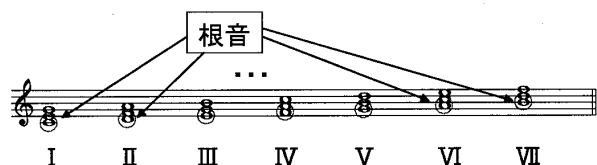


図3 各音度の3和音と和音の音度表記（ハ長調の例）  
（文献[13]より引用，改変）



図4 根音がCのときの和音のコードネーム表記の例

に示すように、当該調における根音の度数で和音に名前（ここではローマ数字）がつけられている。また、図4に示すように、根音が同じでも積み重ねる三度音程の長短によりいろいろな印象の和音の響きが生じる。和声学で最も重要なものが、これらの多種の和音を順次連結することにより生じる和声進行である。

和声進行は曲想に大きく関わる。機能和声やカデンツなどの理論があるが、本稿では、ユーザが和声進行ライブラリから選択することとする。

### 2.6 非和声音の制約

和声進行が与えられたならば、旋律はどのように作れば良いのか？

和声内音だけで旋律を作れば和声とも合致するし、基本的な制約はない。しかし、それでは滑らかな動きの旋律など、豊かな旋律表現には材料不足である。ある和音が響いているとき、その和音の構成音以外の音が使われるとき、その音を非和声音という。非和声音は、経過音、補助音（刺繍音）、倚音、繋留音、先行音、逸音などに分類されるが、このうち本稿で用いるものを以下に示す（図5）[15]。

- 経過音：和音の構成音の間を全音階，半音階的につなぐ
- 補助音：和音の構成音の2度上，または下に動

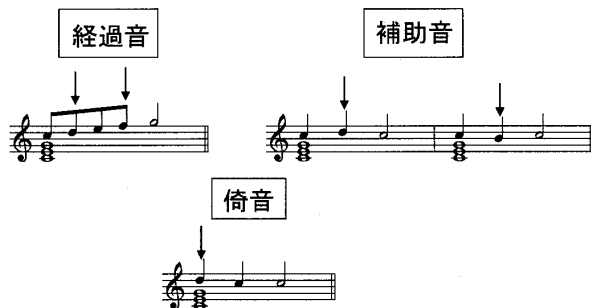


図5 非和声音の分類



図6 禁則となる連続8度，1度，5度の例（文献[13]より引用，改変）

いてもとの音に帰る音（刺繍音とも呼ばれる）

- 倚音：アクセントのある経過音

### 2.7 伴奏からの制約

伴奏は、和声、リズム、曲想の3点で旋律を補助する役割を持つ。特に和声の面で、伴奏は旋律だけでは表現しきれない和音の特色や豊かな響きを作ることができる。和声進行が与えられていて、伴奏音型を選択すれば、単純な伴奏は生成できる。そのときに旋律に対して制約は生じないか？

和声学には声部進行の禁則がある。どの声部間でも連続（平行）同度、五度、八度が禁止される。また、主に外声部間で並達（陰伏）同度、五度、八度が禁止される。最近のポピュラー曲ではこれらの禁則を守らず、聴き苦しい曲が散見される。せめて外声部の禁則違反は目立つので守りたい。その場合は、伴奏の低音進行に配慮してそれとの間で声部進行の禁則を犯さないようにするという制約が生まれる。

### 2.8 音域と跳躍の制約

作曲では、旋律の音域や跳躍の度も曲想を表現する要素として考慮される。明るい曲では跳躍が多く使われ、音域も高めになる。一方、暗い曲、静かな曲では、低めの音域で順次進行を多く用い滑らかな旋律にすることが多い。また、同じ曲の中でも、歌詞の強調したい部分、盛り上げたい部分では音域を高くして大きな跳躍を多用し、それ以外の部分は音域を低く、順次進行や比較的小さな跳躍を使う方法は、歌唱曲の作曲では常套手段である。

音域や跳躍の度数は、和声、リズム、伴奏と同様に曲想に基づき確率的に決定されるものであるが、和声、リズムと異なり、旋律の音程が決定して初めて実現される。このため、こういった音域が出現しやすいかという音程の出現確率、どのくらいの度数の跳躍が起こりやすいかという音程の遷移確率の2つの確率により、旋律設計に曲想を表現する音域、跳躍の度数が反映される。

### 2.9 動的計画法を用いた旋律設計

以上の議論から、曲想を基にユーザが和声、リズム、伴奏音型を独立に選択し、そこから旋律および伴奏を設計するという順序で歌唱曲作曲のプロセスが考えられる。

旋律を設計する声域と、旋律のリズムが与えられたとき、旋律が取りうる音程の候補と、その候補の中から使用する音程を選択するタイミングが決定される。このとき、旋律は図7に示されるように、音程とリズム

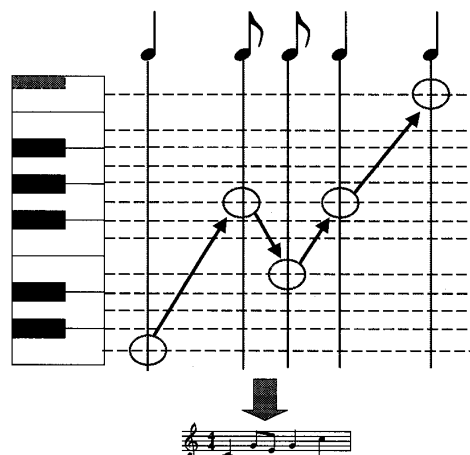


図7 旋律の経路の例。1オクターブ内で旋律を設計することを考えると、旋律の取りうる音程の候補は13音になる。図上部のようにリズムが与えられると、13音の候補を選択するタイミングが決定する。音程とリズムから形成される格子点上を図中のように遷移する経路は、図下部に示した旋律となる。

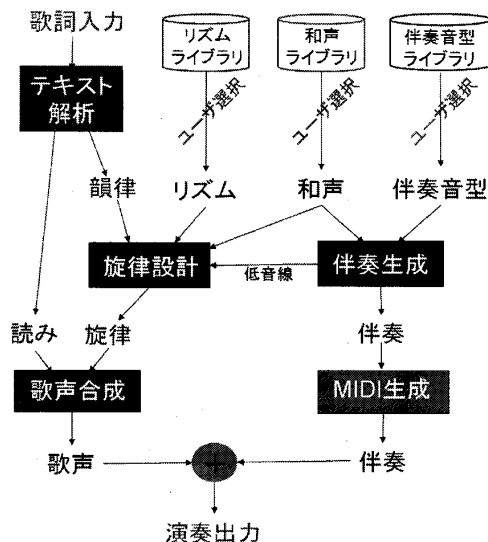


図8 Orpheus のシステム構成

ムからなる格子点上をリズムの進行に従って遷移する経路と捉えることができる。

この中で確率最大の経路を探索する問題を考えよう。その意味は、ある作曲家が以上に述べた様々な制約を尊重し、歌詞の韻律を最大限生かして作曲を行うならば、最も高い確率で作られる曲はどのような曲か、と喩えられよう。

以上の格子点の確率最大経路探索問題を解くには、単純な全探索では  $N$  個の音符がそれぞれ取りうる音高候補が  $M$  個のとき、 $M^N$  通りの経路について確率の積の比較であり、 $N$  が大きい場合は計算量が膨大となるが、動的計画法[16]を用いれば音符数  $N$  に比



図9 和声と伴奏構成音の例 (ハ長調)

例する計算量で効率的に解くことができる。

### 3. 自動作曲システム Orpheus

#### 3.1 ユーザによる入力と選択

以上に述べた作曲の原理に基づき、自動作曲システム Orpheus を構築した[17]。処理の流れを、図8に示す。

ユーザは、漢字かな混じりの日本語テキストを入力するとともに、自分が望む曲想から、(1)リズム木、(2)和声進行、(3)伴奏音型、(4)伴奏音色、(5)合成音声の音色(男/女)の選択肢から選択する。

図9に和声進行の一例を示す。これはJ. PachelbelのKanonからの借用である。

#### 3.2 テキスト解析

テキスト解析部では、入力された日本語テキストに対し、テキスト朗読時の読み、韻律を解析する。

解析ツールとして、Galatea Project[18]のテキスト音声合成システム、“GalateaTalk”を利用した。GalateaTalkは入力された日本語テキストの韻律情報を解析し、朗読音声の合成を行うシステムである。本研究では、この朗読音声合成のための韻律情報を利用した。

解析されたテキストは、2小節を単位とする節に分割される。

#### 3.3 リズムの決定

以上の処理により、各節の音符数が決定するので、これを基にユーザが選択したリズム木から音符数が合致するリズムパターンを抽出する。リズム木の一例は図1に示したが、節の含む音符数に応じたリズムが記述されている。こうしたリズム木は、人手で作成したものであるが、音楽データからの自動生成も可能であり、今後の研究テーマである。

#### 3.4 伴奏生成

伴奏は、和声とともに与えられる伴奏構成音と、ライブラリからユーザが選択した伴奏音型を用いて生成される。生成された伴奏は、テキストベースで楽譜記述言語として出力され、楽譜出力、MIDI出力を行う。

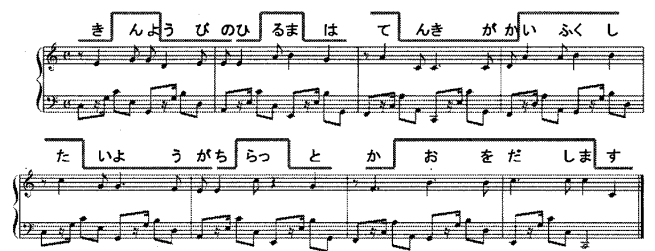


図10 システムの出力例

#### 3.5 旋律設計

旋律設計において、本システムで最も重要となるのは韻律と旋律の上下動の関係である。リズムが決定することで、音程以外の旋律の要素は決定されるため、各音符の遷移の方向を韻律に従って決定した。入力歌詞を朗読する際に、ピッチの変化が起きる位置では上行、または下降音型に、それ以外のピッチがあまり変化しない部分では、平行音型としている。なお、平行音型の部分でも、和声の境界などでは場合により音程を変える必要がある。そのため、発話において不自然でないと思われる条件下でピッチが上下動することを許し、上下動の方向をアクセント句内の位置により制限した。また、アクセント句の境界の音型は、上下動をどちらも許す平行音型とした。

音域については、歌い手の声域をA3-E5の範囲とし、この範囲内で旋律設計を行ったが、この声域の中心から離れるほどその音の出現確率が段階的に小さくなるような確率を設定し、これにより音域の制御を行った。ただし、曲想に応じて音域を変化させることはせず、常にこの確率を用いて音域を制御した。

跳躍音程の制御では、旋律の上行、下降音型の部分では、曲想に応じて跳躍の多い、激しい曲となるか、順次進行の多い、ゆったりとした曲になるかをユーザが選択するようにした。

本システムでは、非和声音として当面は経過音、補助音(刺繍音)、倚音を認めている。これらはすべて、非和声音の出現時に跳躍進行を禁止するという一つのルールにより実現できる。また、調の音階に含まれない非和声音の出現は禁止した。

和声学の禁則として、外声間で連続八度、連続同度、連続五度、並達八度、並達五度、並達同度を禁止している。

以上の制約および確率の下で、2.9節で述べたように動的計画法を用いて旋律を設計する。

音程の出現、遷移について禁止としたものは、出現確率、遷移確率を0とすることで実現した。

表1 作曲例に対する評価結果

	A	B	C	D	E
評価1)	27例	17例	5例	6例	4例
評価2)	15例	22例	17例	2例	3例

### 3.5.1 統合出力と音響出力

本システムで生成された旋律および伴奏は、楽譜出力とMIDI出力される。また、酒向ら[19]による歌声合成システムを用いて旋律の歌声を合成し、これと伴奏を合成することで歌声と伴奏による自動演奏出力を行う。

## 4. Orpheusの自動作曲結果と評価

### 4.1 自動作曲結果例

小節やニュースなどから抽出した歌詞入力に対して、和声、リズム、伴奏、跳躍などの条件を変えながら、出力を得た。出力結果の例を図10に示す。

### 4.2 専門家による作例評価

小節やニュースなどから抽出した歌詞入力に対して、和声、リズム、伴奏、跳躍などの条件を変えながら、出力を得た。これらの結果について、作曲の条件を様々に変えた場合も韻律に従った旋律が生成されることが確認できた。また、禁則、非和声音については実験条件で設定したものに関する違反はなかった。

本システムの自動作曲システムとしての有用性を検証する上では、作曲される曲がどれだけ自然であるかを評価する必要がある。しかし、こうした評価手法については確立されていない。そこで、本実験ではこの点について、作曲を勉強している学生の作例を教師が採点する状況を想定し、システムの出力結果について、和声学や作曲を教えている作曲家1名に、下記の2項目についてAからEの5段階による評価を依頼した。

- 1) システムの設計で設定していないものも含めて、禁則、非和声音の扱いなど音楽理論上の逸脱はないか
- 2) 1)の評価を含まず、総合的に曲が音楽的かどうか

出力結果81例について、作曲家による評価を行った。評価の結果を表1に示す。DやEなど評価の低かったサンプルの理由として、1)は、間接連続5度、導音が主音に進行しない、など、2)は、旋律に大きな跳躍が多い、不自然な動きがある、などであった。

この評価結果では概ね良好な評価が得られており、特に、2つの評価項目どちらについても半数以上の例

が最高評価のAから平均のCの間の評価を受けており、このシステムから重大な禁則が存在する、あるいは、著しく不自然である曲はほとんど生成されないことがわかった。

## 5. 結論

本稿では、歌詞の韻律を旋律設計に反映した歌唱曲の自動作曲手法について述べた。歌唱曲作曲を、リズム、和声、伴奏が独立に組み合わせ可能との仮説により、これを基に旋律が設計される順序でモデル化した。旋律設計においては、旋律を音程間を遷移する経路と捉え、韻律による旋律の上下動に従い、リズム、和声、伴奏の各要素から導かれる音程の出現確率、遷移確率を最大化する経路探索問題と設定し、動的計画法により解いた。

この手法をもとに任意の日本語歌詞を入力とし、ユーザーに選択された和声、リズム、伴奏音型を制約条件として歌唱曲の自動作曲を行い、伴奏と歌声の自動演奏出力を行うシステムOrpheusを構築した。自動作曲結果に対して作曲家による評価を実施し、自動作曲システムとしての有用性を確認した。

すでに、以上に述べたシステムは非和声音の扱いや長い曲の構成やwebベースの試行サービスなど、かなり拡張されており[20]、以下のURLで試行できる。  
<http://orpheus.hil.t.u-tokyo.ac.jp/>

## 参考文献

- [1] G. Papadopoulos and G. Wiggins: AI Methods for Algorithmic Composition: A Survey, a Critical View and Future Prospects, in *Proceedings of AISB '99 Symposium on Musical Creativity*, pp. 110-117, 1999.
- [2] L. Hiller and L. Isaacson: Musical Composition with a High-Speed Digital Computer, in *Machines Models of Music*, M. Schwanauer and D. Levitt Eds, MIT Press, pp. 9-21, Reprint of original article in *Journal of Audio Engineering Society*, 1958.
- [3] C. Roads: コンピュータ音楽 (青柳, 小坂, 平田, 堀内訳・監修), 東京電機大学出版局, 2001.
- [4] 長嶋洋一, 橋本周司, 平賀譲, 平田圭二: コンピュータと音楽の世界—基礎からフロンティアまで, 共立出版, 1999.
- [5] J. A. Biles: Genjam: A Genetic Algorithm for Generating Jazz Solos, *Proceeding of 1994 International Computer Music Conference*, 1994.
- [6] W. Shottstaedt: Automatic counterpoint, In M.

- Mathews and J. R. Pierce, *Current Directions in Computer Music Research*, pp. 225-262, Cambridge, Massachusetts, The MIT Press, 1989.
- [7] R. Ramirez and J. Peralta: A constraint-based melody harmonizer, *ECAI '98 Workshop on Constraints and Artistic Applications*, Brighton, 1998.
- [8] F. Pachet and P. Roy: Musical harmonization with constraints: A survey, *Constraints Journal*, 6(1): 7-19, 2001.
- [9] M. Baroni, R. Brunetti, L. Callegari and C. Jacobini: A grammar for melody: relationships between melody and harmony, In M. Baroni and L. Callegari, eds. *Musical Grammars and Computer Analysis*, Florence: L. Olshki, pp. 201-218, 1984.
- [10] M. Allan and C. Williams: Harmonizing Chorales by Probabilistic Inference, *Proceedings of Advances in Neural Information Processing Systems*, 17, 2005.
- [11] D. Levitt: A melody description system for jazz improvisation, M. S. Thesis, Artificial Intelligence Laboratory, Chambridge, Massachusetts, 1981.
- [12] BAND-IN-A-BOX, <http://www.pgmusic.com/>
- [13] 島岡譲: 和声と楽式のアナリゼ, pp. 8-9, pp. 18-20, pp. 37-39, 音楽之友社, 1964.
- [14] 長谷川良夫: 作曲法教程上巻, pp. 68-89, 音楽之友社, 1950.
- [15] 門馬直美: 音楽の理論, 音楽之友社, 1955.
- [16] R. E. Bellman: *Dynamic Programming*, Princeton University Press, Princeton, New Jersey, 1957.
- [17] 中妻啓, 酒向慎司, 小野順貴, 嵯峨山茂樹: 歌詞の韻律を用いた自動作曲, 日本音響学会春季研究発表会講演論文集, pp. 739-740, 2007.
- [18] Galatea Project, <http://hil.t.u-tokyo.ac.jp/~galatea/>
- [19] 酒向慎司, 宮島千代美, 徳田恵一, 北村正: 隠れマルコフモデルに基づいた歌声合成システム, 情報処理学会論文誌, Vol. 45, No. 3, pp. 719-727, 2004.
- [20] 深山覚, 中妻啓, 米林裕一郎, 酒向慎司, 西本卓也, 小野順貴, 嵯峨山茂樹: Orpheus 歌詞の韻律に基づいた自動作曲システム, 情報処理学会研究報告, 2008-MUS-76, pp. 179-184, 2008.