

Orpheus: 歌詞の韻律に基づいた自動作曲システム

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[†]

概要：本稿では歌詞の韻律を用いた歌唱曲の新しい自動作曲手法を提案する。旋律を音の経路と捉え作曲を経路探索問題として定式化することで、任意の日本語の歌詞を用いた歌唱曲の自動作曲が、歌詞の韻律に基づく制約条件下での最尤経路探索問題を解くことで実現できることを示す。さらにこの作曲原理を実装した自動作曲システム”Orpheus”を用いて実際に楽曲生成を行い、作曲家による生成された楽曲に対する評価を踏まえて、今回の手法によって妥当な音楽性をもった歌唱曲が生成されたことを検証する。

Orpheus: A System for Prosody-based Automatic Song Composition from the Lyrics

SATORU FUKAYAMA,[†] KEI NAKATSUMA,[†]
YUICHIRO YONEBAYASHI,[†] SHINJI SAKO,^{††}
TAKUYA NISHIMOTO,[†] NOBUTAKA ONO[†]
and SHIGEKI SAGAYAMA[†]

In this paper, we discuss a new algorithm for automatic song composition and introduce our new composition system named ”Orpheus”. We show that composing melody on Japanese lyrics can be done automatically by considering musical composition task as an optimal-path search problem under constraints of the upward and downward pitch motions given from the prosody of the lyrics. Valuation on the results generated by ”Orpheus” by a musical composer is also reported, which indicates that our new system can compose a song with a proper degree of musicality.

1. はじめに

自動作曲はその目的によって芸術目的のものと工学的目的のものに分類することができる。芸術目的の自動作曲は、自動作曲によって生まれる新しい音楽の創造や、もしくは新しい作曲行為のコンセプトを提案するために行われる。一方、工学的目的の場合は、既存のスタイルの楽曲の自動生成や、人間の作曲過程の計算論的モデリングなど、多くの場合コンピュータによって人間の作曲を再現するために行われる。本研究で扱う自動作曲は後者の工学的目的での自動作曲である。

芸術的目的での自動作曲の先行事例としては、良く

知られているように W.A. モーツァルト (1756-1791) が幾つかの旋律をさいころを振って組み合わせるとい手法で「音楽のさいころ遊び」という曲を作曲している。コンピュータが発明され音楽を処理させるようになると、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) などが登場した¹⁾。

最近では、IRCAM (Institut de Recherche et Coordination Acoustique/Musique²⁾) で開発された OpenMusic, MAX などをはじめとした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が広く用いられ、乱数やカオスなどのランダム性を使った作曲や、マイクロフォンや様々なセンサにより音や人間の動きを取り込んでリアルタイムに音楽を生成するインタラクティブな楽曲生成³⁾ など、様々な自動作曲システムが盛んに開発されている。

これらの先行事例では確率的手法やコンピュータを用いることにより、偶然性をハイライトした音楽や人

[†] 東京大学大学院情報理工学系研究科

Graduate School of Information Science and Technology, the University of Tokyo

^{††} 名古屋工業大学大学院工学研究科

Graduate School of Engineering, Nagoya Institute of Technology

手ではとても実現出来なかった音楽が生成可能となったことが有意義であると考えられる。その観点から見れば、これらの自動作曲システムの評価は作曲のための手法そのものよりも、結果として生まれた楽曲の芸術性や作曲行為についてのコンセプトの新奇性に対してより重きをおいてなされるべきである。

他方、工学的目的での自動作曲の先行事例としては、ジャズのアドリブをソロを遺伝的アルゴリズムを用いて自動生成する手法⁴⁾や、バッハのコラールを学習により確率モデル化し、このモデルを用いて未知の曲の予測を行う⁵⁾、などの研究が行われている。後者のシステムは曲の予測を行うためのものではあるが、同時にバッハ風の曲を自動生成するシステムとも捉えることができる。

本研究と同じく歌詞を入力として歌唱曲を作曲する工学的自動作曲の先行研究としては「歌詞からラララ」⁶⁾が挙げられる。このシステムでは歌詞の係り受け構造を和声構造に置き換えることで、和声進行の開始、発展、終止といった特徴と歌詞の統語論的特徴を一致させつつ作曲できる。この手法は、歌詞の内容の区切りと和声進行の区切りを一致させることで、作曲家が曲の構造に歌詞の構造を反映させる過程をモデル化したものと言えよう。しかしながら歌詞の内容を一番ニュートラルに伝えられる歌唱パートの旋律についても和声設計と同様に統語論的特徴を用いて作曲しているため、歌詞の聞き取りやすさや歌詞に対する旋律の自然さの面で課題が残る。なぜならば次章でも述べるように、歌詞が聞き取りやすい自然な旋律を作曲するためには、歌詞の音韻論的特徴を旋律に反映する必要があるからである。

そこで本研究では、歌詞の統語論的性質の代わりに歌詞の音韻論的性質である「歌詞の韻律」を用いた自動作曲システムを設計、実装する。

2. 韻律に基づいた歌唱曲作曲モデル

本節では、歌唱曲作曲において歌詞の韻律の果たす役割の重要性を明確にした上で、本研究の手法である、歌詞の韻律を用いた歌唱曲作曲のモデルについて述べる。さらに、このモデルに基づいて実際に旋律を生成する際に使う、動的計画法を用いた旋律の作曲方法についても議論する。

2.1 歌唱曲作曲における歌詞の韻律の重要性

韻律とは、音声発話時の音色以外の全ての要素を含む幅広い特徴のことをいう。そしてこれらの特徴は、通常の発話において、聞き手にわかりやすく発話内容を伝えるための特徴として重要であり、特に歌唱曲作

曲ではアクセントが重要視される。例えば日本語を歌詞とする曲では、「箸」と「橋」、「切る」と「着る」のようにアクセントの高低を変えると意味が異なるものを言葉本来のアクセントと異なる高低アクセントで歌うと、聞き手は歌詞の意味の理解が困難となる。したがって曲を聴く人に歌詞の内容を容易に聞き取らせるためには、歌詞の韻律に従った音の上下による旋律を歌詞に対して作曲することが重要であるといえる。

確かに、現代の歌唱曲全般を見渡せば、歌唱旋律の上下が発話時の韻律に必ずしも従わない楽曲も多い、という反論が予想される。しかし、これは楽曲を繰り返し聴くことや、歌詞テキストが見られることが想定されていたりして、歌詞の了解性よりも音楽的な魅力や歌詞と旋律のミスマッチのおもしろさが優先される場合であるとも考えられる。一方、過去のAMラジオ放送のように帯域が狭くノイズが多い歌声に関しては、歌詞の了解性の重要性は大きかった。このように歌詞の韻律を旋律に反映させることは、今もなお効果的で正統的な作曲法とされている⁸⁾。本システムの応用例である天気予報を歌にする場合など、一度しか聴かれないが歌詞内容が重要であるような自動作曲は、この原則に則すべきであると考えられる。

2.2 韻律に基づいた歌唱曲作曲モデル

次に韻律に基づいた歌唱曲作曲モデルについて述べる。

先に述べたように、韻律は旋律の音の上下に制約を加えるので、我々の求めるモデルは、生成される楽曲の旋律が韻律に逆らわずに作曲されるような作曲モデルである。

ところで、楽曲を構成する重要な要素として、この「旋律」の他に「和声」と「リズム」がある。実際の作曲では、これら3つの要素が互いに複雑に影響を及ぼしながら曲が生成されていると想像されるが、本研究ではこの複雑な作曲過程を、単純化のため、作曲者が作りたい曲の曲想に基づいて特に楽曲に用いたい和声進行と旋律に用いたいリズムの特徴があらかじめ決めておき、その上で日本語の歌詞に対して旋律を作曲する、という場合についてのみ考えモデル化することとする。

したがって曲想をもとに和声進行、リズムをあらかじめ設計し、そこから韻律による制限下で旋律を作曲するという順序立ての歌唱曲作曲モデルを考える。図1にこの歌唱曲作曲モデルの概略を示す。以下プロセスの順を追って、和声の生成、リズムの生成、旋律の設計について説明する。

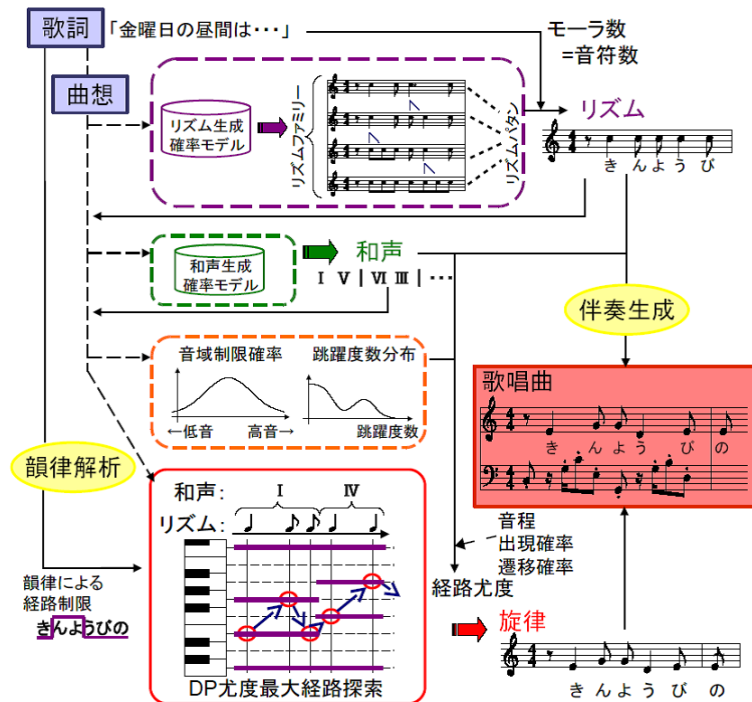


図 1 歌唱曲作曲モデル

2.2.1 和声進行の生成

和声進行は楽曲のスタイルや曲想を強く表す要素である。そこであらかじめ特定の曲想を表すような和声進行の常套句をコーパスとして準備しておき場合によって使い分ける、もしくは「明るさ度合」などの曲想を近似して表すパラメータを準備しておき、そのパラメータに基づいて調整された和音間の遷移確率と和声学の理論に基づいて曲想を反映した和声進行を自動生成する、などの方法が考えられる。

2.2.2 リズムの生成

和声進行と並んでリズムも曲想を強く表す要素である一方、自由度も高い。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。例えばサザンオールスターズの「いとしのエリー」の場合、1番と2番のそれぞれ冒頭二小節のフレーズに含まれる音符数は18個と21個で異なっているが、これらそれぞれの歌詞に付けられた旋律から得る印象は似たものとなっている。すなわちこの2つのリズムは音符数の変動に依らず曲想を表す同一の特徴を持っていると考えられる。本研究では、この音符数の変動に依らない特徴を「リズムパタン」、その特徴を表すリズムを「標準リズム」、同じリズムパタンをもつリズムの集合を「リズムファミリー」と呼ぶ。そして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を「リズム木構造」と定義し、これに基づき標準リズムからの音価の分割、統合が行われ、リズムファミリーが構成されるものと仮定する(図2)。

このような仮説に基づくと、歌唱曲作曲におけるリズム設計は、曲想を基に生成されたリズムパタンからリズムファミリーが作られ、歌詞により決まる音符数に合わせて、リズムファミリーから使用するリズムが決定されるというモデルとなる。

2.2.3 歌唱旋律の作曲

旋律は横軸時間、縦軸音高の二次元平面上での遷移経路であると捉えることができる(図3)。この遷移経路には、和声進行やリズム、歌詞の韻律によって様々な条件が課され、それらの制約は音の出現確率や遷移確率といった形で表すことができる。

例えば和声進行からは時刻毎の和音が決まり、その和音とその楽曲の調からその時刻での旋律のとれる音高には制約が課されるが、これは音高の出現確率として表現できる。また同様にリズムの制約は旋律各音の時間軸上の位置に対する出現確率がとして表され、歌詞の韻律の上下からの制約は次の音へ遷移するときに行きやすいか下行しやすいかについての遷移確率で表せる。

さらに作曲においては、旋律の音域や跳躍の度合も

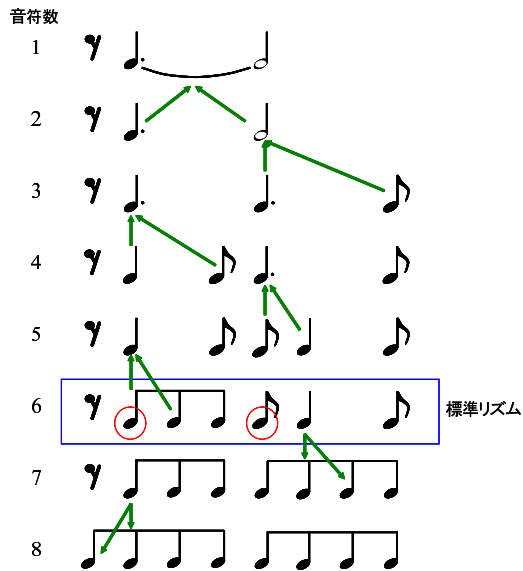


図2 リズム木構造による標準リズムからのリズムファミリーの形成。標準リズム中で赤い丸で示した位置がリズムパターンを表現する。標準リズムがリズム木構造に従って音符数に応じて展開されることで、この位置のリズムが保持され、一定のリズムパターンを持つリズムファミリーが形成される。

曲想を表現する要素として考慮される。これらも同様に、音域は音高の出現確率、跳躍は音高の遷移確率として旋律の遷移経路に制約を与える。また、歌唱曲では、歌い手の声域や技量によってこれら音域や跳躍の度合は制限を受けるため、たとえば歌い手が女声か男声か、歌手なのか一般人なのかなどに基づいて、歌いやすいように音域を狭く制限したり、歌いにくい音程の出現を制限するなどの考慮が必要である。

以上述べたような音の出現確率、遷移確率をかけ合わせることで、各音程、経路に対しての尤度が計算できる。よって、歌唱曲の旋律設計は、考えられる全ての旋律の経路のうち、歌詞の韻律の上下動を満たし、音楽理論的な逸脱をおこさない制限の下で、尤度最大の経路を探索する問題となる。これは、各経路に確率重みと、韻律によるペナルティをつけた動的計画法 (DP) の尤度最大経路探索問題へ帰着できる。

2.3 動的計画法を用いた旋律の作曲手法

最後に、動的計画法を用いた旋律の具体的な設計方法を示す。

旋律の構成音が時刻 i で音高が n (MIDI ノート番号, $1 \leq n \leq N$) であるとき、時間と音高による二次元平面上で点 (i, n) と表す。また旋律が (i, n) から $(i + 1, m)$ に遷移するとき、この遷移経路についての確率重みを $c_i(n, m)$ とする。この確率重みは和声進

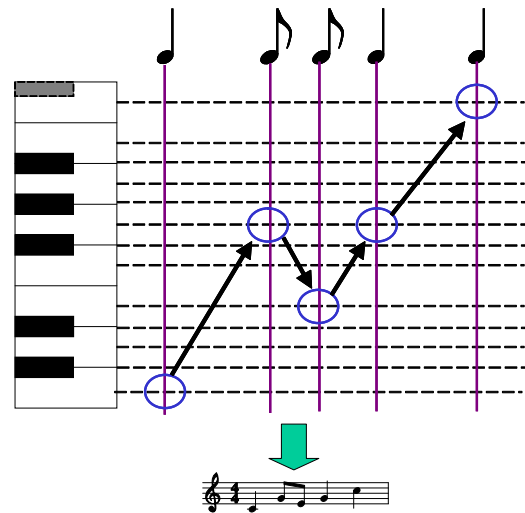


図3 旋律の経路の例。1 オクターブ内で旋律を設計することを考えると、旋律の取りうる音程の候補は 13 音になる。図上部のようにリズムが与えられると、13 音の候補を選択するタイミングが決定する。音程とリズムから形成される格子点上を図中のように遷移する経路は、図下部に示した旋律となる。

行、リズム、歌詞の韻律によって算出される。

さて時刻 1 から時刻 i までの間で旋律の最尤経路が求まったとして、この経路上の時刻 i での音高を $\text{path}(i)$ と表すことにする。この経路の最後の音の音高が m であるときのこの経路の実現確率 $P_i(m)$ は、以下のように時刻 $i - 1$ の情報を用いて再帰的に書くことができる。

$$P_i(m) = \max_{n=1, \dots, N} (P_{i-1}(n) \cdot c_{i-1}(n, m)) \quad (1)$$

そこで、全ての時刻と音高 (i, m) において、

$$n_{\max} = \operatorname{argmax}_n (P_{i-1}(n) \cdot c_{i-1}(n, m)) \quad (2)$$

を計算し、その結果を各 (i, m) と組で保存しておく。すると最尤経路の端 $(i, \text{path}(i))$ からその点と組で保存されている n_{\max} を用いて $(i - 1, n_{\max})$ にバックトラックすることを繰り返し、旋律の最尤経路全体を求めることができる。

3. 自動作曲システム Orpheus の実装

以上のような歌唱曲作曲モデルを実装して、任意の漢字仮名混じり文の歌詞入力から歌唱曲を作曲し、伴奏つき合成音声によって歌う自動作曲システム “Orpheus” を開発した。

これは歌唱曲の作曲システムに入力歌詞の読みと韻律の解析エンジンと合成音声による歌声生成エンジンを接続したものである。前者は日本語自動読み上げソ

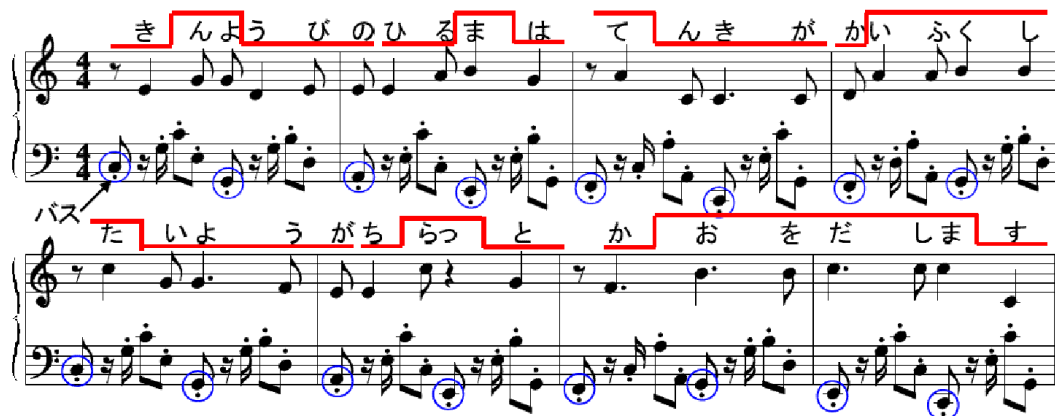


図4 Orpheusによる作曲例
歌詞の韻律にそって旋律が作曲されていることが見て取れる

フト“Galatea-Talk”⁹⁾のテキスト解析によって行い、後者の合成音声による歌声生成は“hts-engine”によって行っている¹⁰⁾。

また、現在は、和声進行、リズム木構造を曲想から自動で生成するのではなく、システムのユーザが作りた曲の曲想に合わせてライブラリから選ぶようになっている。

このOrpheusはcgiプログラムとしても実装され¹¹⁾、web上で公開されている。

4. 出力結果と評価

Orpheusに小節やニュースなどから抽出した歌詞を入力し、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。生成された楽曲の例を図4に示す。

これらの結果について、作曲条件を様々に変えた場合に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システム設計段階で設定したものに関しては違反がなかった。

次に59例の出力結果について、音楽大学で和声学および作曲法を教授している作曲家に評価を依頼した。評価項目は、1) 音楽理論上の逸脱はないか、2) 楽曲に音楽性を感じられるか、の2つで、五段階評価によりあたかも作曲専攻でない学生の作品を見るような想定で採点が行われた。結果を表1に示す。DやEなどの低い評価を受けた出力結果には、導音が主音に上行しない、旋律に大きな跳躍が多い、不自然な動きがある、などの理由が挙げられた。

この評価結果を見る限り概ね良好な評価が得られて

おり、特に、2つの評価項目のいずれ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、もしくは、著しく不自然な楽曲がほとんど生成されていないことがわかる。

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

表2 作曲例に対する評価項目

評価1)	音楽理論上の逸脱はないか
評価2)	楽曲に音楽性を感じられるか

5. 考察

韻律に基づく歌唱曲自動作曲モデルを実装したOrpheusの出力結果は、様々な歌詞に対して韻律に従った旋律を生成することができ、また著しく音楽理論を逸脱するような楽曲も生成されなかったため、概ね良好であったといえる。しかし作曲家による評価の結果、生成される旋律の自然さ、音楽性に問題を残していることも分かった。

例えば、旋律中での滑らかな進行が少なく、大きな跳躍や不自然な進行が生じている。これに対しては、現在のOrpheusは和声進行で与えられた和音の構成音によって旋律の音高の出現確率が決定されており、和音構成音でない音(非和声音)は和音構成音に比べて出現確率が低く設定されていることが理由として挙げられる。もちろんこれは和音と旋律の親和性を保つ

ための妥当な制約であるといえる。しかし経過音、刺繍音など音楽理論の中で意味付けがなされている非和声音ならば、旋律に順次進行をもたらすものとしてむしろ歓迎されることを考えると、現在の Orpheus では必要以上に非和声音の生成が抑制され、結果として跳躍の多い曲が多く生成されているともいえる。したがって音楽理論の枠内で非和声音を積極的に生成するようなアルゴリズムを新たに Orpheus に加えることが必要である。

また作曲に用いられる和声進行パターンや旋律のリズムパターンの数が限られているため、生成される楽曲全体がやや似る傾向にあり、音楽性の面で貧相である感は否めない。これに対しては、和音進行やリズム木構造の自動生成手法を新たに盛り込んだり、旋律に使われる音の出現確率や遷移確率を変更して特定の旋法の音楽を生成させたりすることで、音楽のバリエーションを増やして解決できると考えられる。

6. 結 論

本研究では、歌詞の韻律に基づく制約条件下で歌唱曲を自動生成する自動作曲モデルを提案した。また、そのモデルを用いて任意の仮名漢字混じり文を入力として自動で歌唱曲を作曲する自動作曲システム“Orpheus”を開発し、システムが設計通りの楽曲を生成することを確認した上で、楽曲の音楽性についても作曲家による評価を行い、音楽性にやや問題を残したものの、概ね良好な評価を得ることができた。

今後の課題としては、現在の“Orpheus”システムにまだ実装するに至っていない和声進行の自動生成やリズム木構造の自動生成について検討したい。

また、音楽理論の枠組内で非和声音をより積極的に生成する手法を提案し、音楽性を持った高度な旋律生成モデルを構築することや、旋律に用いられる音の出現確率や遷移確率の設定によって、今回扱った狭義の西洋音楽だけでなく、様々な旋法の音楽や民族音楽などの自動作曲モデルの構築も期待できる。

謝 辞

本手法に対しての評価と有益な助言を頂いた桐朋学園大学音楽学部作曲科金子仁美講師に感謝する。また本研究の一部は、科学技術振興機構 CREST の補助を受けて行われた。

参 考 文 献

- 1) C. Roads (青柳, 小坂, 平田, 堀内 訳・監修): コンピュータ音楽, 東京電機大学, 2001.
- 2) Institute de Recherche et Coordination Acoustique/Musique <http://www.ircam.fr/>
- 3) 長嶋洋一, 橋本周司, 平賀謙, 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.
- 4) J. A. Biles: Genjam: A Genetic Algorithm for Generating Jazz Solos, In Proceeding of 1994 International Computer Music Conference, 1994.
- 5) D. Conklin and I. Witten: Multiple Viewpoint for Music Prediction, Journal of New Music Research, Vol.24, pp.52-734, 1995.
- 6) 早川和宏, 稲垣博人, 田中一男: 歌詞からラララ-言葉から歌への自動変換-, 人工知能学会ことば工学会第3回資料, 1999.
- 7) 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.
- 8) 長谷川良夫: 作曲法教程, 音楽之友社, 1950.
- 9) Galatea Project: <http://hil.t.u-tokyo.ac.jp/galatea/>
- 10) 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.
- 11) 米林裕一郎, 中妻啓, 西本卓也, 嵯峨山茂樹: Orpheus: 歌詞の韻律を利用した Web ベース自動作曲システム, インタラクシオン 2008 予稿集, 2008.