

歌唱曲自動作曲の需要と今後 － 2年間の Orpheus 運用を通じて

深山 覚^{†1} 西本 卓也^{†1} 嵯峨山 茂樹^{†1}

本稿では日本語歌詞を入力とする自動作曲システムにおいて、ユーザから要望のあった多彩な楽器による長い曲を生成するための改良について報告する。筆者らが開発した自動作曲システムは、ユーザが入力する日本語歌詞に対して、簡単な伴奏による8小節の旋律が男性の歌声と楽譜として出力されるものであった。ユーザから得られた意見では、より長く現代のポップスのような作品も作りた要望が見受けられた。意見これに対して、歌声の種類を増やす、ドラムスや伴奏のトラック数を増やすという工夫や、構成を持った長い曲を生成できるようにすることが考えられる。しかし本システムのように、音楽の知識を必ずしも持たないユーザの作曲支援を目的とする場合、手軽に曲の雰囲気や長さを調整できるための工夫が必要である。そこで本稿では歌唱曲の自動生成における伴奏や曲構成についてのパラメタ設定をするためのインターフェイスを提案する。またウェブ上の運用を通じて得られた新たなユーザの反応を報告する。

Demands and Perspectives on Automatic Song Composition – discussions through the 2 years operation of Orpheus

SATORU FUKAYAMA,^{†1} TAKUYA NISHIMOTO^{†1}
and SHIGEKI SAGAYAMA^{†1}

In this paper, we aim to add new functions to automatic song composition from the lyrics to reflect user's comments on the system. Our previous system were able to generate only short songs in length of 8 bars with score and synthesized male singing voice, which were not enough for the demands of the users who want to generate a song with style of modern pop songs. One possible way to answer this demand is to add various types of accompaniment and singing voice. Another way is to extend the algorithm to be capable to generate longer songs with structures. However, for the purpose of composition support for amateur musicians, we should consider how we can implement an interface which users can easily handle complicated settings of parameters. Considering

these factors, we designed an interface for setting parameters of accompaniment and the structure of songs. The evaluations on the songs and the system are also reported based on the questionnaire on the web-site.

1. はじめに

計算機に楽曲を生成させる試みは計算機科学の黎明期より行われてきた¹⁾。近年では計算機を用いた斬新な楽曲生成を目的とする研究の他に、音楽の専門知識を持たない人を対象とした、手軽にオリジナルな楽曲を生成できるシステムの研究が行われており、商品化されているシステムもある²⁾。後者のようなシステムでは生成される楽曲において既存の音楽スタイルの模倣したり、ユーザの指定した雰囲気を出すことが求められる。このようなシステムは、著作権フリーの楽曲生成や作曲支援ツールとしての需要が期待される。本研究ではこの後者の立場から歌唱曲の自動作曲システムについて議論する。

歌唱曲を自動生成できる自動作曲システムを考えたとき、ユーザが各自の好みを反映させることのできる要素の一つは歌詞である。作曲支援システムとしての役割を持った自動作曲システムにおいて、ユーザの歌詞入力に応じて適切に作曲結果が変化することは重要であると考えられる。日本語の歌詞を入力とする自動作曲としては、歌詞の係り受け構造を反映した作曲アルゴリズムが提案されている³⁾。また近年では twitter でのつぶやきを歌詞として作曲するというコンセプトで、予め 5-7-5 の形にフォーマットされた歌詞にメロディをつけることで、楽曲を生成するサービスも行われている⁴⁾。

日本語の歌唱曲の場合、日本語がピッチ変化によって意味が変わる言語であるため⁵⁾、歌詞の言葉の抑揚に合わせた旋律の作曲が重要であるとされている⁶⁾。我々はこのような考察に基づいて日本語歌詞の韻律に基づく自動作曲システム “Orpheus” を提案した⁷⁾⁸⁾。このシステムは web アプリケーションとして公開されている⁹⁾。

本稿では歌詞の韻律に基づく自動作曲システムのユーザ利用履歴を通じて得られた、長い楽曲で現代のポップスのような作品を作りたいという需要を紹介する。次に充実した伴奏を持つ楽曲の制作と長い構成を持った楽曲の制作を支援するための方法について提案し、その効果を検証する。またその結果に基づいて今後の歌唱曲自動作曲システムに求められる機能

^{†1} 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

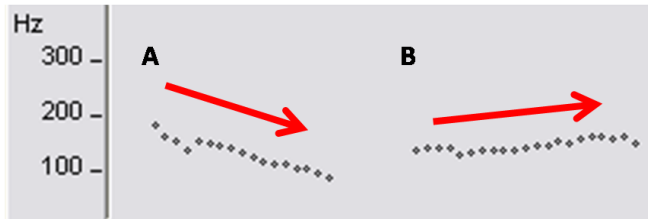


図 1 (A) 神 ka'mi と (B) 紙 kami' のピッチ変化: ピッチの変化によって語の意味が異なって理解される

と、その実現方法についても議論する。

2. 歌唱曲自動作曲システムによる構成をもつ楽曲生成の実現

本節ではまず歌詞の韻律に基づく自動作曲の原理について述べ、それが実装されたシステムの評価をユーザの利用履歴の統計に基づいて述べる。次にユーザの利用履歴とアンケートフォームにおけるコメントから、本システムの需要についての議論し、求められている新たな機能として、歌声の改良と伴奏の充実、また長い歌詞による作曲があることを述べる。最後にそれらの機能を実現するための方法を議論する。

2.1 歌詞の韻律に基づく自動作曲

2.1.1 日本語の韻律と旋律作曲における制約条件

日本語の発音では、アクセント核と呼ばれるピッチが下がる箇所があり、その位置は同じ音素によるものであっても語の意味によって変化する⁵⁾。「神」と「紙」を発音したときのピッチ変化を図 1 に示す。歌詞から旋律を作曲することは、歌詞の語にピッチを付与することであるので、歌唱曲の聞き手の歌詞の意味理解は旋律の音の上下動に影響を受ける。したがって本システムのように、入力された歌詞に基づいて作曲するシステムの場合は、入力歌詞が聞き取りやすくハイライトされた楽曲生成とするため、また入力歌詞に応じて生成楽曲が変化するためにも、歌詞のピッチ上下に従った作曲アルゴリズムとすることが妥当であると考えられる。近年のポップスなどでは、歌詞が元々持つピッチの動きと異なるピッチの上下動で旋律が作曲されているものも多く見られるが、古典的な歌唱曲の作曲においては歌詞のピッチの動きと旋律のピッチの動きを揃えることが推奨されている³⁾。

歌詞の韻律の他には、どのような制約条件が旋律に対して課せられているとみることができるだろうか。まず歌詞の意味理解に基づいた曲想が重要であると考えられる。曲想の表出のされ方は、マイナーコードを多用した暗い曲、ギャロップのリズムを含む快活な曲、と

いったように、和声進行や旋律のリズムの選択の仕方に大きく影響される。さらには楽曲のスタイルやジャンルの背景には音楽文化があり、その音楽文化の経験則が体系化されたものとしての音楽理論が、意識的にしる、無意識的にしる作曲される楽曲にそれが反映されていると考えられる。

2.1.2 制約条件を満たす旋律の探索

前節で議論したような要素の影響によって、旋律の音の現れ方にはある確率的な偏りが生じる。具体的には、和声進行と歌唱音域の制限による音高の出現確率の偏り、リズムによる音の出現タイミングの偏り、平行 8 度の禁止など音楽理論に基づく音高の遷移確率の偏りなどである。したがって、歌詞をよく反映する旋律が作曲されたとすれば、その旋律はこれらの制約を同時に一番良く満たしているのものであると考えられる。すなわち、歌唱曲の旋律生成はこれら制約によって決まる音の出現確率と遷移確率の下で、最も尤度の高い旋律を探索するという問題として定式化できる。

系列長が N であるピッチの列を MIDI ノート番号の列として $X_1^N = x_1 x_2 \cdots x_N$ と表し、各ピッチにおける制約条件の列を $Y_1^N = y_1 y_2 \cdots y_N$ とする。ここでそれぞれの y_n は、ピッチが x_n であるときの音階と調のアノテーションの付いた和音ラベル (c_n)、音の長さ (d_n)、伴奏のベースラインの MIDI ノート番号 (b_n)、ピッチアクセントの情報 (a_n) を含んでいるとする。すなわち $y_n = (c_n, d_n, b_n, a_n)$ とする。ここで Y_1^N が所与のときの条件付き確率 $P(X_1^N | Y_1^N)$ を考える。これは旋律作曲における制約条件 Y_1^N が与えられたときの旋律のピッチ X_1^N の傾向を表していると考えられる。旋律の作曲は、 Y_1^N が与えられているときに条件付き確率 $P(X_1^N | Y_1^N)$ を最大化するピッチの列 X_1^{N*} を求める問題として定式化できる。

$$X_1^{N*} = \operatorname{argmax}_{X_1^N} P(X_1^N | C_1^N). \quad (1)$$

ここで以下のように 1 次マルコフ過程として近似すると、

$$P(x_n | X_1^{n-1}, Y_1^N) \simeq P(x_n | x_{n-1}, Y_1^N). \quad (2)$$

このとき式 (1) は以下のように変形できる:

$$X_1^{N*} = \operatorname{argmax}_{X_1^N} \prod_{n=1}^N P(x_n | x_{n-1}, Y_1^N), \quad (3)$$

ここで $P(x_1 | x_0, Y_1^N) = P(x_1 | Y_1^N)$ である。 X_1^{N*} の候補数は 128^N 個あり、 N が増える

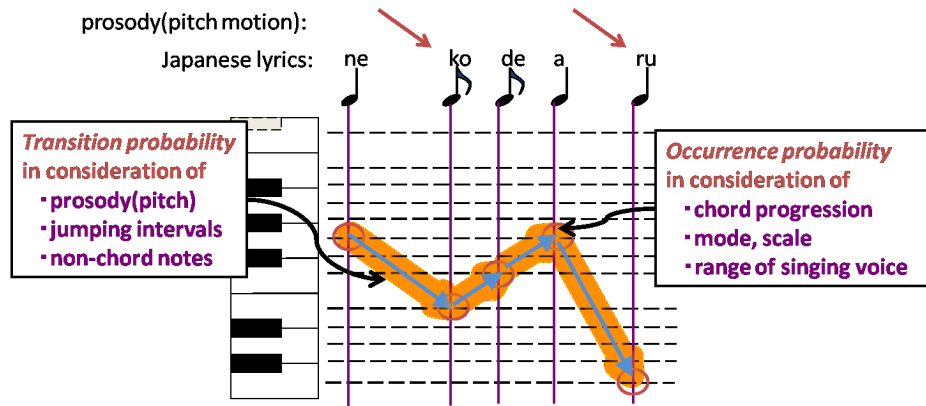


図 2 旋律の作曲は作曲条件によって与えられる音の遷移と出現の確率的偏りのもとで、確率最大の音系列を求める問題として定式化できる。これは動的計画法を用いた経路探索問題として解くことができる。

につれて探索空間が指数的に増大し全探索による解法の計算が困難になるが、動的計画法¹⁰⁾を用いることで効率的に確率最大の解を求めることができる。経路探索として定式化した旋律作曲の模式図を図 2 に示す。また自動作曲システムの構成を図 3 に示す。

2.1.3 楽曲生成例と評価

楽曲の生成例を図 4 に示す。歌詞の韻律の上下に沿って旋律が生成されていることがわかる。アルゴリズムとシステムの評価については 2 種類の評価を行った。まず、本システムは平行 8 度など作曲法に基づく制約条件などを組み込んでいるため、音楽大学で和声学の講義を受け持つ講師に生成楽曲中に禁則を含むかどうかの評価を依頼した。同時に禁則の有無に依らない楽曲自体の評価をも依頼した。これらの評価はそれぞれ 5 段階評価で行われた。評価結果を図 5 に示す。83.1%の楽曲で概ね禁則の違反などがないことが確認された。また 91.6%の楽曲が妥当であると評価された。次に web 上で楽曲とシステムについてのアンケートを実施した。初めの 1 年間の運用中に 56,000 曲が自動生成され、1378 人が 5 段階評価によるアンケートに回答した。評価結果を図 6 に示す。70.8%のユーザが生成楽曲に 3 点以上の得点を付け、システムについての評価では 84.9%のユーザが使って作曲の楽しさを感じるシステムとして、3 点以上を付けた。

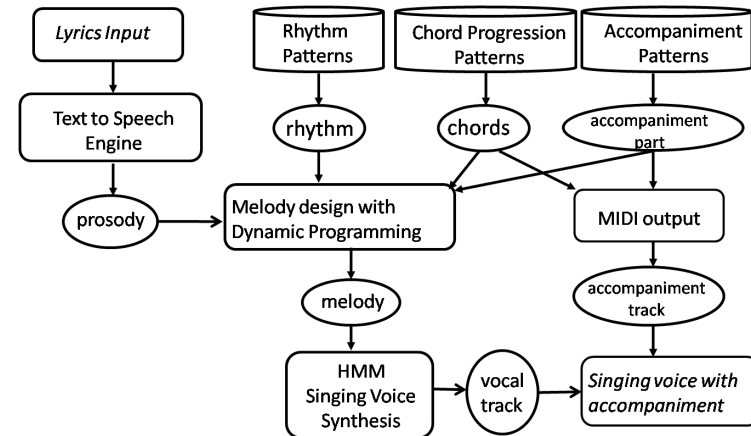


図 3 自動作曲システム Orpheus の構成: 入力された日本語歌詞は韻律の解析がされ、リズム、和声進行、伴奏のパターンの条件と合わせて旋律作曲における制約条件として扱われる。これら制約条件の下で動的計画法によって条件を最も満たす旋律が探索される。さらに生成された旋律に基づいて歌声合成がされ、伴奏の合成音と合わせて出力される。

2.2 構成を持つ長い楽曲生成への要望と実現方法

2.2.1 ユーザの反応に基づいて考えられる改良点

自動作曲システムのウェブサービスの実施中に得られたアンケートでは、より現代のポップスのような曲をも生成したいという要望が寄せられた。これのためには、以下のような改善が考えられる。

歌声のオプションの追加

ドラムス、キーボードを加えるなどの伴奏オプションの増強

長い曲や長い歌詞に基づいた作曲

1 つめの改善点については、新しい HMM 歌声合成システムにより¹¹⁾ 女声のオプションを加えることができた。また Vocaloid シリーズによって歌声合成をする試みも行った。2 つめへの対応としては楽器の種類を増強し、ドラムの有無、伴奏での和音コーラスの有無を選べるオプションを追加した。3 番めの長い曲を生成する場合には、長い歌詞に応じて単純に長い曲というだけでなく、構成を持った楽曲が生成される必要がある。

2.2.2 構成作曲機能の実現

システムの初期バージョンでは 8 小節の旋律と伴奏しか生成できなかった。しかし実際の

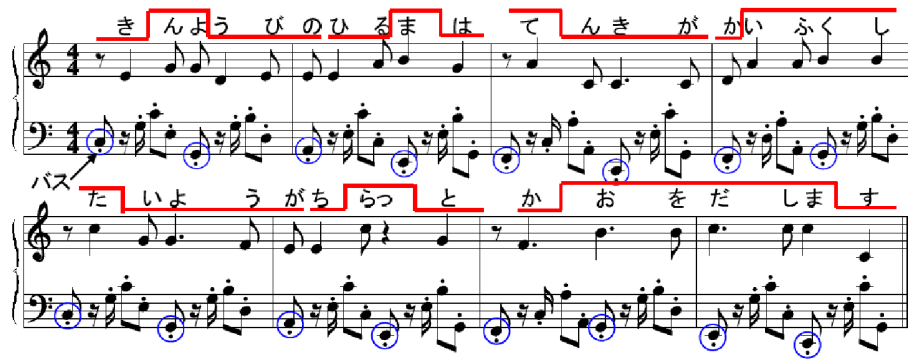


図4 「金曜日の昼間は天気回復し、太陽がチラッと顔を出します。」という天気予報を歌詞として自動作曲した例: 歌詞を読み上げるときのピッチの上下を赤線で表す。旋律の音が歌詞の韻律に沿って生成されていることが分かる。また青丸で囲んである音はバスとして扱っている音である。

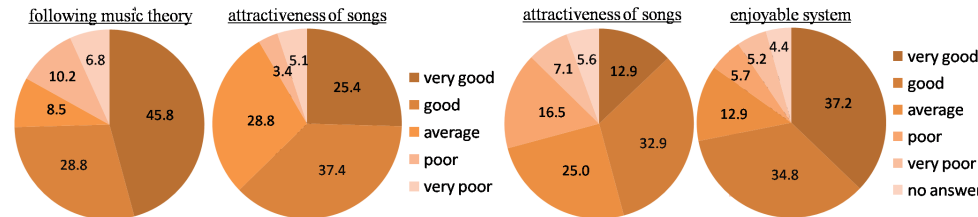


図5 音楽大学にて和声を教授する専門家による評価: 生成楽曲 59 曲に対する古典的な作曲法における禁則があるかないかによる評価 (左) と、楽曲の音楽性の評価 (右)。(円グラフ中の数字は 5 段階評価それぞれが占めるパーセンテージ)

図6 web を通じた自動作曲システムのユーザによる評価: 1378 人のユーザによる楽曲の音楽性についての評価 (左) と、システムを使って作曲の楽しさを感じられるかについての評価 (右)。(円グラフ中の数字は 5 段階評価それぞれが占めるパーセンテージ)

歌唱曲の楽曲では A-A、A-B-A、A-B-C などのように間に部分を挿み、再び初めと部位を繰り返すというような構成を持っていることが多い。このような構成を持った楽曲を生成できるようにすることで、より歌唱曲作曲システムで生成できる歌唱曲のバリエーションを増強できると考えられる。

そのような構造を持つ歌唱曲の作曲のためには、改行を元に歌詞を区切り、それぞれにおいて 8 小節分の作曲パラメタを設定して楽曲生成が行えるインターフェイスが考えられる。



図7 構成を持った歌唱曲を生成するためのインターフェイス: 改行で区切られた歌詞ごとにリズムパターン、和声進行、伴奏音型と伴奏楽器、ドラムスの有無を選ぶことができる。ここでは前半 2 つの部分の和声進行を「あきらめ」「物思いに」として暗い曲調に指定し、最後の部分の和声進行を「前向き」として暗から明へと変化する構成を作曲条件として設定している。各部位の条件設定に基づいて旋律が生成され、連結されたものが生成結果として出力される。

これをプロセスごとに書くと以下ようになる。

1. 入力歌詞から改行を抽出
2. 改行ごとに分割された歌詞で 8 小節の楽曲のオプションを設定
3. 各構成部分ごとに歌唱曲作曲を行い連結して生成楽曲とする

和声進行を接続する際に不具合が起きないように、パターンの継目で近親調へ転調するように、元々の和声進行パターンを移調できるようにした。また各部位で伴奏パターンやドラムスの有無を変化できるように、図 7 のようなマトリクス状のインターフェイスを作成した。また楽曲が生成された後においても、マトリクス状インターフェイスの画面へと戻り、部位ごとのパラメタを調整できるようにした。

3. 増強した伴奏と構成を持つ楽曲生成機能を持つシステムのユーザ評価

前節で述べた伴奏と歌声の増強と、構造生成機能を持たせたシステムを Orpheus Ver 2.0

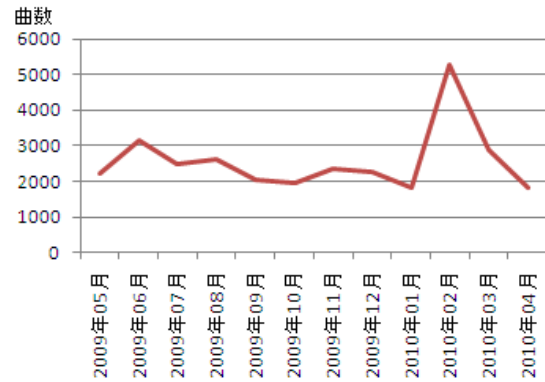


図 8 伴奏の増強と構成を持つ作曲が行える Orpheus Ver. 2.0 が公開されてからの、1ヶ月毎に自動生成された楽曲数。普段は約 2,000 曲から 3,000 曲、メディアでシステムが取り上げられた月は約 5,000 曲が自動生成された。

とし、ウェブアプリケーションとして公開した。公開後に生成された楽曲数は 30,852 曲であった。公開後の生成された楽曲生成数の履歴を図 8 に示す。構造生成機能を持たせたシステムにおいて、ユーザ評価を行った。結果を図 9 から図 11 に示す。「作曲の楽しさ」(図 9) の項目については、よい評価を付ける人数の割合が減ったものの、「歌詞の聞き取りやすさ」(図 10) と「生成楽曲のインパクト」(図 11) については、よい評価を付ける割合が増大した。また、ユーザからのコメントには、改行で区切るの是不自然であるなどといった意見があった。

4. 考 察

ウェブ上に公開した Orpheus Ver2.0 において実際に構成を持った楽曲が一般ユーザによって生成されたことが分かった。作曲支援の観点からは、これによって 8 小節の旋律が生成するだけでは不可能であった、作品を部分から構成する支援が可能となったといえる。楽曲のインパクトの評価が良くなったことは、このように構成をもった長い曲が出来るようになったこと、ドラムスなど、より充実した伴奏を付けられるようになったことが影響している可能性もある。歌詞の聞き取りやすさについての評価が高くなったことは、歌声生成エンジンをより最新のもの¹¹⁾ に差し替えたことが影響していると考えられる。一方、歌詞入力欄の改行によって部分を区切ったが、これについては不自然だという意見もあった。ま

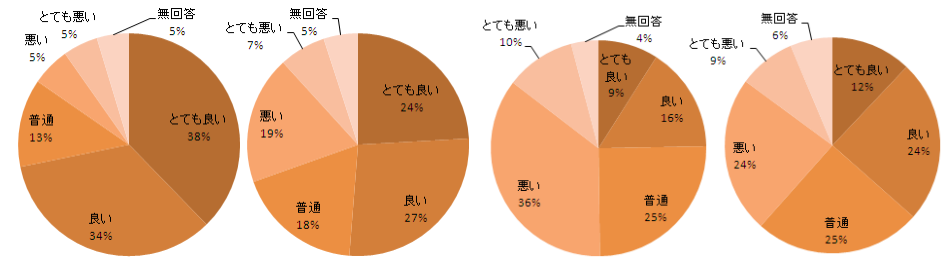


図 9 システムによる作曲の楽しさについての 5 段階評価の比率 [%]: 伴奏の増強と構成を持つ楽曲生成の機能を付け加える前(左)と後(右)を比べると、機能付加後のシステムでは楽しいとする評価が減ったことがわかる。

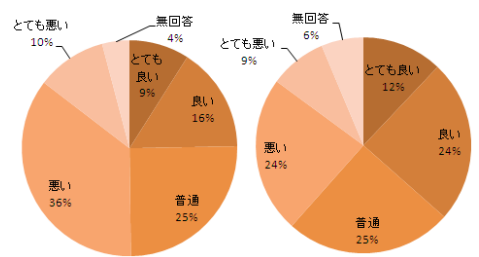


図 10 システムで生成される楽曲の歌詞の聞き取りやすさについての 5 段階評価の比率 [%]: 伴奏の増強と構成を持つ楽曲生成の機能を付け加える前(左)と後(右)を比べると、機能付加後のシステムの方が歌詞が聞き取りやすくなったことがわかった。

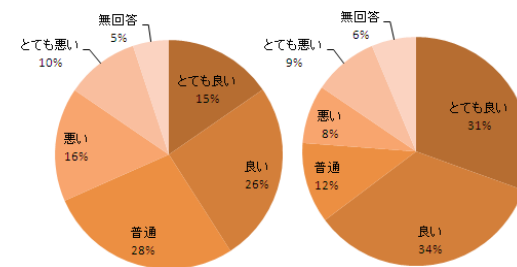


図 11 システムで生成される楽曲にインパクトがあったかについての 5 段階評価の比率 [%]: 伴奏の増強と構成を持つ楽曲生成の機能を付け加える前(左)と後(右)を比べると、機能付加後のシステムの方がインパクトを持った楽曲が生成されるとユーザが感じていることがわかった。

た入力やオプションの設定が煩雑になったために、作曲の楽しさを感じるかについては評価が下がった可能性もある。他に構造をはっきりさせるために、部分の作曲のオプションにスケールの変更や拍子の変更、遠隔調への転調ができるようにする工夫が考えられる。

5. 結 論

本研究では、自動作曲システム Orpheus の初期バージョンにおけるユーザ評価に基づいて、ドラムスなど伴奏音型の充実や長い楽曲を生成することを目的とした。長い曲を生成するために必要な構造生成の作曲支援のためにマトリクス状のインターフェイスを提案した。またそれを実装したシステムを OrpheusVer2.0 として公開し、ユーザの利用履歴を元に構成を持たせた楽曲生成が実際に使われることを確認した。今後ポップスのような楽曲が生成されるようにするには、現代のポップスの旋律や和声進行の分析に基づいて、モーラの詰め込みなど、リズムと歌詞の韻律の関係を導入することが考えられる。

謝 辞

本手法に対しての評価と有益なアドバイスを桐朋学園大学の金子仁美講師から頂いた。本研究の一部は、科学技術振興機構 CREST の援助を受けて行われた。

参 考 文 献

- 1) L. Hiller, 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.
- 2) Band in a box: <http://music.e-frontier.co.jp/product/biabwin/>
- 3) 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研究会第3回資料, 1999.
- 4) iida calling: <http://iida.jp/calling/>
- 5) M. E. Beckman and J. B. Pierrehumbert: Intonational structure in Japanese and English, *Phonology Yearbook 3*, pp.255–309, 1986.
- 6) 長谷川良夫: 作曲法教程 上巻, pp.68–89, 音楽之友社, 1950.
- 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) 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.
- 9) 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.

- 10) R. E. Bellman: *Dynamic Programming*, Princeton University Press, Princeton, New Jersey, 1957.
- 11) 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.