

pâtissier

-アマチュア作詞家のための作詞補助システム-

阿部 ちひろ^{†1} 伊藤 彰 則^{†1}

本稿では、音節数や韻、単語のアクセントに着目した作詞支援システム「pâtissier」を提案する。システムは Ngram 言語モデルをもとに、ユーザの指定した条件を満たす歌詞候補文を生成し、提示する。ユーザはシステムを辞書のように使い、提示文から主体的に言葉を選び作詞を進めることができる。主観評価実験では、被験者がシステムを用いて作詞を行い、システムが実際の作詞においてどのように使われるのかを調査した。システムの使用ログと被験者アンケートから、ユーザはシステムに自身を持つ歌詞のイメージに合うような候補の提示を求めていることがわかった。また、提示された単語は直接用いられるよりも、歌詞のストーリーを組み立てる手がかりとして用いられることが多いという傾向が明らかになった。

pâtissier

-A Lyrics Writing Support System for Amateur Lyricists-

CHIHIRO ABE^{†1} and AKINORI ITO^{†1}

In this paper, we propose a lyrics writing support system focused on the number of syllables, rhyme and word accent. The system generates candidate sentences that satisfy user-specified conditions based on Ngram, and presents them. Users can use the system like a dictionary, and write lyrics by choosing presented sentences. In our subjective evaluations, we have investigated how the system is utilized for writing lyrics actually. The log of using the system and the questionnaires showed that users want the system to present words suitable for their images, and they used the presented words as keywords of a lyrics rather than as they are.

1. はじめに

近年、DTM(Desktop Music)の普及や、動画投稿サイトなどでの楽曲発表の広がりを背景に、個人による楽曲制作が増加している。デジタル技術の発展は、特別な知識を持たない人であっても趣味として楽曲を制作し、自ら発信することを可能にした。これにより、誰もが楽曲を聴く喜びだけではなく、アーティストとして楽曲を制作、発表する喜びを享受できるようになった。このような傾向は今後さらに顕著になっていくものと考えられる。楽曲制作には作詞・作曲・演奏などの過程があるが、作曲や演奏に比べて作詞について学んできた人は少なく、楽曲制作を行う中で苦勞している人が多い部分である。しかしながら、既存の自動作詞システム¹⁾²⁾では歌詞全体を自動的に作成するため、ユーザが自分の好みの言葉を主体的に選ぶことはできない。そこで本稿では、個人による作詞を補助することを目的として、単なる自動作詞ではなくユーザが主体的に歌詞を生成・選択できる作詞補助システムを提案し、ユーザの求める作詞補助についての検討を行う。

2. 作詞補助システム

作詞の手法は、先に書いた詞に曲を付ける「詞先」と呼ばれる方法と、出来上がった曲に歌詞を付ける「曲先」と呼ばれる方法に大別される。J-POPに代表されるポピュラーソングでは後者の手法が多く採用されており、本稿でも曲先での作詞を想定する。曲先では曲のメロディに歌詞を上手く乗せられるよう単語のモーラ数と韻、さらに音程の上下に合わせたアクセントに気を配る必要がある。そこで、特別な知識を持たない人が作詞をする際に必要な補助として、

- 曲に合わせられるような音節数の歌詞を提示する
- ユーザーが意図する韻を持つ歌詞を提示する
- より自然にメロディに乗せられるアクセントを持つ歌詞を提示する

という機能が考えられる。このような機能を備えた作詞補助システムの実現のために、本稿では大量の既存文章から統計的に言葉のつながりを学習した Ngram 言語モデルを用いる。Ngram 言語モデルで基本的な単語連鎖を予測し、その上で音節数や韻、アクセントがユーザの指定条件にある単語列を歌詞候補文として提示する。³⁾

^{†1} 東北大学大学院工学研究科
Graduate School of Engineering, Tohoku University

モーラ数と母音

音節数と韻の条件を設定するためにモーラ数と母音という要素を導入する。モーラとは一般的に拍と呼ばれるもので、日本語の歌では1つの音符に1モーラを当てはめる。ユーザは生成される歌詞候補文のモーラ数を指定することでメロディに合う文を探すことができる。また、モーラがそれぞれの母音を持つかを指定することで生成される文の韻を決めることができる。例えば、3モーラ末尾「あ」と条件指定した場合は「みんな」、「かぼちゃ」といった単語が候補として提示される。

単語アクセント

さらに、単語の音の上下を指定するために単語アクセントという要素を導入する。日本語のアクセントは音の高低で表されるため、作詞においては曲のメロディの上がり下がりと言語のアクセントが矛盾しないことが望ましい。ユーザは何番目のモーラに高い音がかかるかを指定することで、単語のアクセントを考慮することが可能である。例えば、1モーラ目に指定すると「雨」、「箸」などが候補となり得る。

3. システム概要

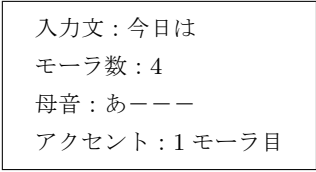
ユーザはまず、歌詞の先頭にくる単語（入力文と呼ぶ）を入力し、入力文に続く歌詞のモーラ数とそのモーラを持つ母音、アクセントの位置を指定する。システムは入力文に続く単語列を生成し、Ngram 確率と各種条件の充足を考慮してスコアを付け、上位のスコアを持つ単語列を歌詞の候補として提示する。システム入力例と提示される候補の例をそれぞれ図1、図2に示す。

3.1 Ngram 言語モデル

歌詞候補文の生成には、Ngram 言語モデルを用いる。Ngram 言語モデルとは、ある単語の生成確率が直前の(N-1)単語にのみ依存すると近似した言語モデルである。単語の出現確率は学習テキストから統計的手法を用いて推定しているため、各単語の生成確率は学習テキストに依存する。

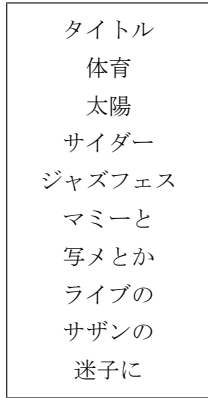
3.1.1 単語 Ngram と品詞 Ngram

単語自体のつながりをモデル化したものを単語 Ngram と呼び、単語の品詞のみを考慮して品詞同士のつながりをモデル化したものを品詞 Ngram と呼ぶ。単語 Ngram では、学習テキスト中に実際に出現した単語の連鎖に高い Ngram 確率が付くため、学習に用いた文に大きく依存する。つまり、歌詞候補文生成の際に学習テキスト中の文と似通った文が生成されやすくなる。品詞 Ngram では学習テキスト中に出現した単語は品詞が同じならば全て同一



入力文：今日は
モーラ数：4
母音：あー
アクセント：1モーラ目

図1 システム入力例
Fig.1 Example of inputs



タイトル
体育
太陽
サイダー
ジャズフェス
マミーと
写メとか
ライブの
サザンの
迷子に

図2 候補提示例
Fig.2 Example of presented words

のものを見なすため、実際には出現していない単語連鎖にも高い確率を与えることができる。候補文の種類を増やせる一方、学習テキストに実在しない単語連鎖が生まれるため日本語として不自然な候補文が生成される可能性もある。

本稿では、式(1)に示す線形結合により2つの言語モデルを併用する。

$$P'(w_i|w_{i-2}w_{i-1}) = \lambda P(w_i|w_{i-2}w_{i-1}) + (1 - \lambda)P(w_i|c_i)P(c_i|c_{i-2}c_{i-1}) \quad (1)$$

ただし、 c_k は単語 w_k の品詞を表すクラス、 $0 \leq \lambda \leq 1$ である。事前実験の結果より、 $\lambda = 0.8$ 、 $P(w_i|c_i) = 1$ とした。

3.1.2 言語モデルと単語リストの作成

Palmkit⁴⁾を用いて単語3gramと品詞3gramを作成した。学習テキストには、Sen⁶⁾で形態素解析を行った砕けた書き言葉のブログ記事⁵⁾を用いた。Senの辞書にはUniDic1.3.12⁷⁾を用い、テキスト中の各単語に読みと品詞、アクセント情報を付加した。辞書にない単語は全て未知語(UNK)1単語に置換した。アクセント情報はUniDicのアクセント型⁸⁾を用いてアクセントの位置を先頭からのモーラ数で表し、1つの単語に複数のアクセントが存在する場合は最も優先順位の高いもののみを採用した。ただし、音の上下がない単語は平板型アクセントに、アクセント型が定義されていない単語(助詞等)はアクセントなしにそれぞれ分類される。さらに、形態素解析により得られた情報を元に、学習テキスト中に出現した全単語をリスト化した。歌詞候補文はこの単語リスト内の単語を組み合わせるにより生成される。学習テキストの詳細とアクセント型の分布をそれぞれ表1、表2に示す。

表 1 学習テキストの詳細
Table 1 detail of corpora

総形態素数	異なり単語数
103,615	10,439

表 2 単語リスト中のアクセント型の分布
Table 2 distribution of accent type

アクセント型	1	2	3	4	5	6	平板型	なし
単語数	2762	1410	932	355	84	8	4252	636

3.2 文の生成

一般に、単語列 $w_1w_2\dots w_n$ の出現確率は 3gram を用いて

$$\log P(w_1w_2\dots w_n) = \log P(w_1) + \log P(w_2|w_1) + \sum_{k=3}^n \log P(w_k|w_{k-2}\dots w_{k-1}) \quad (2)$$

と表される。 w_1w_2 を入力文として、単語列の合計モーラ数がユーザが指定したモーラ数と等しくなるように単語を選び、文を生成する。

文生成アルゴリズム

歌詞候補文の生成は次のような手順で行う。

step1 リスト内の単語 w が入力文 h に続く 3gram 確率を求める

step2 単語 w を母音条件、アクセント条件とマッチングし、条件不適の場合はペナルティを与える

step3 3gram 確率とペナルティから単語列のスコアを決める

step4 単語列の末尾 2 単語を新たに h に定め、合計モーラ数が指定モーラ数と等しくなるまで step1~4 を繰り返す

まず 3gram 確率に基づき単語列を作り、条件に適合しない場合に減点ペナルティを与えることで単語を 1 つ選ぶ度に単語列にスコア付けを行う。減点を大きくすることで各条件の制約を強めることができる。母音条件に対するペナルティは可変、アクセント条件に対するペナルティは固定で 1 とした。

提示文の探索には、ビームサーチを用いる。ビームサーチでは最適解が得られる保証はないが、候補文は必ずしも最適解である必要はない。そのため、計算量と候補文の出来との兼ね合いが取れるビーム幅を用いればよい。リストから単語を 1 つ選び単語列を展開する度にスコア順に単語列をソートし、上位ビーム幅個の単語列のみをさらに探索していく。展開

中の単語列の合計モーラ数が指定モーラ数に等しくなった場合はその時点で探索は打ち切る。最終的に残ったスコア上位の単語列を提示文として採用する。

3.3 ユーザインタフェース

システムのユーザインタフェースを図 3 に示す。システムの使い方は以下の通りである。

入力文の指定

図 3①の入力窓に文頭の 2 単語を入力する

モーラ数の指定

②のボタンでモーラ数を指定する

母音の指定

③のボタンで各モーラがどの母音を持つか指定する (任意)

アクセントの指定

④のボタンでアクセントが先頭から何モーラ目かを指定する (任意)

各種パラメータの調整

⑤のつまみにより各種パラメータの値を変えることができる

音韻つまみ：母音制約のペナルティの値 (初期値 1)

個性つまみ：言語モデルの線形結合比率 (初期値 $\lambda = 0.8$)

精度つまみ：ビーム幅の大きさ (初期値 55)

候補文の提示

検索ボタンをクリックすることで候補文の生成が行われ、右側⑥に候補 10 文がリスト表示される

4. 評価実験

評価実験では被験者が実際にシステムを使って作詞をする 2 種類の課題に取り組み、システムの使われ方とユーザがシステムに望む機能を検討した。いずれの課題においても、システムで検索した条件と提示された候補のログを取り、被験者は各課題終了後に作成した歌詞についての記述式アンケートに回答した。また、実験の最後にシステムに対する感想を自由記述式のアンケートに記入した。システムは任意のタイミングで使用することとし、システムが提示した候補を必ずしも使わなくてもよいこととした。被験者はいずれも作詞経験のない 20 代の学生 4 名である。被験者の音楽経験を表 3 に示す。

4.1 課題 1

課題 1 は、日本古謡『さくらさくら』の 2 番の歌詞を作成する課題である。既にできあ



図 3 ユーザインタフェース
Fig.3 User interface

表 3 被験者の音楽経験
Table 3 Details of the subjects

	被験者 1	被験者 2	被験者 3	被験者 4
作曲経験	あり	なし	なし	あり
楽器・歌唱経験	ピアノ	ピアノ, ギター	ゴスペル	ピアノ

がった 1 番の歌詞と対となるよう、韻の踏み方や曲のイメージなどを意識して言葉を選ぶことを目標としたときにシステムがどのような使われ方をするのかを調べた。

4.1.1 実験手順

- (1) 1 番の歌詞カード (図 4) と楽譜を見ながら MIDI 音源を聴き、曲の雰囲気をつかむ。
- (2) 適宜システムを使い歌詞を考える。作詞中も自由に音源を聴くことが可能である。
- (3) 気に入る歌詞が完成したところで、楽譜中の記入欄に音符との対応が分かるよう、ひらがなで歌詞を記入する。
- (4) 作成した歌詞についての記述式アンケートに回答する。

4.1.2 実験結果と考察

まず、検索ログからシステム使用回数と検索に用いた入力文を調べた。表 4 に示すように、被験者によりシステムの使用回数にはばらつきが見られた。入力文には 2 単語以上指定する必要があるが、入力が 1 単語のみしか行われず候補が提示されないエラーがしばしば見受けられた。また、被験者全員に共通している特徴として、イメージに合うような単語を

表 4 課題 1 システム使用状況
Table 4 Results of problem1

	被験者 1	被験者 2	被験者 3	被験者 4
システム使用回数 (うち検索エラー)	48 (9)	13 (4)	15 (2)	47 (39)
同一入力文と検索回数 (括弧内はエラー)	小路・小道 6 華やぐ (5) 紫の 2 さえずり・さえずりが 4 鳴き声が 2 調べや・調べが 3 かすみ が 2 いざやいざや 2 花に・花が	きれいに 綺麗に 綺麗さは	卯月の 3 新たな 2 期待と不安 3 出会いと別れ 4	焦がれる (5) 表情 (4) 移りゆく (8) 世界 (3) はかなく (4) ゆめ (7)

さくら さくら
やよいの空は
見わたす限り
かすみか雲か
匂いぞ出ずる
いざや いざや
見にゆかん

図 4 『さくらさくら』歌詞
Fig.4 “Sakura Sakura”

さくら さくら
華やぐ小路は
すみれにいろどり
めじろか ひばりか
調べぞ近づく
いざや いざや
触れにゆかん

図 5 課題 1 回答例 (被験者 1)
Fig.5 Answer (subject1)

見つけるために「入力文はそのまま、各種条件を変えて何度も検索する」という状況が多く見られた。母音やアクセント条件は、技巧面の考慮のためというよりもむしろ、提示される候補を変化させるために指定している場合もあると考えられる。さらに、アンケートの結果から作成した歌詞のどの部分でシステムを活用したかを調べた。例として被験者 1 の回答を図 5 に示す。

被験者 1 作詞の流れ (アンケートより抜粋)

1 番の歌詞「雲」から春の鳥である「雲雀」を連想したので、入力文【ひばりか】で候補【音楽】.「音楽」から 2 番は鳥の鳴き声をイメージにすることに。
さらに、1 番の歌詞「かすみ」から白い春の鳥「めじろ」を連想したので、鳴き声のイメージから、入力文【調べが】で候補【近づく】.「聞こえる」だと直球すぎる気がしたので「近づく」を採用した。
なるべく語尾の音韻は同じになるようにした。

被験者 1 のアンケート結果から、1 番の歌詞から連想される単語を入力としてシステムを用い、さらに提示された候補から 2 番の歌詞のイメージを広げる、といった使われ方がされている。提示された候補はそのまま歌詞として使われるだけではなく、歌詞のヒントともなり得ることがわかった。

4.2 課題 2

課題 2 は、もともと歌詞のない曲である、シューベルト作曲『ます』に歌詞を付ける課題である。ある程度自由な条件下でシステムがどのような使われ方をするのか、またその際にどのような補助が望まれるのかを調べた。なお、被験者には曲のタイトルは伏せ「課題曲」というタイトルの楽譜を渡した。

4.2.1 実験手順

- (1) 楽譜 (図 6) を見ながら MIDI 音源を聴き、曲の雰囲気をつかむ
- (2) 歌い出しリスト (図 7) 上の単語から歌詞の最初の単語を選ぶ
- (3) 以降、課題 1 と同様の手順で作詞を行いアンケートに回答する

4.2.2 実験結果と考察

課題 1 と同様に、検索ログからシステム使用回数と検索に用いた入力文を調べた。結果を表 5 に示す。課題 1 の結果 (表 4) と比較して、全体的に検索回数が減少した。自由に作詞する場合には、1 番の歌詞のような入力の手がかりにするものがないため、どこでどのようにシステムを使えばよいか難しく、結果としてできる限り自力で考えようとする傾向が現れたと考えられる。またアンケートでは、課題 1 に比べて課題 2 では出力結果が期待にそぐわないことが多かった、との回答も見られた。加えて、2 つ目の課題ということで被験者がシステムに慣れてきたことも要因として考えられる。課題 2 の回答例として、被験者 2 の回答を図 8 に示す。



図 6 『ます』の楽譜
Fig. 6 Score of "masu"

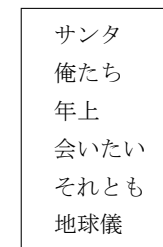


図 7 歌い出しリスト
Fig. 7 Inputs word list

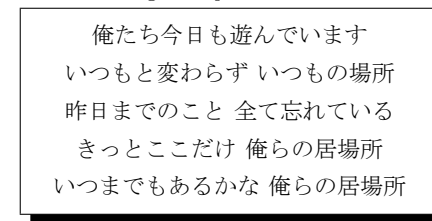


図 8 課題 2 回答例 (被験者 2)
Fig. 8 Answer (subject2)

表 5 課題 2 システム使用状況
Table 5 Results of problem2

	被験者 1	被験者 2	被験者 3	被験者 4
システム使用回数 (うち検索エラー)	51 (5)	5 (0)	7 (0)	14 (6)
同一入力文と検索回数 (括弧内はエラー)	雪空を 2 街並みに・街並みが 笑って・子供が笑い 6 今宵は・今夜は 5 騒いで 5 クリスマスに 9 今夜は騒いで 2 騒ぎ歌って 2 子供たちは 3	俺たち 4	クリスマスは 4 もう 2	川は 2 たどり着く・着く (7)

被験者 2 作詞の流れ (アンケートより抜粋)

曲を聴いて「ヨーロッパ」をイメージ。歌い出しは「俺たち」に決定。
入力文【俺たち】で候補【貫禄】。
これらを総合して、ヨーロッパの貫禄ある農夫の歌にすることに。

被験者 2 の回答は歌詞自体にはシステムが提示した単語は表れていないが、曲のテーマ決定の際にシステムが使われたことがわかった。

4.3 システムについての感想

システムを使ってみて次のような感想が挙げられた。

- 候補が名詞で出てくるとインスピレーションが湧きやすかった
- 想像力を促すために使った
- 課題 1 のようにテーマが決まっていた方が役立った
- 似た単語が出過ぎると感じた
- 動詞の後に「ました」や「でした」が出て困ってしまった

また今後システムに望むこととしては、次のような意見が挙げられた。

- 明るい・暗いと言うようなイメージや感情入力ができたらよい
- シソーラスのように単語の言い換え検索ができたらよい
- 入力文に続く単語だけではなく、入力文を修飾する単語の検索ができたらよい

5. おわりに

本稿では、個人による作詞を補助することを目的とした作詞補助システムを構築し、実際にシステムがどのように使われるのか、またどのような補助が必要とされるのかを主観評価実験を通して検討した。韻やアクセントの考慮といった技巧的な側面をどの程度重視するかは被験者によってばらつきがあったが、いずれの被験者においても自身の思い描く楽曲のイメージに合う歌詞の候補が提示されることを望んでいることが明らかになった。また、提示された候補を単に歌詞として使うのではなく、それを手がかりとして連想される単語から歌詞のストーリーを広げるといったようなシステムの使われ方も見られた。今後はユーザの楽曲に対するイメージを候補に反映させることが課題である。また、ユーザの想像力を広げやすい候補の提示方法についても検討したい。

参考文献

- 1) 伊藤雅光：ユーミンの言語学 (1)-(46), 日本語学 16(4)-20(8) 連載, 明治書院, (1997-2000).
- 2) HRG Oliveira 他：Tra-la-Lyrics: An approach to generate text based on rhythm, Proc. Int. Joint Workshop on Computational Creativity, (2007).
- 3) 阿部ちひろ, 伊藤彰則：統計的言語モデルを用いた作詞補助システム, 情報処理学会研究報告, Vol.2011-MUS-91 No.9, 2011.
- 4) 伊藤彰則, 好田正紀：単語およびクラス n-gram 作成のためのツールキット, 信学技報, NLC 100(521), pp.67-72, (2000).
- 5) にしのそらにみかづき, 入手先(<http://crescent213.blog12.fc2.com/>).
- 6) sen project, 入手先(<http://sen.dev.java.net/>).
- 7) 伝康晴 他：コーパス日本語学のための言語資源：形態素解析用電子化辞書の開発とその応用, 日本語科学, 22 号 pp.101-122, (2007).
- 8) 伝康晴 他：話し言葉研究に適した電子化辞書の設計, 第 2 回話し言葉の科学と工学ワークショップ講演予稿集, pp. 39-46, (2002).