

外国語の歌曲の自動訳詞システムに関する初期実装

西村 綾乃^{1,a)} 伊藤 貴之^{1,b)}

概要: 音楽は世界中で作曲されており、インターネットが普及した現代では外国語の歌曲を耳にする場面も多くなった。しかしながら、母国語ではない言語の歌詞を完全に理解することは難しい。歌詞の意識や直訳を目にすることもあるが、音楽に相応しい詩の形ではなく、そのままメロディに乗せて歌うことはできない。そこで、本稿では外国語歌詞の日本語訳を出力するインタラクティブな訳詞システムを提案し、その初期実装を行った。さらに、3つの童謡を用いて実験を行ったため、その結果を報告する。

An Initial Implementation of Interactive System for Automatic Translation of Foreign Lyrics

AYANO NISHIMURA^{1,a)} TAKAYUKI ITOH^{1,b)}

Abstract: Music has been composed all over the world. Opportunities to listen to songs in foreign language are increasing thanks to the evolution of the Internet. However, it is difficult for us to completely understand the lyrics in foreign languages other than our mother tongue. We sometimes see freely or literally translated lyrics. However, it is often difficult to sing the translated lyrics with given melodies if they do not form styles of lyrics. Therefore, in this paper, we propose an interactive system for translation of foreign lyrics and present its initial implementation. In addition, we introduce and discuss result of our experiments by using three nursery songs.

1. はじめに

古くから音楽と言葉には密接な関係がある。グレゴリオ聖歌の時代においては、音楽とは言葉に従属するものであり、祈りの言葉を強調するための道具であった。そして、長い歴史を経る中で、音楽は地位を確立し、メロディそのものを楽しんだり、歌唱曲として歌詞と一緒に鑑賞するべきものへと変化した。

現在ではインターネットが普及したことにより、世界中の音楽をいつでも気軽に聞くことが可能になった。しかしながら、外国語の歌曲は音楽そのものを楽しむことは出来ても、馴染みのない言語による歌詞の内容を完全に理解することは難しい。中には、歌詞の意識や直訳が掲載されていることもあるが、単なる直訳によって生成された文章はメロディの音数や流れと合致しない。そのため、翻訳した文章をそのままメロディに乗せて歌うことは、音楽的な要

素を無視することになる。

一方で、海外の歌曲の多くが日本語に訳詞され歌われている。しかし、訳詞には語学力や語彙力に加えて音楽の知識も必要であり、非熟達者には難しい作業である。具体的には、訳詞を行う歌曲に対して、原語での意味理解、音楽的なフレーズやリズム構成の理解、歌唱における日本語の発音への理解などが必要になってくる。多数の歌曲で1つの作品となるミュージカルやオペラなどでは、1曲ずつの理解だけではなく、全編を通した話の流れや、メロディフレーズの使い方などの理解も必要となり、その労力は多大なものとなる。

そこで著者らは、外国語歌詞の日本語訳を出力するインタラクティブな自動訳詞システムの開発を目指している。本システムは、ユーザが入力した楽譜情報から自動的に訳詞された譜面を生成し、好みに合わせて編集が可能なアプリケーションを想定している。自動訳詞の手法が確立することによって、オペラやオペレッタ、ミュージカルなどの大量の歌詞を含む作品の訳詞も容易になり、システムを通じて日本語での上演のための補助を行うことで、日本国内での音楽文化の発展に貢献できるだろう。本稿では、英

¹ お茶の水女子大学
Ochanomizu University

a) nishimura.ayano@is.ocha.ac.jp

b) itot@is.ocha.ac.jp

語歌詞を付随した単旋律の MusicXML を入力とすることで、英語歌詞の翻訳を行い、モーラ数を減らし、メロディに当てはめることで日本語歌詞の譜面を作成する初期実装について、手法・実験結果の報告を行う。

2. 訳詞の課題

日本語への訳詞の歴史は近藤朔風らが切り拓いた [1]。代表される訳詞の作品には「野ばら」や「ローレライ」などがあり、これらは現代でも訳詞時と変わらない日本語で歌い継がれている。近藤は原言語の意味と韻律を可能な限り日本語に再現しようとする訳詞を重視した。本提案においても、原言語の意味をなるべく日本語に再現するための自動訳詞を目指す。

日本語への自動訳詞システムの開発に着手するにあたり、まず訳詞特有の課題を列挙する。

原言語の歌詞を日本語に訳すとほとんどの場合、音数（モーラ数）が増えるという問題点にぶつかる。しかしながら、メロディに使われている音符の数は決まっているため、訳詞においてはモーラ数を極端に増やすことはできるだけ避けなければならない。これについて、“訳詞においては、原文の中のどの部分は訳出し、どの部分は省略するかという取舍選択が行われることになる”と、松田は述べている [2]。結果として日本語への訳詞において、原言語の意味を簡略化する必要が生じることが多い。また、モーラ数に関わる課題として、音節の観点から「撥音」すなわち「ん」と発音される音について、モーラ数を 1 とする場合と、直前の発音と合わせて歌うことで、モーラ数に含めないような課題がある。撥音だけでなく、促音、長音、二重母音と呼ばれる発音をモーラとして独立させるかの判断は、その歌唱が与えるリズム感にも影響を与える。

さらに、松田は本文中で、こうも述べている。“曲の終わりに歌いあげる音では、意味のレベルを犠牲にしても「ア」か「オ」の母音にすべき場合がある。”これは、歌唱に適した日本語の発音の課題とも言える。山本は日本語の母音について以下のように述べている [3]。“「あ・お」は会話にも歌唱においてもひびきやすい母音であるが、「え」は歌唱においてとくにひびかせにくい母音である。そこでたえず発声訓練において美しい共鳴をとるよう配慮されている。「い」「う」は歯をくいしばったように口腔中が狭くなり、共鳴のない固い平たい声になりやすいのである。”以上のことから、訳詞を行う際に、選択した単語の母音に対して、言葉の意味を重視するか、歌唱的な面を重視するかといった訳詞の問題が生じることが分かる。

ここで、訳詞をする際に課題となる点をまとめる。

- 課題 1 日本語への翻訳によってモーラ数が増える
- 課題 2 原言語の意味を取捨選択し、簡略化する必要がある
- 課題 3 撥音などの発音を 1 音として独立させるかどうか、リズムと音節を考慮して判断する必要がある
- 課題 4 音楽的な理由から、特定の音に対する母音を限定する必要がある

課題 2 については、モーラ数が増えるのを防ぐために行われる作業であり、課題 1 に従属する課題である。また、これらの課題について、人手で訳詞を行った場合でも、結果にばらつきが出ることは自明であり、システムにおいても訳詞の解を 1 つに絞るのは難しいことが予想される。そこで、ユーザが自分の好みに合わせて自由に調節できるパラメータや、譜面の編集が可能なシステムであるべきだと考えた。

3. 関連研究

歌詞に関連する研究は数多く存在する。阿部らはモーラ数と母音を入力することで、歌詞候補文を生成し提示するシステム「pâtissier」を提案している [4]。伊藤らは入力文のモーラ数を算出し、その結果から入力文を歌詞として割当可能な楽曲リストを作成し、ユーザに提示する、暗記のための替え歌を自動生成するシステムを提案している [5]。

音楽と翻訳に関連する研究では、日本語を中国語に翻訳した歌詞を計算機に歌わせる手法が提案されている [6]。この手法の目的は計算機による中国語の歌唱であり、自動翻訳を完成させているわけではない。歌詞は翻訳後に人手で修正したものを用いている。

一方で、翻訳した文章のモーラ数を減らす課題は、自然言語処理における文圧縮の問題と関連がある。井手上らは、部分的表現を中心として翻訳を行う部分的機械翻訳 [7] を提案している。この手法は、原言語文と原言語文の一部の表現（部分的表現）を入力とし、部分的表現を中心とした簡潔な翻訳を出力するものである。また、体言止めや助詞止めといった文末表現に着目し、新幹線車内の電光掲示板等で流れるような簡潔な表現に文章を加工する手法もある [8]。

これ以外の関連研究もその大半は「歌詞」か「メロディ」のどちらか一方を新しく作り出すものである。具体的には、歌詞の自動生成、あるいは入力歌詞に対応するメロディの自動生成のいずれかの手法が多数提案されている。それに対して本研究では、「歌詞」と「メロディ」がすでに両方存在することを前提として、形式や内容の点で整合性を保ったまま歌詞を翻訳することを目的としている。

4. 自動訳詞システムの構想

自動訳詞システムの構想のシーケンス図を図 1 に示す。现阶段の本研究では、対象となる原語は英語に限定している。ここで、図中の FE はフロントエンド、BE はバックエンドを指す。以下にシステム利用時の手順を示す。

- (1) 訳詞処理と譜面の生成
- (2) 生成された譜面の編集
- (3) 編集した譜面の保存

(1) では、ユーザが訳詞に関するパラメータ（モーラ数の制限、母音の制限、生成された文章の歌詞らしさなど）を設定し、英語歌詞の MusicXML を入力とすることで、訳詞の付随した譜面を生成する。(2) については、生成され

た譜面中のフレーズに対して修正したい箇所があれば、歌詞候補等を提示することで編集が容易になる、インタラクティブな機能を想定している。また、(3)では、ユーザが何度も繰り返しシステムを利用し、多数の楽曲の編集を繰り返すことで、パラメータの調整や単語の選び方の好み、メロディの修正方法を学習できると考えている。

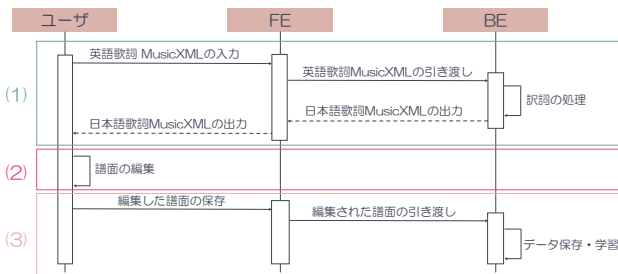


図 1 シーケンス図

5. 実装

初期実装として、バックエンドの処理として行うことを想定している、英語歌詞の MusicXML から訳詞の処理を行い、日本語歌詞の MusicXML を生成する部分の実装を行った。その処理手順を図 2 に示す。処理手順について、以下の 3 つに大別することができる。

- (1) 原言語の翻訳
 - (2) 翻訳文の歌詞化
 - (3) 生成した歌詞のメロディへの当てはめ
- 続いて、各処理手順の詳細について述べる。

5.1 原言語の翻訳

入力された MusicXML の英語歌詞から、Microsoft Translator テキスト API[9] を利用して機械翻訳文を取得する。この時、ピリオド、カンマ、セミコロン、区切りを 1 フレーズとみなし、歌詞に付随するメロディの音符の数を英語歌詞の音数とする。以降、1 フレーズに対する英語歌詞の音数をメロディ音数と呼ぶ。

翻訳は API 依存であり、取得した機械翻訳文に、訳しきれなかった英語が残る場合がある。その場合は、アルファベットを削除する。

この処理では、1 フレーズの機械翻訳文および、メロディ音数を取得する。

5.2 翻訳文の歌詞化

取得した機械翻訳文および、メロディ音数を利用して、歌詞の生成を行う。機械翻訳文を歌詞に近い文章へと加工する処理をここでは歌詞化と呼ぶ。1 フレーズの機械翻訳文のモーラ数がメロディ音数を超える場合、次の 4 つの操作を逐次実行する。この処理を、全てのフレーズに対して繰り返す。

- (1) 敬語表現の削除

- (2) 断定表現の削除
- (3) 助詞の省略
- (4) 文章生成

この時、撥音および促音は 0 モーラとして数える。また、形態素解析には MeCab を用いた。

5.2.1 敬語の削除

機械翻訳を利用した場合、多くの文章に敬語表現が使用されている。そのため、関連研究で述べた文末表現に着目した要約手法の一部を、今回の実装に組み込んでいる。ここでは、モーラを減らす手段として、取得した機械翻訳文から敬語表現である「です」「ます」を削除する。

5.2.2 断定表現の削除

文末表現が「だった」「である」「だ」というような断定の場合も、敬語表現と同様に削除する。

5.2.3 助詞の省略

歌詞のモーラ数を減らすために、以下の 4 つの操作が有効であることが現段階までの実験でわかっている。

- (1) 助詞の省略
- (2) 翻訳文をもとにパラフレーズを得る
- (3) 一人称としての「私」という主語の省略
- (4) 原文の単語から類義語を求め、翻訳を行う

今回は、この中から助詞の省略を適用し、モーラ数を減らすことを試みた。形態素解析の結果、助詞と判断された単語は翻訳文から削除する。

5.2.4 文章生成

上記の 3 つの手順を踏んだ上で、メロディ音数を超える場合は、1 フレーズ中の tf-idf 値が最も高い単語を先頭の単語として与えることで、2-gram を用いてメロディ音数と同等のモーラ数を持つ歌詞を生成する [10]。しかしながら、モーラ数と与えられた単語によっては歌詞の生成ができない場合もある。その場合は、助詞の省略まで行った文章を歌詞として、メロディ音数を超えることを許容する。

5.3 生成した歌詞のメロディへの当てはめ

生成した歌詞をメロディに当てはめる場合、以下の 3 つの場合に分かれる。この時、生成した歌詞のモーラ数を歌詞モーラ数と呼ぶ。

- (1) 歌詞モーラ数 = メロディ音数
- (2) 歌詞モーラ数 < メロディ音数
- (3) 歌詞モーラ数 > メロディ音数

以下に、それぞれの場合の詳細について述べる。

5.3.1 歌詞モーラ数 = メロディ音数

歌詞モーラ数とメロディ音数が等価の場合は、そのままメロディに当てはめることが可能である。従って、メロディの変更は必要ない。

5.3.2 歌詞モーラ数 < メロディ音数

日本語の歌唱では、1 つの文字の母音を伸ばし複数の音符を歌いあげることが許容されている。そこで、歌詞モーラ数よりもメロディ音数の方が多い場合、フレーズの先頭から n 番目の音が $n < n + 1$ となる時、 n 番目の音に「ー」を挿入することで歌詞モーラ数を増やす。

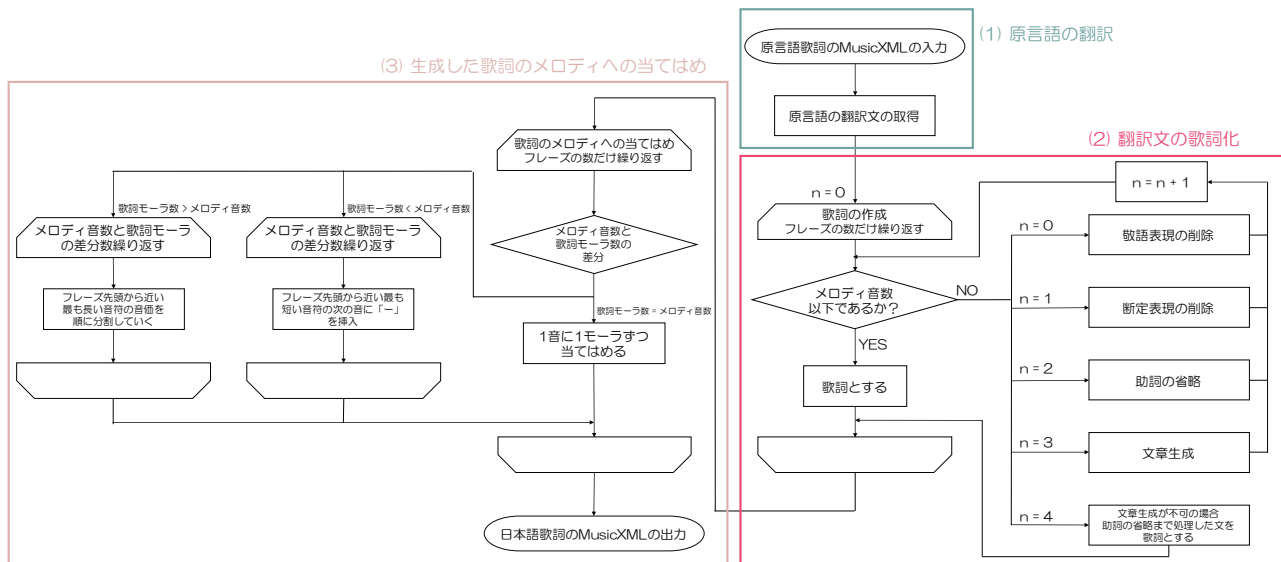


図 2 実装内容の処理手順

全ての音が等価だった場合は、歌詞の末尾に「-」を挿入する。

5.3.3 歌詞モーラ数 > メロディ音数

歌詞モーラ数がメロディ音数よりも多い場合は、メロディに使われている音を分割する必要がある。この時、フレーズ先頭から近い最も長い音符を半分の音長に分割することで、音数を増やした。符点の場合は 2:1 になるように分割する。

6. 結果

本実験では、「London Bridge」「Under the Spreading Chestnut Tree」「The Itsy Bitsy Spider」の3曲の童謡を使用した。また、tf-idf 値を算出するための文書として Project Gutenberg や青空文庫などから作成されたコーパス [11] を利用し、文章生成には日本語版 Wikipedia コーパス用いた [12]。その原曲の譜面と、プログラムを適用した結果をそれぞれ図 3, 4, 5, 6, 7, 8 に示す。

「London Bridge」「Under the Spreading Chestnut Tree」では、人手の訳詞と同じ歌詞部分が存在した。これは、簡単な英語であれば現在のアルゴリズムを適用すれば歌詞として利用できることを示唆している。

一方で、「Under the Spreading Chestnut Tree」「The Itsy Bitsy Spider」において、途中で文章生成を行っているフレーズがある。そのフレーズの歌詞は「幸せ旅なに」や「洗い流させれれせ」「出てきたきた移動」となっており、意味の通じない歌詞となってしまった。

また、「スパイダー水注ぎ口上がった」とあるが、元の翻訳文は「スパイダーは水の注ぎ口に上がった」である。助詞を全て削除したことにより、文章の単語の区切りが変わり、意味の通じづらい歌詞となってしまった。「みず」を「すい」と読む点についても、読み方の結果がツールに依存してしまう。

音楽的な部分では「London Bridge」の最後のフレーズである「マイフェアレディ」という歌詞が、非常に歌いにくい歌詞となっている。英語では「my fair」と、2音節になっている部分が日本語に訳したことで、「ま・い・ふ・え・あ」のように4音節になってしまったためである。

7. 議論

実験の結果から、以下の点が分かった。

- (1) 童謡のような、簡単な歌詞とメロディであれば訳詞の譜面が生成できる
- (2) 文章生成した歌詞の意味が通じない場合がある
- (3) 助詞を全て削除すると本来の意味を損なう
- (4) 単語の音節などを考慮して譜割りを決める必要がある
- (5) 処理の結果がツールに依存する点がある

(2) については、ユーザが修正可能であることを前提とするのであれば、自動生成した文章を提示するよりも、単語の候補のみを抽出し、ユーザに提示することで、人手による歌詞生成に切り替えることもできる。

また、訳詞の課題で提示した課題 1 について、解決には及ばずとも童謡等であればモーラ数を抑えて譜面を生成できることが分かった。

今後は、複雑な英語歌詞を用いることで、取得した機械翻訳文が不適切な翻訳になっている可能性がある。インタラクティブなシステムを作っていく上で、人手による修正が入る前提で、自動で訳詞を行う部分の精度をどこまで高めるか定義を明確にする必要がある。

8. 今後の展望

本稿では英語歌詞の付随した単旋律の MusicXML を入力すると、日本語歌詞の単旋律の MusicXML を出力する自動訳詞システムのための初期実装を行った。童謡などの簡単なメロディと歌詞であれば、人手に近い訳詞が生成で

きることが分かった。その一方で、意味の通じない文章や、音節の区切りが不適切なフレーズなどが出来てしまう場合があるため、より精度を上げる必要がある。次の段階では、訳詞の補助システムとしてどのようなインターフェースを目指すべきかを定義し、実装を進めたい。

インターフェースの定義後は、文章や音楽の解析を含めた訳詞の手法を考えていく必要がある。そこで、楽曲解析には、GTTM[13]を利用することを考えている。文章の解析においても、係り受け解析等を行っていきたい。

また、西欧語の多くは強弱アクセントを持つ発音が多いが、日本語は高低アクセントを持つ言語である。そのため、音高や拍節構造を考慮した日本語の発音に最適な訳詞を行いたい。

Lon - donbridge is fall - ing down. fall - ing down. fall - ing down.
5 Lon - donbridge is fall - ing down. my fair la - dy.

図 3 London Bridge

ろん どんぼし おちる おちる おちる
5 ろん どんぼし おちる まい ふえ あれ だい

図 4 London Bridge の訳詞

Un - der the spread - ing chest - nut tree. There w - e sit both you and me. Oh how hap - py
6 we woul - d be. Un - der the spread - ing chest - nut tree.

図 5 Under the Spreading Chestnut Tree

くり - き の し - た で あな - た と わたし しあわせ
6 たびなに くり - き の し - た で

図 6 Under the Spreading Chestnut Tree の訳詞

The it - sy bi - tsy spi - der went up the wa - ter spout. Down came the rain and
7 washed the spi - der out. Out came the sun and dried up all the rain; and the
13 it - sy bi - tsy spi - der went up the spout a - gain.

図 7 The Itsy Bitsy Spider

すばいだいすいそそぎこうじょうがったあらいなが
7 させれせだしてきたきたいどうすー
13 ばいだいふたたびくちいった

図 8 The Itsy Bitsy Spider の訳詞

参考文献

- [1] 松田直行：訳詞の誕生：近代日本音楽史における訳詞家近藤朔風の位相，駒沢短大文芸，Vol. 37, pp. 15-43 (2007).
- [2] 松田直行：「訳詞」とは何を訳すのか：近藤朔風と森外によるオペラ『オルフェウス』訳詞の比較研究のための序章，駒澤日本文化，No. 7, pp. 165-192 (2013).
- [3] 山本金雄：歌唱における日本語の発音に関する一考察，千葉大学教育学部研究紀要. 第 2 部，Vol. 22, pp. 233-249 (1973).
- [4] 阿部ちひろ，伊藤彰則ほか：patisserie-アマチュア作詞家のための作詞補助システム，研究報告音楽情報科学 (MUS)，Vol. 2012, No. 17, pp. 1-6 (2012).
- [5] 伊藤悠真，寺田努，塚本昌彦ほか：Mnemonic DJ: 暗記学習のための替え歌自動生成システム，情報処理学会論文誌，Vol. 56, No. 11, pp. 2165-2176 (2015).
- [6] 久保一志，江原暉将：日本語歌詞を中国語に翻訳しパソコンに歌わせる，言語処理学会年次大会，pp. 6-9 (2006).
- [7] 山本和英：全文を翻訳しようとしないう機械翻訳：ワードグラフによる部分的機械翻訳の試み（機械翻訳技術の向上），*Japio year book*, pp. 276-279 (2013).
- [8] 和英山本，論史池田，一輝大橋：「新幹線要約」のための文末の整形，自然言語処理 = Journal of natural language processing, Vol. 12, No. 6, pp. 85-111 (オンライン)，入手先 (<http://ci.nii.ac.jp/naid/10016863672/>) (2005).
- [9] : Microsoft Translator API, <https://www.microsoft.com/ja-jp/translator/translatorapi.aspx> (2017).
- [10] 西村綾乃，伊藤貴之：外国語の歌曲を自動訳詞するシステムのための課題と手法の検討，*JSAI 2017* (2017).
- [11] : 日英対訳文対応付けデータ，<http://www2.nict.go.jp/astrec-att/member/mutiyama/align/index.html> (2017/02/24).
- [12] : 日本語版 wikipedia コーパス，<http://dumps.wikimedia.org/jawiki/latest/jawiki-latest-pages-articles.xml.bz2> (2017).
- [13] Tojo, S., Hirata, K. and Hamanaka, M.: Computational Reconstruction of Cognitive Music Theory, *New Generation Computing*, Vol. 31, No. 2, pp. 89-113 (2013).