

# 三大歌集における歌語の共起語変化による美的観念変遷発見

池上 藍羽<sup>†</sup> 岡田 龍太郎<sup>†</sup> 中西 崇文<sup>†</sup>

<sup>†</sup> 武蔵野大学データサイエンス学部データサイエンス学科 〒135-8181 東京都江東区有明 3-3-3

E-mail: <sup>†</sup> s2022003@stu.musashino-u.ac.jp, {ryotaro.okada, takafumi.nakanishi}@ds.musashino-u.ac.jp

**あらまし** 本稿では、和歌における歌語の共起語変化による美的観念の時代変遷発見方式について示す。和歌は、時代が移り変わるとともに構造や用いられる言葉が変化する。和歌を対象として構造や言葉の時系列変化を分析することで、日本人の美的観念の時代変遷を明らかにすることができる。本稿では、三大歌集である万葉集・古今和歌集・新古今和歌集を対象とし、和歌を詠むにあたり中心的な要素となる言葉である歌語に着目する。本方式は、カテゴリに分類した歌語の出現頻度を歌集ごとに差異を導出し、その歌語における重要共起語を文脈語として抽出することを可能とする。本方式は、歌語の文脈語の差異を明らかにすることで、美的観念の時代変遷を示す。

**キーワード** 和歌, 歌語, 共起語変遷, 美的観念変遷

## 1. はじめに

2019年5月、「平成」に代わる新たな元号を「令和」と決定し、新元号の出典である万葉集が注目された。万葉集は、現存する最古の和歌集であり、数多くの和歌が収められている。しかし、古典文学作品を嗜むためには、知識や教養が必要であり、隔たりを感じる人も少なくない。より多くの人の古典文学に対する興味関心を高めることは、文化を継承するうえで非常に重要である。また、古来の文化や思想に触れることにより、過去から多くの知恵を学ぶことで、より豊かな未来を創ることにつながる。

近年、和歌や古典を含む歴史的資料のデータベース化が進んだことにより、テキストマイニングの技術が人文学研究においても用いられるようになってきた。従来、人間が行っていた膨大な労力を必要とする作業に計算機を利用することでより効率的に研究を進めることが可能となる。

また、テキストマイニングにより研究者による主観的な分析ではなく、人文学研究においては困難とされていた数多の仮説の客観的な検証、また、新たな仮説や知見を引き出すことが可能となってきた。和歌データベースを用いることにより経験と勘に頼らない、従来の人文学研究とは一線を画した研究が実現できる。

本稿では、和歌における歌語の共起語変化による美的観念の時代変遷発見方式について示す。和歌は、時代が移り変わるとともに構造や用いられる言葉が変化する。和歌を対象として構造や言葉の時系列変化を分析することで、日本人の美的観念の時代変遷を明らかにすることができる。

本稿では、三大歌集である万葉集・古今和歌集・新古今和歌集を対象とし、和歌を詠むにあたり中心的な要素となる言葉である歌語に着目する。本方式は、カテゴリに分類した歌語の出現頻度を歌集ごとに差異を導出し、その歌語における重要共起語を文脈語として

抽出することを可能とする。本方式は、歌語の文脈語の差異を明らかにすることで、美的観念の時代変遷を示す。

本稿は次の通り構成される。2節では、関連研究について紹介する。3節では、本方式である歌語の共起語変化による美的観念変遷発見について述べる。4節では、本方式を実現する実験システムを構築し、5節で本稿をまとめる。

## 2. 関連研究

本節では、本方式に関連する研究について挙げる。これまで、歌語に焦点をあてた古典和歌研究や、和歌を対象とする、計算機を用いた研究がされてきた。ここでは、特に歌語に着目した関連研究や、和歌を文字列として捉えた関連研究について挙げる。

### 2.1 歌語の枠組みに関する研究

久保田[1]は、「歌語は、非日常的な言語表現で、感動を托すにふさわしく、美しいハレの言葉と考え、非歌語に俗語・漢語・音便形が認められる」としている。また、井手[2]は、「歌ことばの本質を、句切れをもつことばであり、情報量の少なさを補うために歌語が含蓄または曖昧さをもつという考えのもとに、現実にとどのようなやり方で歌語を歌語たらしめているかという原則」を詳説している。以上より、歌語は三十一文字の文学である和歌が含む情報を拡張させる役割を持つと考えられている。我々は、和歌にとって歌語は核となる存在であり、和歌を読み解くうえで欠かせない重要な要素であると考えた。

佐藤[3]は、歌語の定義を「原則として、口頭語や文章語には用いられず、歌にのみ用いられる語・表現」と示している。しかし、佐藤は、歌語を扱った諸論者の論点は一律ではなく、扱う歌語の範囲が論者によって異なることから、「歌語をどのように捉えるか」という基本的枠組みが、論者の共通理解となっていない」と

論じている。また、佐藤は、辞書においても定義は様々であり、最も多く共通して示されているのは「特に和歌に多く用いられる特殊な言葉や表現」ということである。しかし、「一般に書き言葉や話し言葉では使われない言葉」を加えるかどうかは辞書によって異なるとしている。以上より、どの単語が歌語に含むのか定義は曖昧であり、体系的に示すことは極めて困難であると考えられる。

本研究では、三省堂が出版している『雅語・歌語 五七語辞典』[4]に記載されている語句を歌語として定義した。編者である西方は上代の万葉集をはじめ、昭和前期に至るまでの和歌、連歌、短歌などから採集した五音七音表現を意味ごとに分類した。

## 2.2 和歌表現史の研究

百留[5]は、「和歌表現の史的変遷を、語学的手法を用い、レトリックの変遷から解明」している。その際に、レトリックとして歌語や枕詞に着目して考察している。百留は、従来の研究では「勅撰集や万葉歌人における和歌表現を明らかにしたりする限定的・個別的な研究が中心になされてきた」と述べている。百留の研究は、「特定の資料や個人のパーソナリティなどに依拠しない、これまでとは違った新たな語学的手法」である。

百留は、歌語の問題点について「文学的考察には歌語を体系的に捉えるという観点」や「歌語・歌ことばの史的変遷を貫く大きな枠組み」がみられないと論じている。

百留は「和歌のレトリックを歌語のみや掛詞のみといった単体の表現として捉えるのではなく、レトリックによる複合体の表現」として捉え、相互関連的に把握することで、和歌表現の多角的追求を可能にしている。

百留は、歌語と共起している単語を時代ごとに調査し、共起語から歌語が持つ概念を明らかにし、和歌表現の史的変遷を分析した。これにより、時代による歌語がもつ概念の変化が明らかにされた。

この研究では、体言を中心に考察されており、取り上げられている歌語も少数に限定されていた。

本研究では、体言だけではなく、用言も考察対象とした。古文用 UniDicS[6]を用いて和歌を形態素解析し、見出し語化をすることで、動詞や形容詞などの用言を考察対象にすることが可能となった。

また、計算機を用いて共起関係を分析することで、人間がこれまで重要視してこなかったような単語まで網羅的に把握をすることが可能である。

## 2.3 パターン抽出による古典和歌の研究

竹田ら[7]は、和歌における言い回しに着目し、和歌の類似性を明らかにしている。古典研究に計算機を用

いることで、既存の仮説を検証するのではなく、研究者が新たな仮説を得るプロセスの実現を目指した。この研究は、計算機によって作業の効率化を図り、人手では困難とされていた仮説の検証を可能とし、さらに新たな仮説の契機となるデータを提示する。

和歌表現における類似性の研究はこれまで名詞や動詞を中心とする自立語に着目されてきたことを示し、竹田らは、自立語のみならず、助詞や助動詞の付属語に着目した。パターンによる符号化と最小記述長原理を用いて、特徴パターンや類似歌、差異表現の抽出を実現した。ワイルドカード付きパターンを用いて、定数部分を付属語や用言の活用語尾に限定することによって回しのパターンを抽出している。竹田らは、パターンの類似性指標として SRS(String Resemblance System)と呼ばれる新たな体系を定義し、和歌の類似性を計量することを可能とした。

## 2.4 文字 N-gram による古典和歌の研究

近藤ら[8]は、N-gram により、平安和歌資料における、歌語などを含む言い回しの表現の網羅的な研究を可能とした。近藤らは和歌を文字列として捉え、N-gram により、語彙や言い回しを含む表現などを抽出できるとした。N-gram で文字列を抽出する方法はひらがな且つ清音に表記体系が統一された和歌データにおいて精度は高くなり非常に有効であると示している。N-gram を用いることの利点として、形態素タグを付ける作業を必要としないことや、単語単位ではなく文字列の長さに着目しているため掛詞に関する要素や類句の抽出が可能であることが挙げられる。また、N-gram において、長い N を設定することで、語句のみならず、掛詞に関する要素や慣用的な表現、類句の網羅的な抽出が可能である。N-gram を利用することで、あらかじめ単語を指定することなく、文字列を抽出できるメリットがある。

しかしながら、近藤らは、この手法では、単語として成り立たないような偶然発生した文字列も抽出してしまうと述べている。また、文字列抽出では活用によって語尾が変化する用言についての分析が困難である。

現在では、古文用 UniDicS[6]を用いることにより、古文においても形態素解析が可能となっている。本研究では、歌語の共起語を分析することを目指すため、文字列単位で分析する N-gram ではなく、単語単位で分析する形態素解析という手法をとった。

## 2.5 本研究の位置づけ

本研究では、歌語の共起語から歌語の潜在的な意味を見出すことで、美的観念の時代変遷を明らかにする。実験に用いる和歌のデータセットは国際日本文化研究センター[9]の和歌データベースから、かな文字で清音表記された三大歌集を抽出した。これまで定義が曖昧

だった歌語を、西方[4]の『雅語・歌語 五七語辞典』を参考に歌語データベースを構築し、和歌データベースとマッチングをすることで、和歌データから歌語を抽出し、どういった歌語が、いつの時代の歌集で、どのように用いられているか明らかになる。また、歌語を和歌における美的観念を反映した言葉として扱うことで、時代ごとに歌語の共起語を分析し、歌語の潜在的な意味を明らかにすることが可能である。

本研究では、工藤[10]が開発したオープンソースの形態素解析エンジンである MeCab と、国立国語研究所が開発した解析辞書の古文用 UniDicS[6]を用いて和歌の形態素解析を可能とした。また、形態素解析により見出し語化をすることで、これまで中心的に研究がなされてきた名詞などの体言だけではなく、動詞や形容詞などの用言を考察対象にすることを可能とした。計算機を用いて単語の共起関係を分析することで、これまでの研究では重要視されてこなかった単語の潜在的意味を網羅的に把握することを可能とする。

### 3. 歌語の共起語変化による美的観念変遷発見

本節では、提案方式である、和歌における歌語の共起語変化による美的観念の時代変遷発見方式について提示する。

3 節の構成について述べる。3.1 節では、本研究の全体像について述べる。3.2 節と 3.3 節では、本研究で扱う和歌と歌語のデータセットについて述べる。3.4 節では、形態素解析機能について述べる。3.5 節では、歌語抽出機能について述べる。3.6 節では、潜在的意味抽出機能について述べる。

#### 3.1 全体像

本節では、本研究における提案手法の概要を述べる。提案システムの全体像を図 1 に示す。本システムは、形態素解析機能、歌語抽出機能、潜在的意味抽出機能及び、和歌データセット、歌語データセットで構成される。形態素解析機能は、MeCab[10]と、解析辞書である古文用 UniDicS[6]を用いて、和歌データを形態素解析する機能である。歌語抽出機能は、歌語データセットをもとに作成した MeCab で用いるユーザー辞書を用いて、和歌から歌語を抽出する機能である。潜在的意味抽出機能は、単語の意味は共起している単語によって形成されるという分布仮説[11]に基づき、歌語の共起行列[12]を作成し、ベクトルに表現された歌語を、cos 類似度を用いて意味的類似度を計算することで、歌語の潜在的な意味を抽出する機能である。和歌データセットは、国際日本文化研究センター[9]の和歌データベースから、三大歌集である万葉集、古今和歌集、新古今和歌集のうち、清音で表記されたデータを抽出したものである。歌語データセットは、上代から

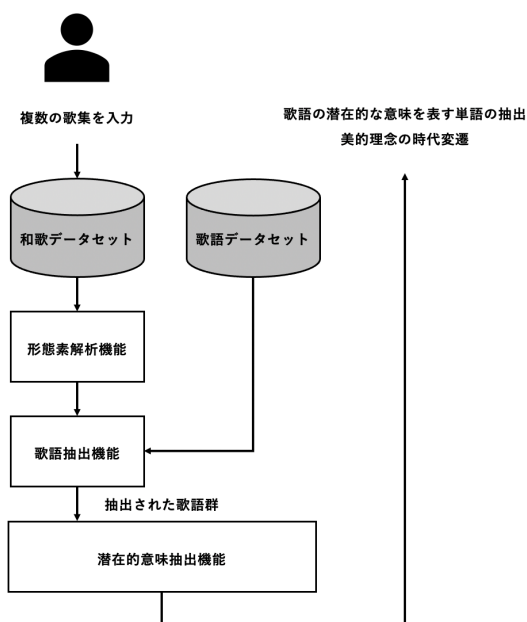


図 1 本方式の全体像

昭和前期に至るまでの和歌、連歌、短歌などから採集した五音七音表現が収録された西方[4]による『雅語・歌語 五七語辞典』から、使用例をもとに歌語を「天象」や「心」など 20 の大分類と、さらに心の中でも「恋」や「憂」など 272 の中分類にカテゴライズされたものを 3673 語収録したものである。これらの機能により、歌語の潜在的な意味を表す単語を抽出することが可能となる。これにより、三大歌集における美的観念の時代変遷を明らかにすることが可能となる。

本研究の目的は、三大歌集における美的観念の時代変遷を明らかにすることである。本研究で用いる三大歌集とは、奈良時代に成立した万葉集、平安時代に成立した古今和歌集、鎌倉時代に成立した新古今和歌集の総称を指す。本研究において、歌語は、和歌の構造における中心的要素であり、当時の美的観念を反映した言葉とする。

#### 3.2 和歌データセット

本研究では、国際日本文化研究センター[9]の和歌データベースを実験に用いる。和歌データベースから、三大歌集である万葉集、古今和歌集、新古今和歌集のうち、清音で表記されたデータを抽出した。また、本研究で用いた歌の数は、万葉集が 4536 首、古今和歌集が 1111 首、新古今和歌集が 2005 首である。形態素解析するにあたり、漢字で表記されたデータを用いる方が精度は高いが、万葉集と新古今和歌集においては、漢字表記のデータがなく、清音で表記されたデータのみであったため、三大歌集で比較する際には、全て清音表記のデータを格納することにした。

### 3.3 歌語データセット

本研究では、歌語データセットを構築するにあたり、上代から昭和前期に至るまでの和歌、連歌、短歌などから採集した五音七音表現が収録された西方[4]による『雅語・歌語 五七語辞典』から、使用例をもとに歌語を「天象」や「心」など 20 の大分類と、さらに「心」の中でも「恋」や「憂」など 272 の中分類にカテゴライズされたものを 3673 語収録した。本研究では、三大歌集における 20 の大分類のシェアを調査する。

また、実験に用いる和歌データは歴史的仮名遣いが用いられた清音表記データのため、歌語データセットを構築する際は和歌データに表記を合わせ、全て歴史的仮名遣いに修正した。

### 3.4 形態素解析機能

本研究では、和歌データにおいて、体言のみならず、動詞や形容詞など、用言には心情を表す語が多く存在しているため用言も対象とした。しかし、用言は形態素解析による見出し語化が必要である。

和歌などの古典文学資料においては、歴史的仮名遣いで書かれているものがほとんどであり、そういった資料を形態素解析するためには古文に対応した辞書が必要である。本研究では、形態素解析用の辞書として、古文用 UniDicS[6]を用いることとした。それにより、歴史的仮名遣い資料を直接扱うことが可能となる。

上代には、平安時代初期に生まれた、平仮名や片仮名とは大きく異なる上代特殊仮名遣と呼ばれる特殊な仮名遣と語法が存在した。そのため、本研究における和歌データセットのうち、奈良時代に成立した万葉集と、平安時代以降に成立した古今和歌集、新古今和歌集は、解析辞書を使い分ける必要がある。古文用 UniDicS のうち、万葉集の解析においては上代 UniDic を、古今和歌集と新古今和歌集の解析においては中古和文 UniDic を用いた。

### 3.5 歌語抽出機能

歌語抽出機能では、構築した歌語データセットにより、形態素解析をした三大歌集を含む和歌データから歌語を抽出し、三大歌集における、意味ごとに分類した歌語のカテゴリ別のシェアを抽出する。

上記 3.3 節で述べた、「天象」や「心」などの 20 の大分類と、大分類をさらに細かく分類した、「心」の中でも「恋」や「憂」などのような 272 の中分類のカテゴリを用いて、歌語のカテゴリ別シェアを導出する。それにより、歌集ごとに和歌に詠み込まれた大まかな内容が明らかになる。

### 3.6 潜在的意味抽出機能

本研究では、歌語の出現頻度による共起関係を明らかにする。本研究では、抽出した歌語の前後 2 単語を共起語とする。また、単語の意味は共起語により形成

表 1 作成した共起行列の例

	トシ	ハル	...	ヨソ メ	フル マイ
トシ	2	3	...	0	0
ハル	3	2	...	0	0
...	...	...	...	...	...
ヨソ メ	0	0	...	0	0
フル マイ	0	0	...	0	0

されるという分布仮説[11]に基づき、歌語の共起行列[12]を作成し、共起行列の各行を抽出することによって、各単語をベクトルで表現することが可能となる。共起行列は、行、列とも和歌に出現する歌語を各要素とし、単語前後 2 単語で共起した場合は 1、それ以外は 0 とする行列である。なお、共起行列は表 1 のような正方行列であり、万葉集は 895 行 895 列、古今和歌集は 658 行 658 列、新古今和歌集は 743 行 743 列である。

ここで、ある単語  $x$  とある単語  $y$  を考えたとき、分布仮説により、単語  $x$  に共起する単語群と単語  $y$  に共起する単語群が類似しているのであれば、単語  $x$  と単語  $y$  は類似した意味を持つと仮定する。ここで、共起行列の各単語に対応する行を取り出し、二つの行ベクトルの  $\cos$  類似度を計算することは、この二つの単語間の意味的類似度を計算することに値する。 $\cos$  類似度を用いて、ベクトル化した単語の類似度を計測することで、歌語の潜在的な意味を見出すことができる。

歌語の潜在的な意味を見出すことで、当時の美的観念を抽出することが可能となる。これにより、美的観念の時代変遷を発見することが可能となる。

共起語を抽出する際、「こと」や「もの」など、どの和歌においても、高頻度で出現する単語を含む場合、どの単語においても類似するとみなされる恐れがある。そのため本研究では、単語の出現回数による影響を軽減するために、Pointwise Mutual Information (PMI) [13]を用いることとする。また、本研究では、PPMI (Positive PMI)を用いる。PPMI によって補正された共起行列  $C'$  は下記のように表される。

$$PMI(x, y) = \log_2 \frac{P(x, y)}{P(x)P(y)} = \log_2 \frac{\frac{C(x, y)}{N}}{\frac{C(x)}{N} \frac{C(y)}{N}} = \log_2 \frac{C(x, y) \cdot N}{C(x)C(y)},$$

$$\text{PPMI}(x, y) = \max(0, \text{PMI}(x, y)),$$

$$\mathbf{C}' = [\text{PPMI}(\mathbf{c}_i, \mathbf{c}_j)].$$

$\text{PMI}(x, y)$  は単語  $x$  と単語  $y$  が共起する出現確率を表し、共起行列を  $\mathbf{C}$ 、コーパスに含まれる単語数を  $N$  とする。  $\mathbf{c}_i$  は共起行列  $\mathbf{C}$  の  $i$  番目の行を抽出したベクトルである。

ここで、共起行列の次元削減を行うために、固有値分解を用いる。

$$\mathbf{C}' = \mathbf{V}\Sigma\mathbf{V}^{-1}$$

固有値分解によって導出された  $\mathbf{V}$  を次元削減された共起行列として用いる。これにより、重要ではない要素を削ることが可能となる。

次元削減された共起行列  $\mathbf{V}$  の  $i$  番目、  $j$  番目の行である  $\mathbf{v}_i$ 、  $\mathbf{v}_j$  の  $\cos$  類似度  $\text{sim}(\mathbf{v}_i, \mathbf{v}_j)$  を計算することによって、それに対応する単語同士の意味的類似度を計算することができる。

$$\text{sim}(\mathbf{v}_i, \mathbf{v}_j) = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{|\mathbf{v}_i||\mathbf{v}_j|}$$

## 4. 実験

本節では、本手法の実験内容と結果、考察について述べる。

4.1 節では、実験環境について述べる。4.2 節では、三大歌集における歌語の偏りについて述べる。4.3 節では、三大歌集における歌語の共起語変遷についての検証結果を述べる。4.4 節では、本研究による実験結果について考察を行う。

### 4.1 実験環境

本節では、3 節で提案したシステムを実装し、共起語を抽出した。

本研究では、形態素解析をするにあたり、MeCab と解析辞書として、古文用 UniDicS の上代 UniDic と、中古和文 UniDic を用いた。

補正した共起行列  $\mathbf{C}'$  を固有値分解したことにより、次元圧縮された共起行列  $\mathbf{V}$  は 100 行 100 列の行列となる。

### 4.2 実験 1(三大歌集における歌語の偏り)

本節では、三大歌集で用いられた歌語における 3.3 節で述べた大分類の割合について集計する。実験結果は表 2 に示す。

三大歌集に共通して多く見られたのは、「水」や「山」など、「地理」に関する歌語であった。また、季節や、「朝」、「夜」など時間帯に関連する歌語である、「時」に関する歌語も多く見られた。

万葉集において、「消」や「興」、「美」など、「状態」に関する歌語が多く用いられているという特徴が見られた。このことから、万葉集は写実的な和歌が多いといえる。また、万葉集において、「食」や「住」、「仕事」、

表 2 三大歌集における歌語の偏り

カテゴリ名	万葉集	古今和歌集	新古今和歌集
天象	6.24%	7.35%	9.91%
地理	11.56%	12.24%	12.55%
位置・形	3.64%	3.97%	5.03%
数・量	7.82%	6.96%	6.67%
時	8.90%	10.34%	11.19%
音・色 灯・火	3.97%	4.71%	4.29%
状態	7.06%	5.82%	6.39%
心	9.08%	11.34%	9.72%
体	5.52%	6.11%	5.74%
衣	3.00%	3.36%	2.06%
食	2.19%	0.76%	0.93%
住	1.65%	0.67%	1.09%
仕事	4.72%	3.96%	4.07%
往来	7.91%	5.94%	6.37%
技芸	0.32%	0.30%	0.33%
思考	4.40%	4.71%	4.66%
生死	1.34%	1.95%	1.77%
神仏	0.65%	0.37%	0.36%
人	3.37%	3.76%	2.71%
動物・植物	6.64%	5.38%	4.16%

「動物・植物」など日常生活に関する歌語や、「行」や「歩」、「舟」など「往来」に関する歌語が他の歌集よりも多く見られた。これらは、万葉集には皇族や貴族のほか農民が詠んだ歌や、地方の歌が多く収められていることから、このような特徴が見られたと考えられる。

古今和歌集において、「心」に関する歌語が多く用いられているという特徴が見られた。紀貫之は、『古今和歌集仮名序』[14]において、「やまとうたは、人のこゝ

ろをたねとして、よろづのことはとぞなれりける。」と述べている。これは、和歌は人の心を種として様々な言葉となったものであるという意味であり、和歌において、心に思っていることが重要視されていることがわかる。以上より、古今和歌集において、「心」に関する歌語が多く用いられたと考えられる。また、表 3 にあるように、「心」に関する歌語は、歌集によって多く用いられているものが異なっていた。万葉集では、大分類「心」において「睦まじ」などを含む中分類「睦」や、「逢瀬」などを含む中分類「逢」に関するものが多く、古今和歌集や新古今和歌集では、「憂し」などを含む中分類「憂」に関するものが多く見られた。

新古今和歌集においては、他の歌集に比べ「天象」に関する歌語が多く見られた。また、新古今和歌集において、古今和歌集の成立時期には重要視されてきた「心」に関する歌語の使用割合が減少している。これは、新古今和歌集成立時期において、歌語の抽象化が進み、象徴的な表現として用いられる機会が増えたことより、直接的に感情を表現する言葉である「心」に関する歌語の使用が減少したと考えられる。

#### 4.3 実験 2(三大歌集における歌語の共起語変遷)

本節では、三大歌集における歌語の共起語変遷についての検証結果を述べる。本実験では、三大歌集において、多く出現した歌語の大分類のうち、上位 3 件である「地理」、「心」、「時」に注目する。また、季節とは、如何なる人にも等しく訪れるものであり、和歌において、万人に共通する美意識の基盤である。以上を踏まえ、本実験では、上位 3 件のうち、和歌において最大のテーマである季節に関する歌語の大分類「時」に着目し、その中で三大歌集に共通して出現回数の多かった「アキ」、「ハル」における共起語変遷を示す。各歌集における「アキ」、「ハル」の共起語をそれぞれ、表 4、表 5 に示す。尚、表は、歌語抽出機能において現代仮名遣いに変換された単語で記す。

三大歌集における、「アキ」の共起語変遷において、万葉集では、「ハギ」や「ハナ」など植物、または、「シラツユ」など自然に関する歌語が共起語として抽出された。また、古今和歌集においては、「ウシロメタシ」や「マドウ」、「ワク」、といった共起語が抽出された。「ウシロメタシ」には、「先が気がかりである」といった意味がある。また、「マドウ」には「思い悩む」、「途方に暮れる」という意味があり、「ワク」には「区別する」という意味の他に「理解する」という意味がある。また、新古今和歌集では、「ツキ」といった自然に関する言葉に加え、「カナシ」、「ユメ」などの単語が抽出された。「カナシ」は、「心が引かれる」や「いとしい」の他に、「切ない」、「悲しい」といった意味がある。また、「ユメ」は、現代語と同じ睡眠中にもつ幻覚である

表 3 大分類「心」における中分類の内訳

順位	万葉集	古今和歌集	新古今和歌集
1	思：804	思：213	思：286
2	睦：660	心：151	心：223
3	逢：490	憂：120	憂：200

「夢」という意味の他に、「非現実的なこと」や「儚いこと」、「迷い」、「煩惱」という意味がある。「アキ」の共起語変化において、万葉集では秋に関する植物や、自然に関する歌語が多く見られ、古今和歌集や新古今和歌集においては、心情を表す歌語が見られるようになった。このことから、「アキ」は季節としての「秋」から、物悲しさや儚さの象徴として用いられるようになったと考えられる。

三大歌集における、「ハル」の共起語変遷において、万葉集では、「サク」や「ウメ」、「ニシキ」、「フユ」といった共起語が抽出された。「ハル」の共起語として「ウメ」が抽出されたことから、万葉集成立時期においては「ウメ」が春を代表する花であったと考えられる。また、「フユ」が共起語として抽出されたことから、「ハル」と「フユ」を比較した和歌が多かったと考えられる。また、古今和歌集では、「アラレ」、「クラシ」、「チル」、「オリ」、「オモウ」などが抽出された。冬の季節である「アラレ」が共起語として抽出されたことから、万葉集同様「ハル」と「フユ」を比較した和歌が多かったと考えられる。「クラシ」は「はっきりとしない」という意味、「オリ」は「季節」という意味、「オモウ」は「感じる」、「恋しく思う」などの意味がある。また、新古今和歌集では、「ニオイ」や「チル」、「ユキ」、「ハナ」、「アラレ」などの共起語が抽出された。「ハル」の共起語として「ユキ」や「アラレ」が共起語として抽出されたことから、他の 2 つの歌集と同様「ハル」と「フユ」を比較した和歌が多かったと考えられる。また、「ハル」と「ユキ」や「チル」が共起していることから、新古今和歌集成立時期において、花が散る様子を雪に例える表現が用いられるようになったと考えられる。また、万葉集においては「ハル」の共起語として「サク」が抽出されているのに対して、古今和歌集、新古今和歌集では「チル」が共起語として抽出された。このことから、万葉集においては、花が咲いている状態を詠んだ和歌が多く、古今和歌集や新古今和歌集においては、花が散っている状態を詠んだ和歌が多いこ

表4 「アキ」における共起語変遷

順位	万葉集	古今和歌集	新古今和歌集
1	スタル 0.418017	コチ 0.304806	ツキ 0.353439
2	ノウ 0.379018	ウシロメタシ 0.299478	カナシ 0.347545
3	ハギ 0.367616	スム 0.298904	タノモシ 0.318466
4	ハツ 0.357869	マドフ 0.282218	クニ 0.292127
5	シラツユ 0.322289	ヒカリ 0.281639	ツク 0.291183
6	ハタ 0.298537	ワク 0.276938	ユメ 0.245635
7	ハナ 0.293763	ホコ 0.271563	ソトモ 0.243579
8	キサキ 0.269188	タラチネ 0.268937	アケ 0.242285
9	ケゴロモ 0.265690	オヤ 0.268102	ミナト 0.236545
10	アナ 0.254703	オオシ 0.267452	コトワリ 0.232542

表5 「ハル」における共起語変遷

順位	万葉集	古今和歌集	新古今和歌集
1	サク 0.418550	アラレ 0.370667	ニオイ 0.303579
2	ウメ 0.297421	カタチ 0.324673	カヨウ 0.297330
3	ヨメ 0.294437	クラシ 0.280130	チル 0.274557
4	ミヤコ 0.274386	カシラ 0.279385	ユキ 0.263313
5	マツ 0.271563	チル 0.250121	ハナ 0.260799
6	ニシキ 0.269017	モチ 0.245264	オオウ 0.249196
7	ヤマ 0.265310	オリ 0.238831	ヒカリ 0.249017
8	フユ 0.261468	ヌスム 0.236374	カタ 0.239601
9	オオル 0.261086	オモウ 0.232836	アラレ 0.238414
10	モユ 0.260664	チグサ 0.222344	シラム 0.232138

とがわかる。以上より、「ハル」の共起語変化において、万葉集では春に関する自然現象や、春の華やかさを表す単語が見られ、古今和歌集では、「オモウ」といった心情を表す単語が見られた。また、三大歌集全てにおいて、「ハル」の共起語として「フユ」や「アラレ」、「ユキ」といった冬に関する歌語が抽出された。また、「ハル」の共起語として、万葉集においては「サク」、古今和歌集、新古今和歌集では「チル」と花の状態に関する共起語に変化が見られた。このことから、時代が経つにつれ、花が咲き誇っている状態から花が散っている状態へと美的観念が大きく変化したといえる。

#### 4.4 考察

本節では、三大歌集における歌語の分野別使用割合や歌語の共起語から、歌語の潜在的な意味を抽出した。奈良時代に成立した万葉集においては、4.2 節の表 2 にあるように、「消」や「興」、「美」など、「状態」に関する歌語が多く見られたことや、4.3 節の表 4、5 にあるように、自然に関する共起語が多く、心情を表現する歌語があまり見られなかったことから、万葉集に収められている和歌は、直接的な表現が用いられるこ

とが多く、写実的なものが多いと考えられる。また、4.2 節の表 2 にあるように、「食」や「住」、「仕事」、「往来」、「動物・植物」に関する歌語が多く用いられていたことから、日常生活に関しての和歌が多いといえる。また、「往来」に関する歌語が多く用いられていたのは、地方の歌が収められていることが影響していると考えられる。

平安時代に成立した古今和歌集においては、4.2 節の表 2 にあるように「心」に関する歌語が多く用いられているという特徴が見られた。また、4.3 節の表 4、5 にあるように、「アキ」、「ハル」において心情を表現する「ウシロメタシ」、「マドウ」、「オモウ」などの歌語が共起語として抽出された。これより、古今和歌集に収められている和歌には心のうちを表現した和歌が多いと考えられる。

鎌倉時代に成立した新古今和歌集においては、4.2 節の表 2 にあるように、古今和歌集の成立時期には重要視されてきた「心」に関する歌語の使用割合が減少している。このことから、新古今和歌集において、心情を直接表現せず、婉曲的な表現を用いている和歌が

増加したと考えられる。

以上の実験より、三大歌集における歌語の共起語変化により、当時の美的観念や、歌集における表現の違いが抽出できた。

## 5. おわりに

本稿では、歌語の共起語変化に着目した、三大歌集における美的観念変遷発見方式について述べた。本研究では、和歌において中心的な要素となる歌語を共起語から、歌語における潜在的な意味を抽出した。これらを成立時期の異なる複数の歌集で比較することで和歌に反映されている当時の美的観念における時代変遷発見方式を実現した。

また、本方式を検証するにあたり、国際日本文化研究センター[9]による和歌データベースと、西方[4]の『雅語・歌語 五七語辞典』を用いて、本方式の有効性を検証する実験を行った。工藤[10]が開発した MeCab と、古文用 UniDicS[6]を用いて和歌の形態素解析をし、見出し語化をすることで、用言を考察対象とすることが可能となった。

三大歌集における、歌語の潜在的な意味の変遷発見方式を実現したことにより、古典文学の表現史における新たな視点や分析、また、当時の歌人の感性や作品における新たな解釈を発見することが可能であると考えられる。

また、このような計算機を用いた古典文学作品による網羅的な研究は、これまで発見されてこなかった表現や解釈など、新たな知見をもたらし、人文学研究の発展に寄与できると考えられる。

今後の課題として、古典文学愛好者を増やすための分析可視化ツールの実現、それぞれの時代が持つ思想、もしくは風俗が書かれた文献と和歌などの古典文学の特徴の相関を導出することによるこれまでに見出せなかった深い時代背景の発見、詠み人知らずなどの匿名作品の分析による著者背景の新たな発見が挙げられる。

## 参 考 文 献

- [1] 久保田淳, 「歌語の変遷」, 『言語』 第二十五巻第六号, 1994
- [2] 井手至, 「歌ことばの本質」, 『女子大國文』 第百十六号, 1994
- [3] 佐藤武義, 歌語論の展開と現状, 国際文化研究科論集, 1996
- [4] 西方草志, 『雅語・歌語 五七語辞典』, 2012
- [5] 百留恵美子, 和歌表現史の研究, 2005
- [6] 古文用 UniDicS, <https://unidic.ninjal.ac.jp/>
- [7] 竹田正幸, 福田智子, 失われゆく情報の復元・保存技術-人文学における情報処理-(文献学・データベース共有・史料編纂): 1. 古典和歌からの知識発見-モバイルスーツを着た国文学者-, 情報処理, 2002
- [8] 近藤みゆき, 平安時代和歌資料における特殊語彙抽出についての計量的研究と利用ツールの公開, 文部省科学研究費補助金 1998 年度研究成果報告書『特定領域研究 人文科学とコンピューターコンピュータ支援による人文科学研究の推進一』, 1999
- [9] 勢田勝郭, 和歌データベース, 国際日本文化研究センター, <http://db.nichibun.ac.jp/sp1/ja/category/waka.html>
- [10] 工藤拓, MeCab, <https://taku910.github.io/mecab/>
- [11] Zellig S. Harris., "Distributional structure", Word, Vol.10, pp.146-162, 1954
- [12] 斎藤康毅, 『ゼロから作る Deep Learning②—自然言語処理編』, 2018
- [13] Kenneth Ward Church and Patrick Hanks, "Word association norms, mutual information, and lexicography", Comput. Linguist, Vol.16, Number 1, pp.22-29, 1990
- [14] 紀貫之, 『古今和歌集仮名序』, 905