

中古中世歌合の構造化

富士池優美
yfujiike@ninjal.ac.jp

人間文化研究機構 国立国語研究所

1. はじめに

中古中世の歌合は、中古から中世にかけて和漢混淆文が一般化する過程において、和歌の実作に基づき、和歌のあり方や歌ことばの用法について評論が加えられた資料である。歌合コーパスは①和歌、平仮名漢字交じり文、和化漢文といった異なる文体が混在し、語彙・文体の研究上重要な位置にあること、②序文・歌・判詞・日記等、多様な要素を持ち、必要に応じた検索を可能にするコーパスのタグ付けには慎重な設計が求められることの2点において重要である。本発表では、中古の代表的な歌合である「天徳四年内裏歌合」を例に、形態論情報付きコーパス構築の前処理として、資料の特性に応じたテキストの構造化(タグ付け)のあり方を検討する。

2. 歌合コーパスの意義

歌合とは、和歌を左右に分けてつがわせ、その優劣を判定した文芸的な遊戯であるとともに、その評論でもある。平安時代中期の遊楽の中で代表的な行事となり、平安時代末期から鎌倉時代初期にかけて、特に高度な文芸的内容を持つ行事に発展した。本研究で対象とする中古中世の歌合は、和歌の実作に基づき、和歌のあり方や歌ことばの用法について評論が加えられた資料と言える。歌合の中で特に注目されるのが、和歌評論である判詞、歌の勝負に関する評や議論を記した部分である。

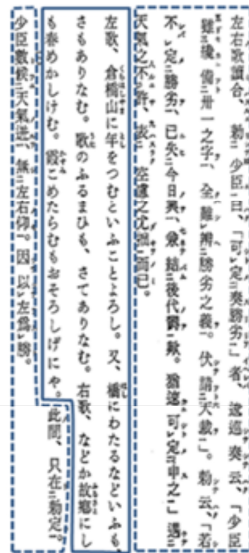


図1 歌合の判詞
実線: 和文体、破線: 和化漢文体

歌合の判詞では判詞として平仮名漢字交じり文、和化漢文といった異なる文体を交えて、一つの論をなすことがある。図1では和文体(実線)の箇所「なむ」「けむ」といった口語的表現が見られるのに対して、和化漢文体(破線)の箇所では「いはく」

「いへども」といった漢文訓読的表現が見られ、判詞の文体によって用語に差がある様子が見てとれる。

中古中世は和漢混淆文が一般化する過程にある。平仮名文(和文)についてはこれまでの日本語史研究の中心資料であるとともに、コーパスの構築も進められている。また、和漢混淆文についても、コーパスに基づき、和漢の両系統の連続性の解明、文体的変異の構造を記述を目的とする研究が進められつつある。歌合判詞に見られるような文体の混在は中古中世の物語、説話等にはない特徴であり、和漢混淆文が一般化する過程の実態を探るにあたり、歌合判詞は語彙・文体の研究上重要な位置にある。しかし、歌合には用語索引類がなく、既存の電子化資料は紙面に基づく外形的な電子化であるため、言語研究を目的とした場合には不足があり、語彙・文体的特徴を捉えるのは現状では困難である。当時の言語意識を捉える材料としてのみ用いられてきた歌合であるが、判詞を和漢混淆文の一形態としてその語彙を記述する研究は進んでおらず、その方面の研究が求められている。

歌合は序文・歌・判詞・日記といった多様な要素を持つ。これらの多様な要素の扱いの困難さも、歌合を対象にした語彙研究の進展を阻む一因であったと考える。このような多様な要素を持つ資料の電子化方法については、「日本語歴史コーパス 平安時代編」において、中古和文を対象に地の文・会話・手紙・和歌・詞書のタグ付けを行っているほか、中近世の狂言、洒落本といった資料の扱いの検討が進められている。歌合についても同様に、どの要素をどのように構造化するのが適切かということ、検討する必要がある。

3. 歌合テキスト構造化の問題点

3.1 外形的要素

歌合は序文・歌・判詞・日記等、多様な要素を持つが、その開催目的や規模等によって様式が異なる。

図2は天徳四年内裏歌合の様式を示したものである。ここでは、番ごとに題、番が示され、左右それぞれに方、歌人、歌が示され、左右どちらかに判が

付き（勝敗、図2では左の「勝」）、番の後に判詞が付く。

図3は天徳四年内裏歌合の全体構造である。冒頭に目録（題・家人・講師・判者）が置かれ、末尾に5種の日記が添えられる。歌合の全体構造は大きく、{歌合全体一番一歌}の三層構造になっている。この構造は歌合共通のものであるが、題の付き方、目録・序文・日記、題、判、判詞といった要素は歌合によって有無が異なる。

このように歌合の各要素は、階層的な構造を持ちながらも個々の要素の配置が一定しないため、ある程度の自由度を持たせることが重要な課題となる。

3.2 和化漢文部分の扱い

歌合は、和歌、平仮名漢字交じり文といった和文系統と和化漢文といった漢文系統の文体が入り交じることがある資料である。中でも図1や図2に示したように、判詞という一つの要素の内部に和文と和化漢文が混在することが問題になる。勅判は和化漢文で判者の判詞は平仮名主体といった形で、判者によって文体が異なるものもあれば、定型的な表現のみ和化漢文として表記されるものもあり、その入り交じり方は歌合によって様々である。この実態を踏まえると、形態素解析を前提とした歌合コーパスとしては、和化漢文と和文とで同様の扱いができることが望ましい。具体的な対処としては、表記が省略された部分の補読、和化漢文を書き下し形式にするといった和化漢文を和文風表記に改めた上で和化漢文表記であった情報を残すといったテキスト整形処理が考えられる。これらを形態素解析の前処理として行う必要がある。

4. 歌合コーパスの構造化

4.1 設計方針

歌合コーパスの底本としては、校訂済み本文を持ち、形態論情報整備の際に参考となるルビや注とい

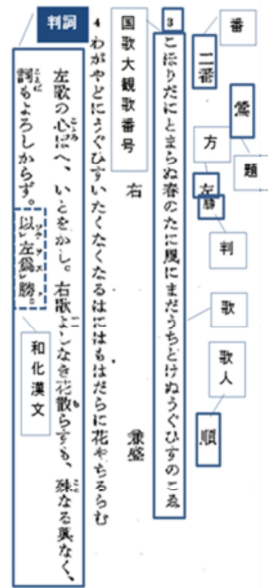


図2 歌合の様式

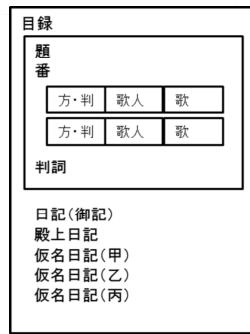


図3 歌合の構造

った情報を持つものを選択した。例えば「天徳四年内裏歌合」では『歌合集』（日本古典文学大系）を用いた。既存の電子テキストに国文学研究資料館の「大系本文データベース」があるが、序文・歌・判詞・日記といった要素の情報がないため、言語研究を目的とした場合には不足がある。そこで、歌合特有の構造に対し、可能な限り細かく情報を付与することとした。これによって、和歌と和歌評論の組み合わせと言える歌合を、和歌は他の資料の和歌と、判詞は歌論・歌学書と、それぞれ比較するといったことが可能になる。

本コーパスでは、得られた用例が属する要素、またその語の代表形（語彙素）・読み（語彙素読み）・品詞といった形態論情報を取り出せることを目指す。そのため、記述にはXMLを用い、国立国語研究所が作成した「太陽コーパス」、BCCWJ、「明六雑誌コーパス」、構築中の『虎明本狂言集』コーパス、「洒落本コーパス」、「日本語歴史コーパス 平安時代編」、和漢混淆文を対象としたコーパス等の仕様を参考に、歌合特有の構造にあわせて構造化する。構造化したデータには、形態素レベルでの代表形・読み・品詞といった形態論情報を付与する。

3.2節で示したように、歌合コーパスにおいては判詞という一つの要素の内部に和文と和化漢文が混在することが問題になる。これについては、和化漢文で書かれた部分に対して和文風表記になるようテキストの整形を行うこととした。具体には、図1破線部分のような、和化漢文部分に一二点やレ点といった返り点が付されている箇所を全て書き下した上でコーパス化した。また、形態素解析にあたり問題となる、表記が省略された部分の補読や、読まない文字のタグ付け、漢字で表記された助詞・助動詞を仮名に開くといった、漢文系統の文体のコーパス化において必要な処理を行うこととした。

4.2 構造とタグセット

歌合のコーパス化にあたっては、3.1節に示したように、配置が一定しない要素について自由度を持たせてマークアップする枠組みを設定することが求められる。要素の洗い出しにあたっては、『歌合集』所収歌合のほか、平安初期歌合や大規模な歌合である『六百番歌合』等をあわせて検討した。

(1) 文書の構造に関する要素

一般的な文書構造に関する要素として、テキスト全体を表す<text>とそれを構成する<front><body>から成る。<title>は歌合全体のタイトルである。

s 要素 テキストは文に分割する。割書の前後等、文の断片を認める場合がある。和歌内部については

文を認定しない。<s>に<s>が含むような階層性は持たせない。

SUW 要素 言語研究での基本的な単位である短単位を表す。テキスト全てが短単位に分割される。中古和文 UniDic による解析結果を人手修正して付与する。(図 4 では省略)

(2) 歌合特有の構造に関する要素

budate 要素・dai 要素・ban 要素 <budate>は和歌集の部立にあたる、複数題をとりまとめた題を表す。「天徳四年内裏歌合」では題と番の対応は 1 対 1 であるため該当する要素はない。<dai>は歌題を表す。<ban>は歌の番(左右で一对)をまとめたもので、@banID 属性(必須)で番の通し番号を記述する。一つの題に複数番が対応する「六百番歌合」では「春上」を<budate>とし、<dai>に「元日宴」「余寒」「春氷」「若草」等の題が入り、各題は 6 番から成る。

banNum 要素 テキスト上の番の番号を表す。

hidari 要素・migi 要素 方(かた)を表す。左方を<hidari>、右方を<migi>とする。

syobu 要素 勝負を表す。左右のどちらか勝った方が「勝」となり、引き分けの場合は左方の syobu 要素が「持」となる。勝負を付けない歌合も存在する。

yomibito 要素 歌の詠み人を表す。

uta 要素 和歌を表す。@no 属性で新編国歌大観番号、@utaID で歌合内の通し番号、@type 属性で重出歌であることを記述する。

han 要素 判詞を表す。

nikki 要素 日記や歌合前後の贈答歌のような付属的な文章を表す。

nikkiTitle 要素 日記類の題(名称)を表す。

kotobagaki 要素 歌合歌には原則として詞書はないが、日記や歌合前後の贈答歌には詞書が付される

ことがある。和歌集には原則として詞書があるが、和歌とは別に扱われる。またその形式は必ず 1 文になるといった、地の文とは異なる特徴を持つことから、歌合においても独立した要素とすることとした。

(3) 語・文字単位で外形・機能を表す要素

warigaki 要素 底本テキストで割書であったことを示す。

kanbun 要素 正格漢文ないし和化漢文のような漢文の影響を受けた表記であったことを示す。対象は文全体ないし文の一部である。

kanbun 要素 正格漢文ないし和化漢文のような漢文の影響を受けた表記であったことを示す。対象は文全体ないし文の一部である。

hodoku 要素 和化漢文のような漢文の影響を受けた表記において、ルビにはあるが本文に対応する文字がない場合、hodoku 要素として補入する。

fudoku 要素 和化漢文のような漢文の影響を受けた表記において読まれない文字(不読文字)を表す。漢文における置き字に相当するもので、形態素解析の対象外となる。

ruby 要素 文字の右側に付された、文字・文字列の読みを表す振り仮名等を表す。@rubyText 属性(必須)にルビ文字列を記述する。右側漢字傍記を含むほか、<kanbun>内において、和化漢文内の助詞・助動詞を仮名に開いた際の元の文字(漢字)の記述を含む。

odoriji 要素 踊り字・繰り返し記号であったことを表す。これらは開いた形とし、@originalText 属性(必須)に元の文字・文字列を記述する。

missing 要素 破損・欠損を表す。

(4) 位置情報を表す要素

pb 要素 底本テキストのページ開始を表す(空要素)。@n 属性でページ番号を記述する。

```
<text><title>天徳四年三月三十日内裏歌合</title><pb n="78"/><front>内裏和歌合<warigaki><s><kanbun>天徳四年三月
卅日、清涼殿於此事有。</kanbun></s></warigaki> (略) </front><body><dai>霞</dai><ban banID="0001"><banNum>
一番</banNum><hidari>左</syobu>勝</syobu><yomibito>朝忠</yomibito><uta no="0001" utaID="0001"><ruby
rubyText="くら">倉</ruby><ruby rubyText="はし">橋</ruby>の<ruby rubyText="やま">山</ruby>のかひより春がすみ
としをつみてやたちわたるらむ</uta></hidari><migi>右</yomibito>兼盛</yomibito><uta no="0002" utaID="0002">ふる
さとは春めきにけり<ruby rubyText="み">三</ruby><ruby rubyText="よし">吉</ruby>野の<ruby rubyText="み">御
</ruby><ruby rubyText="かき">垣</ruby>の<ruby rubyText="はら">原</ruby>をかすみこめたり</uta></migi>
<han> (略) <s>左歌、<ruby rubyText="くら">倉</ruby><ruby rubyText="はし">橋</ruby><ruby rubyText="やま">山
</ruby>に<ruby rubyText="とし">年</ruby>をつむといふことよろし。</s> (略) <s>右歌、などか<ruby rubyText="ふる
さと">故郷</ruby>にしも春めかしけむ。</s><s><ruby rubyText="かすみ">霞</ruby>こめたらむもおそろしげにや。
</s><s><kanbun>此ノ間、只勅定ニ在リ。</s><s>少臣數々天氣ノ逆ルヲ候フニ、左右ノ仰セ無シ。</s><s>因ツテ左ヲ以
テ勝ト爲ス。</s></kanbun></han> (後略) </ban> (後略)
```

図 4 「天徳四年内裏歌合」冒頭の形式化例

<s><kanbun>仰セニ云ハク、</kanbun></s><s><kanbun>「定メ申ストコロニ據ル<ruby rubyText="可">べ</ruby>シ」</kanbun></s><s><kanbun>ト<ruby rubyText="イヘレバ">者</ruby></hodoku>レ</hodoku><hodoku>バ</hodoku>、少臣奏シテ云ハク、</kanbun></s><s><kanbun>「左方<ruby rubyText="之">ノ</ruby>申ス所、謂ハレ無キニ非ズ。</kanbun></s><s><kanbun>此クノ如キ<fudoku>之</fudoku>事、只時<ruby rubyText="之">ノ</ruby>議ニ隨ハム。</kanbun></s><s><kanbun>但シ、人<ruby rubyText="之">ノ</ruby>誤リニ依ツテ何ゾ其ノ歌ヲ惡マム」</kanbun></s><s><kanbun>ト。

図5 「天徳四年内裏歌合」和化漢文部分の形式化例

5. 形態素解析例

図6に平安時代初期の歌合である「寛平御時后宮歌合」の形態素解析例を示す。題・左右・詠み人・歌・歌番号といった情報を「日本語歴史コーパス 平安時代編」の仕様に合わせ、本文種別や話者として持たせた。

6. おわりに

ここまで、歌合テキストの構造化の際の課題を事例に基づき検討し、形式化の例を示した。

歌合は特有の階層的構造を持ち、それが個々の要素の配置が一定しないという課題があった。歴史的資料を対象としたコーパスの構築にあたってはどこまで情報を与えるかということが重要であるが、本研究では歌合特有の構造に対し、可能な限り細かく情報を付与し、その情報を形態素解析結果に対応させて持たせることとした。また、歌合は一つの要素の内部に和文と和化漢文が混在するという特色を持っており、この扱いが大きな課題となっていた。

本研究で対象とした中古中世の歌合コーパスにおいて、ジャンル特有の要素の記述方法を確立し、文中に異なる文体を持つ歌合を対象とした記述を試行することによって、今後、和漢混淆文をはじめとした歴史的な資料を対象としたコーパスの仕様検討のための一材料となるものと考えている。

参考文献

市村太郎・河瀬彰宏・小木曾智信 (2012), 「近世口語テキストの構造化とその課題」, 『情報処理学会研究報告. 人文科学とコンピュータ研究会報告 2012-CH-96(1)』, 1-8.
 市村太郎・河瀬彰宏・小木曾智信 (2013), 「洒落本コーパスの構造化 —仕様と事例の検討—」, 『第3回 コーパス日本語学ワークショップ』, 249-258.
 小沢正夫・松田成穂 (1994), 『古今和歌集』, 新編日本文学全集 11, 小学館
 久保田淳・山口明穂校注 (1998), 『六百番歌合』, 新日本古典文学大系 38, 岩波書店
 小林正行・市村太郎 (2013), 「『虎明本狂言集』コーパスの構造化 —仕様と事例の検討—」, 『第3回 コーパス日本語学ワークショップ』, 323-332.
 萩谷朴・谷山茂校注 (1965), 『歌合集』, 日本古典文学大系 74, 岩波書店

関連 URL

大系本文データベース <http://base3.nijl.ac.jp/>
 中古和文 UniDic <http://www2.ninjal.ac.jp/lrc/index.php?UniDic%2F%C3%E6%B8%C5%CF%C2%CA%B8UniDic>

付記 本研究は JSPS 科研費 25770179 の助成を受けたものである。

キー	語彙素読み	語彙素	出現発音形	品詞	解析活用型	活用形	本文種別	話者
春	ハル	春	ハル	名詞-普通名詞-副詞可能			題	
歌	ウタ	歌	ウタ	名詞-普通名詞-一般			題	
廿	ニジュウ	二十	ニジュウ	名詞-数詞			題	
番	バン	番	バン	名詞-普通名詞-助数詞可能			題	
左	ヒダリ	左	ヒダリ	名詞-普通名詞-一般			左右	
紀	キ	キ	キ	名詞-固有名詞-人名-姓			詠み人	詠み人
友則	トモノリ	トモノリ	トモノリ	名詞-固有名詞-人名-名			詠み人	詠み人
花	ハナ	花	ハナ	名詞-普通名詞-一般			歌	歌番号(00001)
の	ノ	の	ノ	助詞-格助詞			歌	歌番号(00001)
香	カ	香	カ	名詞-普通名詞-一般			歌	歌番号(00001)
を	ヲ	を	ヲ	助詞-格助詞			歌	歌番号(00001)
風	カゼ	風	カゼ	名詞-普通名詞-一般			歌	歌番号(00001)
の	ノ	の	ノ	助詞-格助詞			歌	歌番号(00001)
たより	タヨリ	便り	タヨリ	名詞-普通名詞-一般			歌	歌番号(00001)
に	ニ	に	ニ	助詞-格助詞			歌	歌番号(00001)
たぐへ	タグエル	類える	タグエ	動詞-一般	文語下二段-八行	連用形-一般	歌	歌番号(00001)
て	テ	て	テ	助詞-接続助詞			歌	歌番号(00001)
ぞ	ゾ	ぞ	ゾ	助詞-係助詞			歌	歌番号(00001)
鶯	ウグイス	鶯	ウグイス	名詞-普通名詞-一般			歌	歌番号(00001)
さそふ	サソウ	誘う	サソウ	動詞-一般	文語四段-八行	連体形-一般	歌	歌番号(00001)
しるべ	シルベ	導べ	シルベ	名詞-普通名詞-一般			歌	歌番号(00001)
に	ニ	に	ニ	助詞-格助詞			歌	歌番号(00001)
は	ハ	は	ハ	助詞-係助詞			歌	歌番号(00001)
やる	ヤル	還る	ヤル	動詞-非自立可能	文語四段-ラ行	連体形-一般	歌	歌番号(00001)

図6 形態素解析例