

# 『源氏物語』における「紫上系」と「玉鬘系」における 語の使用傾向について

## Quantitative Analysis of the Word Frequency Difference Between “Murasaki no Ue Group” and “Tamakazura Group” in “The Tale of Genji”

土山 玄

Gen Tsuchiyama

同志社大学 研究開発推進機構, 京田辺市多々羅都谷 1-3

Doshisha University, 1-3 Tatara Miyakodani, Kyotanabe, Kyoto

あらまし:『源氏物語』の主人公である光源氏の出生から栄華を極めるまでが描かれた第一部と称される33巻は「紫上系」と「玉鬘系」という2つのグループが内在している可能性が従来から指摘されている。このような問題を扱うとき、従来は、記述内容の検討や成立に関する歴史的事実の考証という観点から研究をおこなうのが主であったが、本研究では計量文献学の方法を用いる。計量文献学は文章から得られる計量的なデータを収集し、これを分析することによって当該文献の文体的特徴を把握して、結論を導き出すものである。

そこで、本研究では第一部における「紫上系」及び「玉鬘系」と称される2つのグループを採り上げ、計量的な分析を行った。分析の結果、「紫上系」と「玉鬘系」との間において、助動詞の語の使用傾向に相違が認められた。

**Summary:** Some researchers of Japanese Literature believe that the current order of chapters of “The Tale of Genji” is different to the original order. The first part of the story is divided two groups: the *Murasaki no Ue Group* and *Tamakazura Group*. Herein, we investigate the difference in style regarding the word frequency between the two groups using principal components analysis and the chi-square test. In this study, we will report the feature words between the two groups.

キーワード: 源氏物語, 計量文献学, 多変量解析, 紫上系, 玉鬘系

**Keywords:** The Tale of Genji, stylometry, multivariable analysis, Murasaki no Ue Group, Tamakazura Group

### 1. はじめに

『源氏物語』とは、平安時代の女流作家である紫式部が著した『紫式部日記』の記述から、紫式部の手により執筆されたと考えられる現存最古の長編物語の1つである。『源氏物語』の研究の歴史は長く、平安時代にまで遡るが、江戸時代の本居宣長は著名であり、著書の『源氏物語玉の小櫛』において語句や年立などに新説を提示し、従来の研究における誤りを正すとともに源氏物語の基調をなす中心理念として「もののあはれ」を論じていることはあまりに有名である。

この『源氏物語玉の小櫛』において、本居宣長は『源氏物語』の成立過程についても言及しており、第2巻「帚木」の冒頭は第1巻「桐壺」からの連続に違和感が

あることを指摘している。これは「帚木」の冒頭には「いひ消たれ給ふとが多かんなるに」や「かかる好きごと」とあるが、これらが指す内容が直前の「桐壺」に叙述されておらず、「帚木」以降に描かれている内容を指しているとするものである。また、同様に「帚木」の冒頭にある「語り伝えむ」についても後述のことを指すと本居宣長は論じている。

また、和辻哲郎が本居宣長の研究を受けて『源氏物語』の論を発展させている。和辻哲郎の論ずるところによれば、「桐壺」における光君は母に酷似した継母藤壺に思慕の情を抱いており、葵の上を好きになれなかったと描かれているが、「帚木」では冒頭から光源氏は好色人として描いており、ここに違和感が残ると問題提

起している。また、「帚木」の冒頭に本居宣長が指摘した前述のような記述があるのは、「帚木」執筆時の読者が「桐壺」に描かれていない光源氏についての何らかの知識を有していることが前提とならなくてはならないと論じている。

他にも青柳（1939）も源氏物語の成立過程について研究をしており、第5巻「若紫」を中心に「桐壺」から第23巻「初音」までの23巻の執筆順序について考察している。これは「若紫」を書き出しの帖とし、第7巻「紅葉賀」、第8巻「花宴」、第9巻「葵」、第10巻「賢木」、第11巻「花散里」、第12巻「須磨」の前半を書いた後、一度前へ戻り「帚木」、第3巻「空蟬」、第4巻「夕顔」、第6巻「末摘花」を書き、次に「須磨」の後半から続けて、第21巻「少女」まで10巻の執筆を行い、再び前に戻って「桐壺」を書いてから第22巻「玉鬘」、第23巻「初音」と物語を進めたと論じている。またこれら23巻を若紫グループと帚木グループという2つに分け、若紫グループに登場する人物は帚木グループにも登場するが帚木グループに登場する中心的な人物は誰一人として若紫グループには登場しないということを指摘している。

このような成立過程の研究、すなわち成立論に関する確固たる見解が提出されたのは武田（1954）によるもので、その論によれば、第一部と称される33巻までが「紫上系」と「玉鬘系」という二つの系列に分離されるとしている。これら二つの系列は、はじめに「紫上系」に属する17巻が書かれ、その後「玉鬘系」に属する16巻が執筆され紫上系の中に挿み込まれたという論である。上述の青柳（1939）は帚木グループの主要な登場人物が若紫グループには登場しないことを指摘したが、武田はこれを受けて、初出が「紫上系」となる登場人物は「玉鬘系」においても登場するが、初出が「玉鬘系」である人物は「紫上系」に登場しないということを指摘している。このことにより、「紫上系」が先行して構想および記述され、「玉鬘系」が後に記述挿入されたとされる。くわえて、「紫上系」と「玉鬘系」との間には文体や技巧などといった点においても相違も見出せると論じている。

また、大野（1984）によれば「紫上系」は本紀の形式を取り入れていること、一方、「玉鬘系」は列伝の形式を取り入れていることが指摘されている。本紀とは年月を追う記述の形式であり、列伝とは個人の行動などの重要な点の記述に主眼をおいた形式である。このようなことから、「紫上系」と「玉鬘系」の文章の間には質的な相違が認められると言える。

なお、一般に源氏物語は三部構成であると考えられており、第二部は第34巻から第44巻までの8巻であ

り、第三部は匂宮三帖及び宇治十帖と称される13巻であり、「紫上系」と「玉鬘系」にそれぞれ属する巻は表1に示す通りである。

表1 「紫上系」及び「玉鬘系」

紫上系		玉鬘系	
01 桐壺	13 明石	02 帚木	24 胡蝶
05 若紫	14 滯標	03 空蟬	25 蛍
07 紅葉賀	17 絵合	04 夕顔	26 常夏
08 花宴	18 松風	06 末摘花	27 篝火
09 葵	19 薄雲	15 蓬生	28 野分
10 賢木	20 朝顔	16 関屋	29 行幸
11 花散里	21 少女	22 玉鬘	30 藤袴
12 須磨	32 梅枝	23 初音	31 真木柱
	33 藤裏葉		

このように、『源氏物語』の成立過程に関するが、論説はすでに報告されているが、客観的なデータは主要な登場人物の出現に関する情報のみである。そこで、本研究では従来の人文学における手法ではなく、計量的な観点から『源氏物語』の第一部の量的な特徴に検討を加えたい。

文章を対象とし、統計的な分析手法を行い、文献に考察を加える学問の1つに計量文献学がある。計量文献学は、著者の文体に関わる習慣的特徴を解析することを通じて、著者の識別、文献の成立年代、あるいは成立の順序を推定する学問分野である。

本研究において取り扱う『源氏物語』をはじめ、古典文学などの歴史的な文献には、著者や成立時期について議論の対象となっている場合が少なからず存在する。このような問題を扱うとき、従来は、記述内容の検討や成立に関する歴史的事実の考証という観点から研究をおこなうのが主たる方法であったが、計量文献学はこのような方法とは一線を画し、文章から得られる計量的なデータを収集し、これを分析することによって当該文献の文体的特徴を把握して、結論を導き出すのである。

このように文体を取り扱うこと背景には、文体に著者の個性が現れるという理解があり、これは経験的に首肯しうるものと考えられる。ただし、文体という概念は多様であり、学問分野によってその内容は異なる。計量文献学においては、文体とは計数可能な記述形式、すなわち、文章を構成する量的な要素であり、語の意

味や記述内容は考慮しない。この点において計量文献学は人文的文献研究とは立場を異にしている。

本研究では『源氏物語』の第一部の33巻における語の頻度を採り上げ、これに統計的な分析を加えることで、第一部の諸巻の量的な特徴を明らかにする。次いで、「紫上系」と「玉鬘系」の間における量的な関係性を可視化し、どのような語の出現傾向に顕著な相違が認められるか検討する。

## 2. 関連研究

著者の識別を目的とした文章の計量的な研究は広く報告されているが、それに比べて文献の成立過程に関する研究事例は多くない。しかし、研究の歴史は長く、L. Campbell は1867年にプラトンの30余りある『対話篇』の執筆年代を推定するために、語の出現頻度を用いて計量的な分析を行っている。

プラトンの執筆時期は50年から60年に及ぶと考えられているが、『法律』がプラトンの最後の著作であることをアリストテレスが言及していることを除き、『対話篇』の執筆順序は不明であった。プラトンの思想には発展がみられることから、プラトンの思想を体系的に理解するためには、著作の執筆順序を明らかにすることがきわめて重要であった。Campbell (1867) では、プラトンの後期の『対話篇』とされる『ティマイオス』、『クリティアス』、『法律』に共通に用いられていながら、他の『対話篇』においてはほとんど用いられない語彙を調査し、これらの語彙の生起状況の調査することで、『ティマイオス』、『クリティアス』、『法律』に加え、『ソピステス』、『政治家』、『ピレポス』が後期の『対話篇』もまた後期の執筆であると指摘した。

日本における研究事例としては、金 (2009) があげられる。金 (2009) では芥川龍之介の著作を対象に執筆年代の推定を行った実証的な研究であり、助詞の出現率の経時的な変化を明らかにすることにより、高い精度での執筆年代の推定が可能であることを指摘している。

『源氏物語』を対象とした計量的な研究事例としては、最終10巻である宇治十帖において論じられている他作者説を検討するものが多い。安本 (1957) においては平安時代に紫式部 (973-1014) によって著されたとされる『源氏物語』を分析対象としている。『源氏物語』を宇治十帖とその他の44巻に二分し、統計的検定を行っている。検定に用いた項目は名詞の使用度、用言の使用度、助詞の使用度、助動詞の使用度など12あり、これらの使用度は、各巻からランダムに1000字抽出し、その1000字における各項目の頻度によって

求められている。従って、『源氏物語』の全文が分析の対象になっているのではない。

検定の結果、宇治十帖の文体は作り物語、用言的、緊密かつ連続的な構想による詳細な描写を特徴とし、一方、他の44巻の文体は歌物語、体言的、飛躍的、断続的な構成による直感的描写を特徴とすると考察されている。それゆえ、宇治十帖の作者は他44巻の作者と同一人物であるとは言い難い、と結論づけている。

他方、新井 (1997) は、各巻の中央部から各巻の長さに応じて標本を抽出し、五十音図の頭子音行別頻度や母音列別頻度に対して統計的検定を行っている。検定の結果、宇治十帖の作者が他の諸巻の作者と別人であるとは考えられないと述べている。

上掲の『源氏物語』を対象とした2つの計量的な研究はそれぞれに意義を有するが、物語の全文を用いて分析をおこなってはいない。村上・今西 (1999) では『源氏物語』の全文を対象とし、多変量解析の手法を用いて本格的な計量的研究を行った。この研究では助動詞の出現率を用い、数量化Ⅲ類により『源氏物語』の成立巻序の推定を行っている。分析の結果、紫上系、第二部、玉鬘系、第三部の順序で成立した可能性を報告している。

## 3. 分析

### 3.1 データ

源氏物語の写本系統は青表紙本系、河内本系、別本と3系統に大別される。本研究では、青表紙本系の大島本を主な底本とする『源氏物語語彙用例総索引 自立語編』及び『源氏物語語彙用例総索引 付属語編』を電子化したデータベースを分析に利用した。『源氏物語語彙用例総索引』は源氏物語の本文すべてについて、形態素解析を行ったものである。すべての語に、表記形、表記形仮名読み、終止形、終止形仮名読み、品詞コード、活用形コード、意味コードなどの情報が付与されている。なお、単語認定については、『源氏物語大成 索引篇』の単語認定基準に準拠している。

### 3.2 方法

先にふれたように、本研究において採り上げる分析項目は語の頻度である。語の頻度とは各語の対象における出現頻度である。本研究における分析では頻度ではなく分析対象となる第一部33巻の各巻における品詞別の延べ語数に対する割合を求め、これについて分析を行った。また、語の頻度を集計に際して、全語彙を一括して集計するのではなく、品詞別に集計

した。本研究においては、名詞・代名詞・動詞・補助動詞・形容詞・形容動詞・副詞・連体詞・助詞・助動詞について、語の集計を行った。

分析においては、相関係数行列を用いた主成分分析を行った。主成分分析とは、多次元データに対する次元縮約の手法であり、もとのデータの変数より新たに合成変数を求めることで、情報の縮約を行う。分析結果は、主成分分析によって求められた分析対象の第1主成分と第2主成分の主成分得点の散布図によって表現される。散布図においては第1主成分を横軸、第2主成分を縦軸とし散布図を描いた。また、各主成分に含まれている情報は寄与率によって評価され、本章においては第1主成分及び第2主成分の寄与率は散布図に併記した。

次いで、「紫上系」と「玉鬘系」との間において顕著な出現傾向の相違が認められる語、すなわち特徴語の抽出にはカイ二乗検定を用いた。カイ二乗検定とは2つの項目の間における独立性の検定に使用される手法であり、2つの項目間に関連があるのであれば、p値が小さくなる。特徴語の抽出に際しては、第一部の33巻において出現するすべての語を「紫上系」と「玉鬘系」ごとに集計し、ある語がその語を除くすべての語の出現傾向と同様であるのか検定をし、本研究ではp値の低い語を特徴語とし、抽出した。

### 3.3 分析結果

分析においては、まず語の頻度について上述の品詞別に主成分分析を行った。本稿では分析結果の一部を示す。大野(1984)において指摘されているように、「紫上系」と「玉鬘系」では形式が異なることが報告されているが、図1に示すように、名詞について主成分分析からは語の出現傾向の顕著な相違は認められない。図1は名詞の出現頻度上位50語についての分析結果である。黒字で表記されている巻が「紫上系」であり、灰色で表記されている巻が「玉鬘系」である。また、図1における、実線の楕円と破線の楕円は各巻の第1主成分及び第2主成分の主成分得点から推定される95%信頼楕円であり、実線が「紫上系」、破線が「玉鬘系」の信頼楕円である。なお、第1主成分の寄与率は32.5%、第2主成分の寄与率は10.0%となる。

名詞の他に代名詞・動詞・補助動詞・形容詞・形容動詞・副詞・連体詞においても、図1と同様に「紫上系」と「玉鬘系」との間において、顕著な語の出現傾向の相違は認められなかった。従って、「紫上系」と「玉鬘系」の間において、名詞・動詞、形容詞・形容動詞・副詞などの修飾語の出現傾向に大きな相違はないと考えられる。

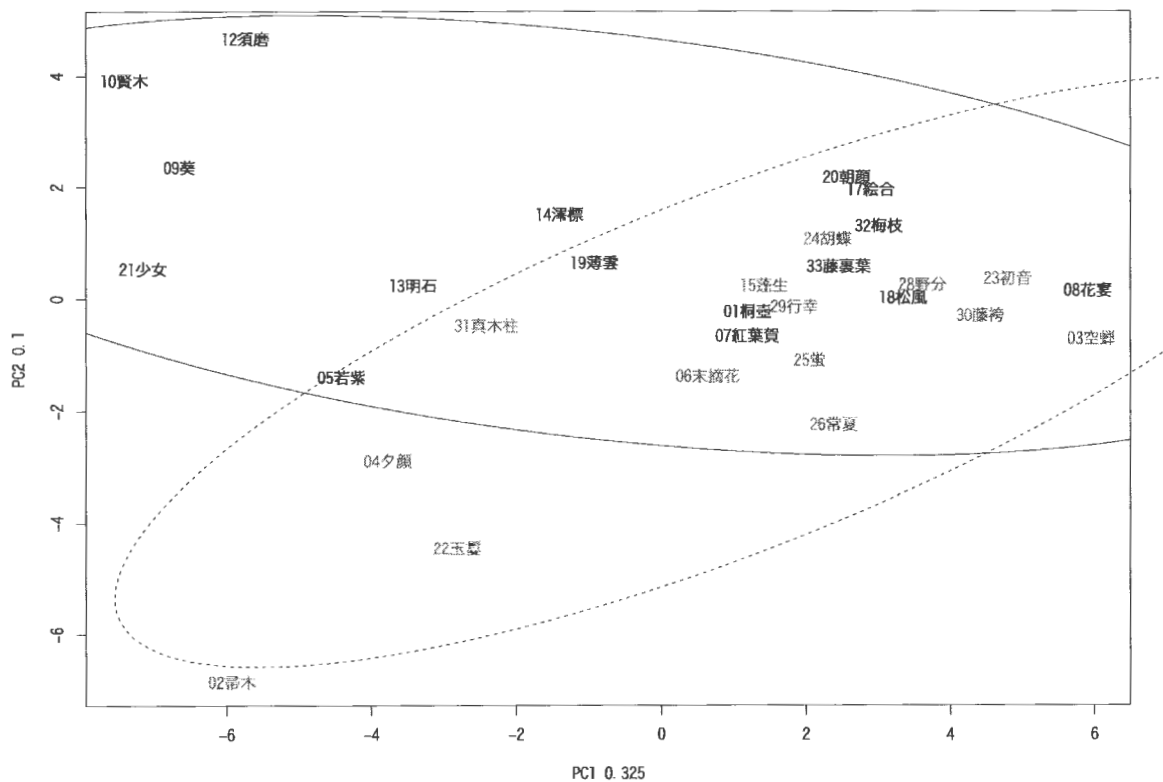


図1 名詞出現頻度上位50語に対する主成分分析の結果

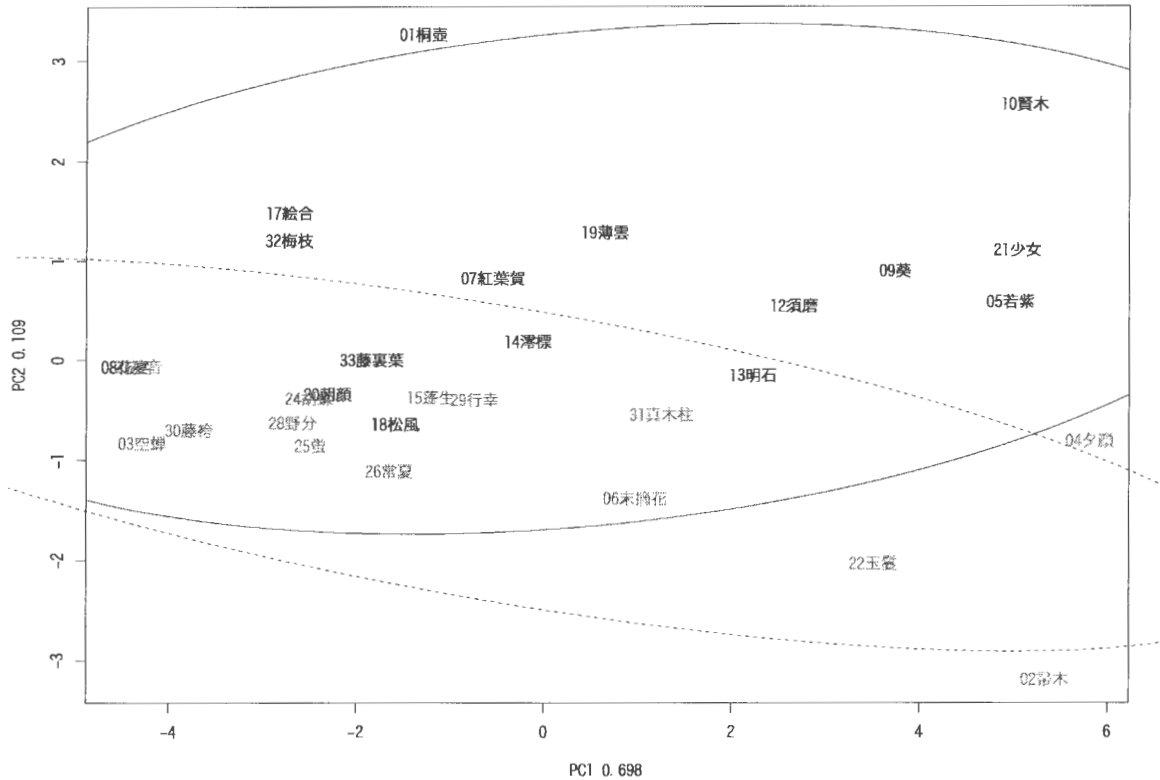


図2 助詞 15 語に対する主成分分析の結果

しかし、助動詞において、「紫上系」と「玉鬘系」の間に語の出現傾向の相違が認められた。図2は助動詞の出現頻度上位15語の主成分分析の結果である。図1と同様に「紫上系」と「玉鬘系」の95%信頼楕円は重複するが、第一部の33巻が混在する図1とは異なり図2においては、縦軸の第2主成分の正の領域に「紫上系」が、負の領域に「玉鬘系」が位置する傾向が認められる。なお、第1主成分の寄与率は69.8%、第2主成分の寄与率は10.9%である。図2において認められた傾向は変数を増加させても大きくは変化しない。

表2 紫上系における助動詞の特徴語

	p 値	紫上系	玉鬘系
ス	0.0000	0.0387	0.0186
サス	0.0000	0.0207	0.0105
リ	0.0000	0.0964	0.0742
ル	0.0000	0.0426	0.0284
ヌ	0.0028	0.0795	0.0681
キ	0.3237	0.0633	0.0598
ラム	0.4430	0.0156	0.0141
マシ	0.5521	0.0096	0.0087
ツ	0.6154	0.0325	0.0311
ゴトシ	0.7104	0.0008	0.0006

図3は助動詞の全語彙を用いて行った主成分分析の結果である。図3においては、「紫上系」と「玉鬘系」の諸巻が図2に比べ混在していると言えるが、図2と同様に、第2主成分の正の領域に「紫上系」が、負の領域に「玉鬘系」が位置する傾向が認められると考えられる。

表3 玉鬘系における助動詞の特徴語

	p 値	紫上系	玉鬘系
ム	0.0000	0.0940	0.1236
タリ	0.0000	0.0946	0.1160
マジ	0.0020	0.0086	0.0132
ケリ	0.0216	0.0774	0.0865
メリ	0.0414	0.0190	0.0233
ベシ	0.0558	0.0622	0.0691
ジ	0.0745	0.0096	0.0124
ラル	0.1487	0.0140	0.0167
ズ	0.4297	0.1244	0.1283
ナリ	0.7924	0.0815	0.0827

次に、助動詞について特徴語の抽出を行った。用いた手法はカイ二乗検定である。表2、表3は「紫上系」及び「玉鬘系」における特徴語であり、p値の昇順に整理した10語である。分析の結果、「す」や「さす」といっ

た使役・尊敬の助動詞は「紫上系」に多く出現すると言  
える。また、動詞の未然形に接続する、受身・尊敬・可  
能・自発の助動詞である「る」は「紫上系」に、「らる」は  
「玉鬘系」に多く出現すると言える。

#### 4. 結びにかえて

本研究では『源氏物語』における第一部33巻の語の  
頻度を採り上げ、これに統計的な分析を加えることで、  
第一部の諸巻の量的な特徴について検討した。分析  
の結果、助動詞において「紫上系」と「玉鬘系」の間に  
語の出現傾向の相違が認められると考えられる。次い  
で、「紫上系」と「玉鬘系」の間における助動詞の特徴  
語を抽出した。その結果、使役・尊敬の助動詞、受身・  
尊敬・可能・自発の助動詞の出現頻度が、「紫上系」と  
「玉鬘系」との間において、特徴的に相違することが明  
らかになった。

#### 参考文献

池田亀鑑(1987). 『合本源氏物語事典』, 東京堂出  
版.  
和辻哲郎(1970). 『日本精神史研究』, 岩波書店.  
青柳秋生. (1939). 源氏物語執筆の順序—若紫の巻  
前後の諸帖に就いて—. 『日本文学研究資料叢  
書 源氏物語Ⅲ』, 有精堂, 32-52.

武田宗俊(1954). 『源氏物語の研究』, 岩波書店.  
大野晋(1984). 『源氏物語』, 岩波書店.  
Campbell, L. (1867). *The Sophistes and Politicus of  
Plato*. Clarendon Press.  
金明哲. (2009). 文章の執筆時期の推定—芥川龍之  
介の作品を例として—. 行動計量学, 36(2),  
89-103.  
安本美典. (1960). 『文章心理学の新領域』, 創元社.  
新井皓士. (1997). 源氏物語・宇治十帖の作者問題:  
一つの計量言語学的アプローチ. 一橋論叢,  
117(3), 397-413.  
池田亀鑑. (1985). 『源氏物語大成 索引篇』, 中央公  
論社.  
村上征勝・今西祐一郎. (1999). 源氏物語の助動詞の  
計量分析. 『情報処理学会論文誌』, 40(3), 33-38.  
上田英代・村上征勝・今西祐一郎・樺島忠夫・上田裕  
一. (1994). 『源氏物語語彙用例総索引—自立語  
編—』, 勉誠出版.  
上田英代・村上征勝・今西祐一郎・樺島忠夫・上田裕  
一. (1996). 『源氏物語語彙用例総索引—付属語  
編—』, 勉誠出版.

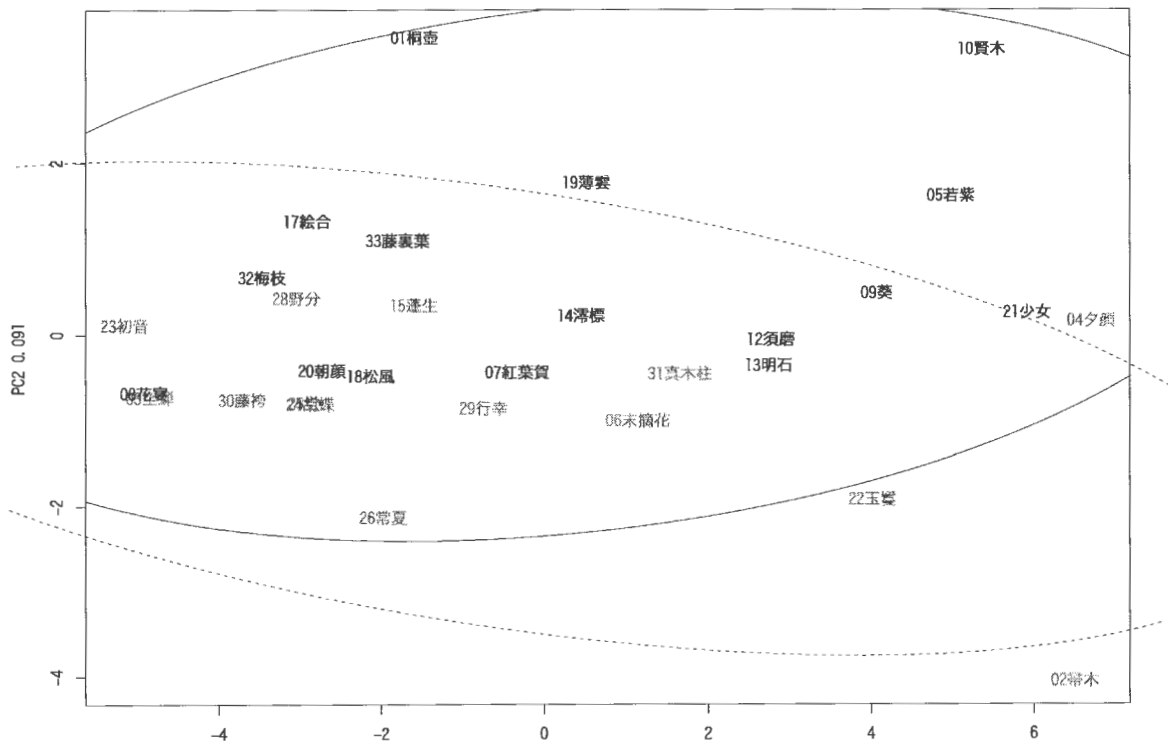


図3 助詞全語彙に対する主成分分析の結果