

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・一 問題の所在(宇治十帖作者別人説)

十一世紀初頭に成立した『源氏物語』は、とかく独創性に欠けると評されがちな日本学芸史にあって、掛け値なしに世界文化史的意義を認められている言語芸術作品である。全巻五四帖、総語数およそ四十万語とされる大河小説的規模、「若菜」や「夕霧」の巻などの芳醇な浪漫性、「若菜」の巻以下に顕著な近代小説的心理描写、全巻を貫く美意識と文体(臚写体)にみられる品格の高さ、そしてこの時代にして世界に類例のない女性の手になる秀作であるという社会史的意義等、そのいずれをとっても伝紫式部の「源氏」は日本文化の真に誇るべき遺産であることに疑問の余地はない。

新井 皓 士

だが、この作品の成立過程、構成原理、作者「紫式部」の生涯、その他に関して、古来謎とされ論議の絶えない事柄もまた少なくない。「極端なことをいえば、源氏物語の本文そのものにしても、(一)断定することはできない状態にある。」第四五帖以下の通称「宇治十帖」の作者問題もまたその一つである。

『源氏物語』は周知のように、光源氏を主人公とする正編(全体の約五分の四)と、その没後の展開、すなわち匂宮(光源氏の孫)と薫大将(光源氏のいわば戸籍上の子)を中心とする続編、合わせて五四帖(巻)の形で現在に伝わっている。しかし、散逸を疑われる巻(輝く日の宮)や表題のみ伝えられる巻(雲隠れ)の問題は別にしても、現存する五四帖の正編と続編の接合部(第四

二、四三、四四帖) および統編それ自体について、作者別人説あるいは後人補筆説が古来絶えない。このうち統編「宇治十帖」については紫式部の娘「大式三位」を作者とする説(十五世紀後葉の一条兼良『花鳥有情』にみえる)があり、接合部の特に第四四帖「竹河」については、斎院宣旨と略称される女性の作とする伝承(十六世紀初頭の花屋『玉栄集』にみえる)のみならず、二十世紀後半に入って有力な作者別人説が浮上し(武田宗俊、池田龜鑑、石川徹)、いまだ完全な決着をみないという。

これら諸説は方法論からみれば、主として物語の内容上の矛盾や文体の差異を伝統的な文献学ないし文学研究の視点から分析し解釈立論したものと見えよう。これに対して統計学的手法をこの問題に応用した先駆的試みに、安本美典の「文体統計による筆者推定」(一九五八)、⁽²⁾「現代の文体研究」(一九七七)がある。すなわち金子元臣校訂『定本源氏物語』を用いて、各巻(帖)の頁数、和歌の使用度、文の長短、直喩・声喩・色彩語の使用度、心理動詞による終結文の使用度、品詞別単語頻度を計測し、ノンパラメトリック検定を行った結果が前者であり、おそらく同じデータに単語長データを加えて因子分析法

を適用した結果が後者である。前者では断定こそ避けられてはいるが、「宇治十帖別人作者」の可能性が示唆され、後者では「このデータから少なくとも、宇治十帖は、他の四十四帖と、文体がやや異なっていることはいえるであろう。」という、より慎重な結論が提示されている。

一・二 『源氏物語』のテキスト

『源氏物語』には成立時の原本というべきものがなく、約二百年後に校訂された河内本、青表紙本と呼ばれる二系統の写本によって本文が伝えられている。前者は(河内守)源光行・親行父子が整理した混態本で声点(濁点)や句読点なども補なわれ、藤原定家の整理になる後者はいわゆる定家仮名遣いが準用され句読点を有さない。後者の名称は表紙の色に由来すると言われるが、定家自筆は四帖しか残存せず、十五世紀末と思われる三条西実隆奥書の「青表紙証本」(岩波古典文学大系の底本)および飛鳥井雅康筆写本(大島本、小学館日本古典文学全集の底本)が、この系統を伝えている。本研究に着手した時点において偶々滞独中であつた筆者は、はじめ飛鳥井本(及び明融本)に準拠した玉上琢磨校訂の角川文庫

本により機械可読テキスト・データを作成したが、その後「日本古典文学全集」を参照して若干の修正を施した。修正の主要点は助動詞などの撥音表記および「御」の字の扱いに関わるものである。なお仮名遣いは（定家仮名遣いではなく）いわゆる歴史的仮名遣いが基準になっている。

ところで、表音文字（アルファベット）を用いる言語が計量言語学の対象となる場合、計測単位として文字（頻度分布、連接関係、エントロピーなど）以上に着目されるのは単語であり、その出現頻度や語彙分布などの記述統計学および推測統計学的検証には長い伝統と蓄積がある。これはいうまでもなく、単語の切れ目をスペースで示す表記上の特徴が計測を容易にし単純な機械処理をも可能にするからである。しかし膠着語であり漢字仮名まじりを表記特徴とする日本語では、単語に関してこのような形態上の目安がなく、またあえてローマ字表記をとってこの問題を回避しようとするれば、そもそも単語および文節そのものの定義問題が生ずる。この事情は文の長さを計測する場合にも波及し、単語数により文章長を量る印欧語的手法を安易に採用することはできないし、

漢字仮名まじり文では当用漢字の使用等テキスト編集方針が影響するゆえ、これを単に文字数で量ることは余程慎重でなければならぬ。

印刷文化のおかげで一見安定したテキストが供給される現代においてもこのような事情であるとすれば、写本時代のテキストに関しては一層慎重でなければならぬことは言うまでもない。しかし、こと平安朝の和文については文字表記上むしろ計量言語学にとって有利ともいえる歴史的背景がある。すなわち女手とも呼ばれた「かな」の発達がそれであり、現代の校訂本では読解の便のために相当数の漢字が適用されている『源氏』をはじめ、物語文学や和歌は専ら「かな」によって表記されていた。これはとりもなおさず、テキストをかな文字表記に戻せば、むしろ元の姿に近くなることを意味している。むしろ当時であっても固有名詞や一定数の漢字³⁾あるいは変体がなの使用があったものと想定されるが、物語にせよ和歌にせよ、もともと単に目で読むというより音読が前提にあったと考えられるから、音節文字である「かな」の文に還元することはむしろその本質にかなうといえよう。文字情報に頼りがちなフィロロジィーに対して音として

のことにば意識的に立脚するのが二十世紀の言語学(リ
ンギステイクス)だと言われ、アルファベット系言語
の場合もたとえば単語の長さを量るのに音素・文字(書
記素)・音節といった選択肢が提起されうるが、「かな」
はある意味ではこれら諸要素の接点をなすものでもある。

以上のような理由から、機械可読テキストデータ作成
にあたって筆者は、少なくとも日本語古文に関しては専
ら平仮名を基本としているが、今回『源氏』テキストの
標本を抽出する際に更に次のような諸点に留意した。ま
ず、文の切れ目の解釈は底本に従う。ただし和歌につい
ては、独立して一文の体をなす場合は文章長の調査では
カットし、地の文や会話文中で前後と一体化している場
合は残すこととした。このため一般には全角文字で入力
したが和歌のみは半角カタカナ入力とし容易に区別でき
るようにしておき、また和歌のみを採録したデータ(フ
ァイル)を別途用意し、他との比較に供することとした。
会話(引用)部分については試行錯誤の末に、現代風の
引用符を省き、地の文と切れ目なく続く場合は会話部の
最初の句点を文の切れ目とみなす原則をとっている。次
に人名ないし固有名詞に相当するものは底本の漢字表記

を踏襲し、他は原則としてひらがな表記とする。例外は
助動詞「む、なむ、けむ、らむ」でこれらに含まれる
「む」を「ム」とし、その撥音化した「ん」と同様に別
枠扱いとした。それはこれらが写本によって「南、剣、
乱」と書かれていることもあり、『源氏』成立期には少
なくとも文字としての「ん」は確立していなかったと思
われるからであり、写本によって「む」と「ん」の使用
状況はかなりまちまちとみられるからである。同様の理
由で読み方(ルビ)に「おほん、おほむ」その他の可能
性がある「御」の字については、明確に「み」とすべき
場合以外は漢字をそのまま残すことにした。

二・一 「かな」とやまとことば

音節文字「かな」の分布や「かな」で表された文章の
長さの分布は、文体の特徴ないし相違をどの程度反映す
るか、これらの分布から得られる統計量に基づき『源
氏』正編と続編について一様性検定を試みた場合、どの
ような結論が得られるか、これが本研究の出発点となっ
た着想であるが、それは単に欧文の場合の単語単位の計
測を文字単位におきかえるだけでなく、日本語の特徴

と歴史的発展を生かすものでもあることに簡単にふれておこう。

古代の日本には独自の文字がなく、中国文明と接触する過程で「漢字」を受容し、「外国語」としての漢文を習得利用すると同時に、その字音と字訓を「交え用ゐる」ことよって日本語を記録したこと、奈良時代の『古事記』『日本書紀』『万葉集』がその代表的成果であることは周知の事実である。そしてこれらにおいて日本語の音を表すため用いられている漢字（いわゆる万葉仮名）を詳細に分析すると、「上代特殊仮名遣い」と総称される法則性が認められ、これが当時のやまとことばの音韻体系を精確に反映するものであること、すなわち「一字一音に用いられた万葉がなの数は、記紀・万葉を通じて九七三に達するが、（……）同一音に対して幾通りもの使用字があり、これらを類別すると八七に分けられる」ことが明らかになっている。これは現代日本語の清濁音節合計のみならず、平安時代後半に登場する「いろは」四七プラス濁音節の計六七より二〇（古事記の場合は二一）多いが、その差は当時の音韻体系では「エキケコソトノヒヘミメヨロ（モ）」とその濁音節に、甲類・乙類

と呼ばれる二種の区別があったからであり、それは母音の差から、換言すれば奈良時代のやまとことばには、イエオの列（段）が二種に分かれ母音が八種類あったことから生じている。しかしこの区別は平安時代になると消滅し、僅かに残ったア行のエ（衣）とヤ行のエ（江）の区別も、『源氏物語』が成立するころには失われていたことは、その少し前に流布したと考えられる「いろは」歌が示している。

音韻体系がこのように変化する一方で、それを固定する媒体としての文字の体系にも、平安時代には大きな発展があった。『日本書紀』では「字」を「な」と読ませているが、やがて「乎古止」点と並んで經典読解用の略字メモから「(カタ)カナ」が、また漢字草書体から「(ひら)かな」が発達すると、これと区別するため漢字はあらためて「真字」(まな)と呼ばれるようになる。そしてどちらかといえば仏教世界にとどまるカタカナに対し、「仮字」とされ「女手」ともされる(ひら)かな文字が、おそらく当初は一段低い文字として、しかし日本語の特性を柔軟に表現しうる文字として着実に普及し始めたのである。

このことを如実に示すのは十世紀の始めにおける『古今和歌集』の成立であろう。中国における正史は近代に至るまで継続し二四史(二五史?)を数えるが、それに對抗するわが六国史は十世紀始めの『三代実録』をもってその幕を閉じ、いわばその代替機能を担うものが勅撰和歌集「八代集」であるとすれば、そのはしりをなす『古今集』は、『経国集』などの先行する三つの勅撰漢詩集の跡を襲い、平仮名による最初の勅撰集として「仮字」の意義と役割を明らかにするものであった。遣唐使が廃止されて百年余り、平安京が開かれて二百年余りの歳月が経った十一世紀初頭の『源氏』は、変態漢文・和風漢文を含め男子ないし公の世がなお漢文脈にとらわれている時代において、当時はあくまでも私的な営みでありながら結局後世にこの時代を代表するものとなる仮名文化が凝縮したものであり、助詞・助動詞を重要な要素とする日本語の特性に対応した日本語固有の文字による表現が、あるいは和歌として、あるいは日記として、あるいは『竹取』『宇津保』『落窪』などの物語として生成発展した「かな文学」の集大成ともいえる作品である。

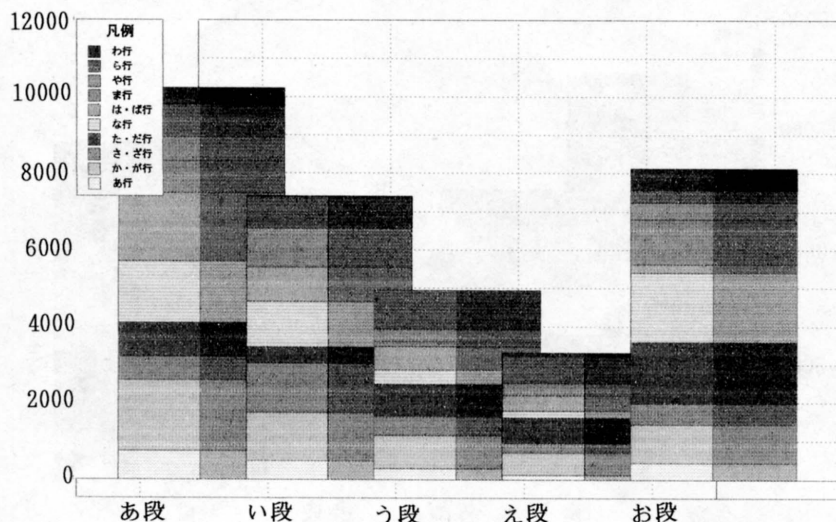
二・二 「かな」分布の代表例

それでは実際ひらがなで表した場合、日本語の音節にはどんな分布の傾向がみられるのか、皆曇(Siddhan)すなわち梵語研究から十一世紀に導入され今日も用いられる五十音図の形で文字を集計し、縦軸を行(あかさたな……)、横軸を列(段、あいうえお)として、積み重ねて表した幾つかの例を示そう。

最初は『古今和歌集』の文字(音節)分布で、たとえば一番左の積み重ね棒の高さはあ段(列)文字の総計を意味し、積み重ねの段落は下から順に「か行、さ行」となっている。ただし繁雑になるのを避ける意味もあって清音と濁音(たとえばか行とが行)は合算してある。グラフから明らかのように、母音列でみると「アオイウエ」という頻度順であり、頭子音行でみると「か行」「な行」「た行」が多く、「や行」や「わ行」は極めて少ない。

次は八代集の最後を飾る『新古今集』の文字分布をグラフ化したものである。紀貫之撰の『古今集』に対して、

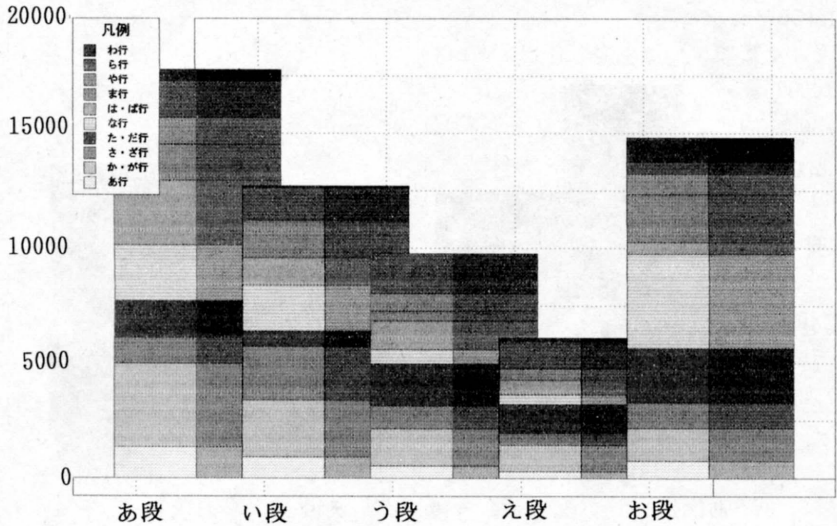
古今和歌集



藤原定家が撰者である『新古今集』は歌風において新境地を示すとされるが、母音列の頻度順位は「あおいうえ」で同じであり、頭子音行でみても、「か行」「な行」が多いなど、分布の傾向はかなり似ている。実は頭子音行の順位は、古今集では、カナタハラマサアヤワであり新古今では、カナタラハマサアヤワ、となって微妙に違うのであるが、その差は前川の提起している相互相関値をとってみても、ほとんど問題にならないほど小さい(一・三九)。なお特に助動詞に含まれる「む」は時代が進むとともに「ん」と表記される率が高くなるようであるが、名詞などに含まれるそれも含めてすべて合わせても和歌では一パーセント前後にすぎないようである。

それでは奈良時代に編纂された万葉集はどうであろうか。すでに述べたように奈良時代にはなお、アとウ以外の母音イ、エ、オには甲乙二種があり、母音体系としては五音ではなく八音であったと推定されるが、その具体例として大野普の『日本語の文法を考える』⁽⁷⁾に、一字一音の万葉仮名で表記されている巻五、一四、一五、一七、一八、一九、二〇の分析から作成された詳細な音節別使

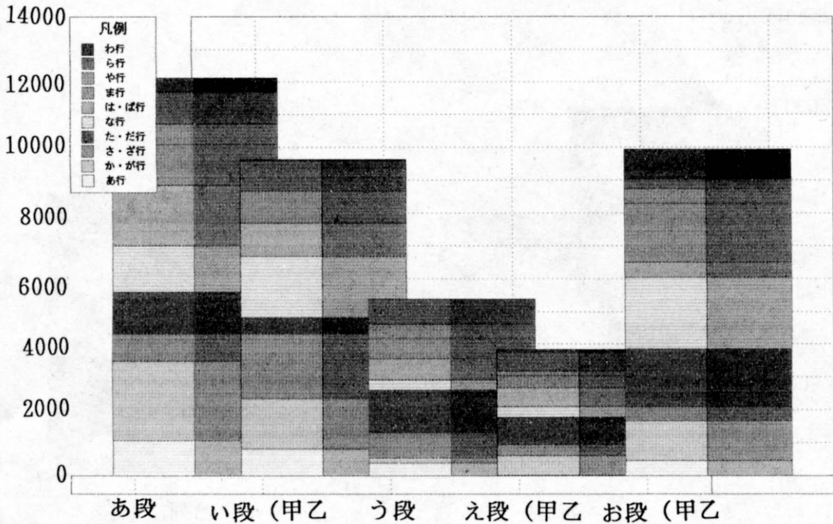
新古今和歌集



用度数表が掲載されているので、これを利用してみよう。まずこの表自体をながめると面白いことに、八母音がすべての頭子音について出現するのではなく、アヤワ行はやや特殊としても、サタナラの行はイエ段に乙がない(つまりシセチテニネリレに当たる音節は一種しかない)こと、ハマ行のオ段にも乙がない(ホモに当たる音節は一種しかない。ただし古事記では二種のモが書き分けられている)ことに気づく。濁音も同様で、これが先にあげたように、母音が五から八に増えても文字(音節)総数が単純に増加していない理由であり、なぜ頭子音によってこのような違いがあるのか、音韻学的検討が必要であろう。しかしここでは比較の為むしろ平安時代を基準に考えることにして、甲乙二種の差をあえて無視し二種ある音節は一種に合算し母音を五つに集約した形でグラフに描いてみる。

結果はやはり母音列の順位はアオイウエとなり、頭子音行の順位はカナタマハサラアヤワとなっている。古今・新古今ではハ・ラ行の下位にあったマ行がこれらより上位にきているのは万葉集ではモの使用度が高かったことがその主因かと思われ、またラ行が下位になるのは

万葉集の五十音分布 (甲乙類合算)

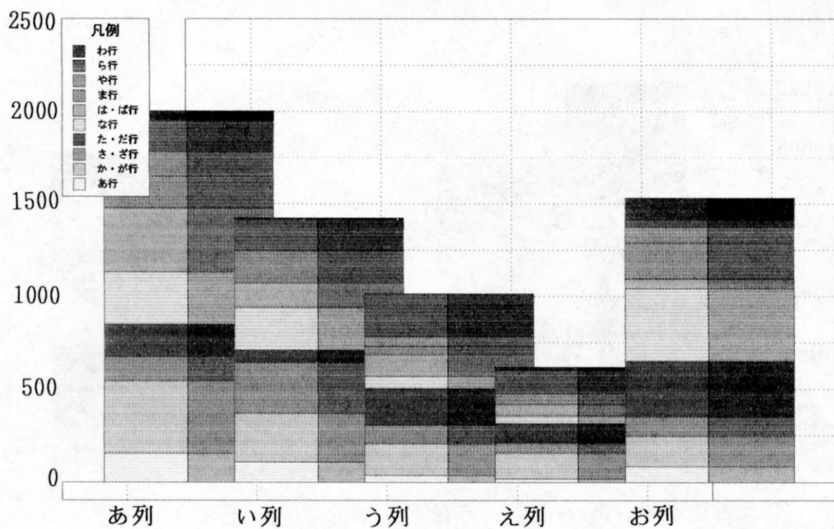


上代日本語にはラ行音で始まる語がほとんどなかったことを反映しているかと思われる。

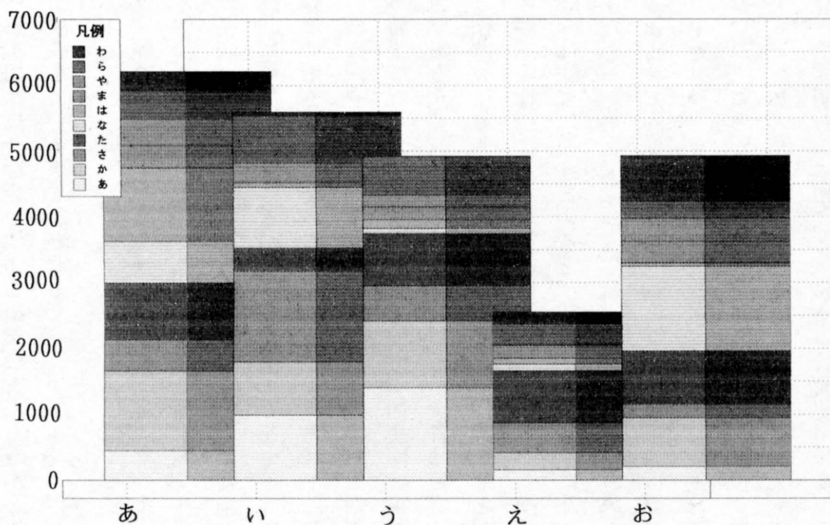
漢文脈の日本語、つまり漢文訓読が下地になっている文と和語を基本とする文では、使用語の相が異なり、それは名詞はもちろん形容詞、副詞のみならず助動詞や補助動詞にも及ぶとされるが、音節文字にもその差はあらわれるであろうか。そのことを検証するテキストとしては、ほぼ同一の題材と趣向を詠じた漢詩と和歌を対置させる詩集、たとえば『新撰万葉集』や『和漢朗詠集』が適当であろう。前者は一首ごとの対応、後者は漢詩一に對し和歌数首の例もみられるが、ここでは紫式部と同時代の藤原公任撰になる『和漢朗詠集』をとるとし漢詩の訓読みは底本に従うとすると、結果は図のようになる。

すなわち和歌にあっては、母音列の順位は万葉・古今・新古今と同様にアオイウエとなり、頭子音行ではカタラハサマアヤワの順位となつて、大きな変動はみられない。ところが漢詩については、母音列の順位がアイウオエとなり、頭子音行の順位はカタアサナハラマヤとなつて、明らかに和歌とは異なる音節使用頻度がみら

和漢朗詠集(和歌)



roueishu-kansi



れる。またこのグラフにはあらわれていないが、パプペ
ポが出現すること、撥音「ん」の使用度が非常に高いこ
と(5%弱)も、和歌の場合とくらべて顕著な相違をな
している。文章博士の家系であり漢詩・漢文の素養が十
分あったとみられる紫式部ですら、つれづれに書を読ん
でいて、「おまへはかくおはすれば、御さいはひはすく
なきなり。なでふをんながまんなふみ(真字書)は読む。
むかしはきやう(経)よむをだに、人はせい(制)しき⁽⁸⁾」
と非難される時代に、専ら男子のものであった漢詩漢文
の世界と、男女ともに参加した和歌の世界におけること
ば(音)の違いを、上は如実に反映しているといえよう。
なお、朗詠集の漢詩に「ん」が多い理由として、名詞自
体に含まれるもの、ん音便(うかんで、いんで等)、漢
文訓読読みの動詞(ぎんず、くんず等)が多いことがあ
げられよう。

この項の最後に、個々の字(音節)で出現頻度が高い
ものを参考までに列挙しておく、

万葉集 のしにもかとなはまき
古今集 のなかとはしにもきる

新古今 のかなしはきもとるみ
朗詠和 のかなはしとにるくも
朗詠漢 うのんしいかにくはり

三・一 標本抽出と検定

やまとことばの一特徴が少なくとも和歌の場合、音節
文字かなの分布傾向、すなわち母音列、頭子音行の頻度
順位として顕在化することを前節で示したが、特に母音
列の順序(アオイウエ)は筆者が試みに調べた限りでは、
『和泉式部歌集』『紫式部歌集』『長秋詠草』(俊成)、『百
人一首』、あるいは『伊勢物語』や『和泉式部日記』中
の和歌においても変わりが無い。その一方、『万葉集』
でも一音一字表記ではなく字訓表記をとり長歌を多く含
む一、二巻などでは、イ列がオ列より少し多くなってア
イオウエの順になるところをみると、アオイウエの順位
はむしろ平安時代に顕著な特徴といふべきかもしれない。
しかしこれについては今後なお検討することにして、こ
こでは特に分布が似通っていた『古今集』と『新古今
集』について五十音図の行別、列別の頻度数による分割
表を作り、カイ二乗一様性検定を試みてみよう。この両

者は相互相関値でみる限り極めてよく似た分布を示していたが、母音列データによる自由度四の検定統計量は四九・四〇二、また頭子音行データによる自由度一〇での検定統計量は七八・三五五となるので、いずれも有意水準五%、一%とも、両者の分布が一様性をもつ、とする仮説は棄却される。『古今集』の文字数は三万四千余り、『新古今集』は六万一千以上であるが、この結果は一見似通った音節分布でもカイ二乗検定が判別の方法として有効であり、作者問題検討の一手段となりうることを示している。

さて『源氏物語』はのべ語数四〇万ほどであり、そのうち助詞・助動詞が二〇万ほど、また異なり語数は一万三千から一万四千ほどになるとい⁽⁹⁾う。これをもし全文ひらがな入力すれば恐らく百万字を越えるであろう。全文入力による悉皆調査は別の機会に譲るとして、さしあたり標本調査を行うとすれば、どのような方法で標本を抽出するのが適当であろうか。まず考えられるのは、全頁に通し番号をふって無作為抽出を行うことであろう。しかし『源氏』の各巻にはかなり長短のばらつきがあり、

機会的に抽出を行うとまったくデータにあらわれない巻が生ずる可能性がある。これは『源氏』の成立事情や構成を考えるとなるべく避けるべきであると思われる。そこで筆者はまず各巻の巻頭と巻末および中央部から一定量の文字データをとって予備調査を行った。その結果、巻頭と巻末のテキストは作者の意識が比較的強く働いためにやや特異値を示す傾向があり、無意識な文体特性をみるための安定したデータとしてはむしろ各巻中央部のテキストをとるべきであると判断した。むろん念のため無作為抽出方式⁽¹⁰⁾も採用して後で比較してみたところ、検定結果などに大きな差異はないことが判明したが、第二巻から第四一巻までのうち『源氏』の代表的巻の一つである「明石」をはじめ五巻ほどが無作為抽出方式では標本と無縁となつて、筆者の懸念を裏打ちした。

「宇治十帖」の作者問題を検討する為のテキスト・データは、以上の予備作業の末、「桐壺」「匂宮」「紅梅」「竹河」を除く各巻の中央部から各巻の長さに応じて機械的に採取し標本とした。「匂宮」以下については既にふれたが、第一巻の「桐壺」を省いたのは、写本上および内容上この巻が他との均質性を欠いているともいえる

からである。すなわちこの巻は三条西本には欠如し飛鳥井本ではこの巻だけ別人の筆になるとされておき、また主人公誕生から思春期まで一気に紹介する筋書きは他の各巻とは物語の性質やテンポが異なっており、その意味で特異性を示す可能性はある。むしろこれらの巻は別途入力しテキストデータ・ファイルを作成している。

正編と仮に呼んだ部分については、物語の展開の上から定説に従って、第二巻「帚木」から第三三巻「藤裏葉」までを第一部とし、第三四巻「若菜」より第四一巻「幻」までを第二部にわけ、これに第四五巻から第五四巻までの「宇治十帖」という、つごう三セットの標本を作成し、一様性検定等に使用することとする。すなわち音節文字を無意識に使用される互いに独立な変数とみなす意味から、固有名詞（及びそれに相当する特殊呼称）は漢字表記のままにして区別し、他の部分は一・二の原則に従った音節文字データの標本である。

対照のために筆者は更に、以下の三つのグループに分けられるテキストから標本を作成して検定等に利用した。

(一)『紫式部日記』と『栄花物語』、(二)『竹取物語』『落窪』『宇津保物語』『狭衣物語』、(三)『手枕』（本居

宣長）。これらは散文を主とするものであるが、中に含まれる和歌のみを抽出して別のファイルを作成したことはすでにふれたところである。(一)は全四十巻からなる『栄華（花）』のうち巻三十までは紫式部に近い赤染衛門が作者に擬され、特に巻八は『紫式部日記』を直接利用しているとみなされるところから、頭子音行および母音列の観察度数（頻度）によるカイ二乗一様性検定の性能を検証するためである。(二)は『源氏』に先行する代表的物語類、およびかつては『源氏・狭衣』と並び称されたもので、特に『狭衣』は相当程度に『源氏』を摸しているゆえ、作者問題の検討の上で、他と同様の一様性検定のみならず様々な角度から検証する必要があると思われるもの、また(三)は国学の大家が戯れに『源氏』に欠ける巻を補おうと試みたものであり、『狭衣』同様に様々な検討に値するであろう。

三・二 文章長、頭子音行、母音列による分割表の検定

はじめに『源氏』の一部、二部、三部（宇治）及び『狭衣』の文章の長さを比較してみると次のような統計

量がえられる。「源氏」は三部を合算したもの)

文数	平均	標準偏差	変異係数
第一部	二九六	五五・六八	四三・〇七
第二部	一四〇	六〇・六三	四七・四三
第三部	二〇二	六六・六三	四六・七〇
源氏	六三〇	五五・四九	四四・五五
狭衣	一五九	六〇・〇三	五九・三三

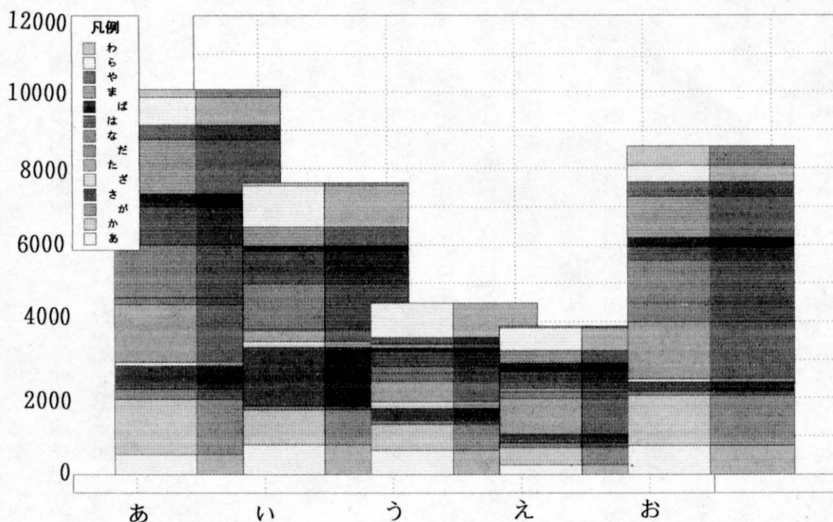
ここにいう文章長は一・二で述べた原則に従っており、漢字等の扱いが異なるゆえ単純に比較できないが、参考までにあげておくと、安本は正編四四帖の文章長平均を約五五字、宇治十帖のそれを約六十字とし、また大野・丸谷によるといわゆるa系列(紫の上系)が五三字、b系列(玉鬘系)は八五字となっている。特に後者とは差がかなり甚だしいが、この点に関してはいずれ精査するとして、さしあたり右の統計量をもとに、標本平均の差の検定(±検定)を行ってみると、『源氏』と『狭衣』では五%水準で仮説が棄却され標本平均の差は有意であるとなるのに対し、源氏三部間相互の検定ではいずれも仮説保留の結果となる。

次に『源氏』、『狭衣』、『手枕』の音節文字データの分析をしてみよう。五十音図の形で網羅的に表示することは、紙幅の関係上不可能であるので、ここではまず二・二節同様、『源氏』における仮名分布をグラフで示し、次に観測度数を頭子音行と母音列別に集計したものを、『源氏』一・二・三各部及びその合算、『狭衣』、『手枕』の順に表の形で示すことにする。なおグラフは一・二・三部合算値に基づいている。

	一部	二部	三部	源氏	狭衣	手枕
あ行	1299	655	947	2901	837	504
か行	2065	1171	1565	4801	1292	819
が行	320	146	228	694	229	105
ひ行	1326	668	932	2926	958	480
ひ行	226	105	199	530	153	93
た行	2100	1085	1554	4739	1443	722
た行	491	244	360	1095	341	178
な行	1830	931	1227	3988	1178	609
は行	1343	697	1022	3062	851	469
ば行	445	220	309	974	361	116
ま行	1486	852	1106	3444	1039	495
や行	385	225	295	905	271	134
ら行	1671	853	1230	3754	1079	538

(15) 源氏物語・宇治十帖の作者問題

源氏物語



わ行	375	177	254	806	234	113
ムン	174	74	121	369	111	51
御	130	77	78	285	101	86

あ列	4410	2386	3274	10070	3101	1529
い列	3392	1759	2478	7629	2219	1202
う列	1948	1031	1501	4480	1221	745
え列	1740	908	1225	3873	1180	556
お列	3872	1945	2750	8567	2535	1343

(ただし、母音列の集計値には、ム、ン、御の度数が含まれていない)

頭子音行をそれぞれ一クラスとみなした分割表を念頭において、まず『源氏』の各部分でカイ二乗一様性検定を行ってみる。たとえば第一部と第二部の頭子音別分布は一様性をもつ、とする帰無仮説をたて、有意水準5%、自由度一五として、検定統計量を算出すると二一・三七八となる。自由度一五で有意水準(危険率)〇・〇五及び〇・〇一のパーセント点はそれぞれ二四・九九五八、三〇・五七七九であり、検定統計量はこれより小さな値ゆえ、仮説は保留される。同様に第一部と第三部(宇治十帖)、第二部と第三部について検定統計量を算出する

と、それぞれ一六・三六三、二一・七九二となって、いずれも一様性仮説は保留される。母音列についても第一部と第二部、第一部と第三部、第二部と第三部の順に検定統計量はそれぞれ四・〇六九、五・一五六、二・三五二となり、自由度四における五％点、一％点は九・四八八、一三・二七七であるから、仮説は保留され、これらの標本が同一母集団に属する確率は極めて高いといえよう。むろんこのことが直ちに『源氏』の正編と続編の作者を同一とみなす積極的理由にはならないが、少なくとも別人である可能性を相当程度狭めるものであるう。

なぜなら同様の方法で先にあげた諸テキストと『源氏』の音節文字分布の一様性検定を行うと、いずれも本文および和歌の頭子音行あるいは母音列の分布が『源氏』とは異なり同一母集団に属するとはみなせない結果を示すからである。引用関係を指摘される『柴花』巻八と『紫式部日記』の場合にも一様性仮説は棄却されるが、ここでは『源氏』の文体を意識的にうつつしたとも思われる『狭衣』と『手枕』についてのみ、頭子音行データと母音列データによる『源氏』(三部合算データ)との一様性検定の結果を明示し、この項を閉じることしよう。

すなわち『狭衣』の場合、頭子音行による分割表では検定統計量が三六・七九七(自由度二五)となり有意水準五％、一％、〇・一％とも仮説棄却となるが、母音列では一一・五二四(自由度四)となり、有意水準五％の場合のみ帰無仮説棄却となる。更に詳しくみると第一部との組み合わせでは頭子音行、母音列とも仮説は保留され、第三部との組み合わせでは行列共に仮説棄却となって、『狭衣』が『源氏』でもやや憂愁の気配の濃い「宇治十帖」より正編により近いことを窺わせている。『手枕』は短いながら完結した巻をなすゆえ、抽出標本データと単純に比較するのは危険であるが、母音列の分割表による一様性検定では仮説が保留され、頭子音行では検定統計量が五九・五一六という大きな値となって仮説が棄却される。五％点のみならず〇・一％点をも遙かに越えるこのような値での仮説棄却は作者問題の判定にかなり決定的な論拠となりうると思われる。

筆者はまた助動詞および助詞の意味・機能の変遷という面からの分析を試みつつあるが、現在のところ助動詞「なり」については『源氏』『狭衣』と『手枕』の間に顕著な相違を見出したので簡単に報告しておきたい。「な

り」には古くは動詞・助動詞の終止形をうける伝聞・推定の「なり」(第四類)と、体言および連体形をうける指定の「なり」(別類)とがあり、時代とともに前者は衰滅し後者がもっぱらとなるが、『源氏』『狭衣』では両者の比率がほぼ一対十程度であるのに対し、『手枕』には前者の例がみあたらない。これと同様の標識語を一段と増やし、また連辞の分布その他多面的な調査を行う一方、多変量解析の計量言語学的応用の可能性をさぐり、国文学分野ではたとえば『和泉式部日記』の作者問題の解決などに多少とも寄与することができれば、というのが筆者の希望であり今後の課題である。

三・三 結語

以上によって筆者は、音節文字かなの頻度分布は日本古典文学の特徴をかなりの程度うきほりするものであること、五十音図の頭子音行別、母音列別の頻度データは文体の差異を検証する一つの有効な手段であること、また『源氏物語』正編と続編「宇治十帖」の作者は別人とは考えられないこと、を主張しかつ論証した。

(1) 『源氏物語』一(日本古典文学全集一、二、校注・訳 阿部秋生・秋山虔・今井源衛)、解説(三頁)

(2) 「文体統計による筆者推定——源氏物語、宇治十帖の作者について——」(『心理学評論』一九五八・二)、現代の文体研究(『岩波講座』『日本語』一〇、一九七七)

(3) 原本が残っていないので精確な数は求めようがないが、固有名詞や官職名などを除いて大体百余りの漢字を使っているのではないかと推定される。

(4) 『古事記』序(日本古典文学大系一、倉野憲司・武田祐吉校注) 四九頁

(5) 『かな』(小松茂美) 三七頁

(6) 『文章を科学する』(前川守) 七〇頁

(7) 『日本語の文法を考える』(大野普) 一九二頁

(8) 『紫式部日記』(日本古典文学全集一八) 二四〇頁

(9) 『光る源氏の物語』(大野普・丸谷才一) 上一五頁

(10) 対照用の無作為抽出標本は次のような手順によった。手元にある二種の独訳本のうち二巻本のマネッセ版の頁配分が比較利用しやすいので、一様乱数表により抽出すべき頁番号を定め、次にそれと対応する『源氏』の本文を抜き出して標本とした。勿論こんな面倒な手続きは実は不要で、乱数もコンピュータで簡単に得られるのだが。

(一橋大学教授)