

梶井基次郎『檸檬』における文体の計量分析

和田弘名

1. はじめに

大阪生まれの作家梶井基次郎は、1925年同人誌『青空』に発表した「檸檬」から、1932年『中央公論』に発表の「のんきな患者」まで、作品はわずかに20編ほどしか残していない。しかし、短い作家人生ではあったが、その作風は第一期（1925～1926）の東京時代、第二期（1927～1928）の伊豆湯ヶ島時代、第三期（1929～1932）の大阪時代に分けられ、「軽やかで明るい」、「暗くて険しい」、「軽やかで平俗さを志向」が各期の特徴とされている¹⁾。これら各期の作風の特徴と、文長、句長、漢字率など各種の計量処理から得られた結果との関連を考察し、その文体の特徴を探ろうとしている。

テキストとして、作品集『檸檬』²⁾の中の全20編をデジタル化して用いた。発表年とセンテンス数は次の通りである。ただし、発表年はテキストに表示されているものを採用したので、初出誌の年月とは異なっている。

1924. 10	檸檬	141	1927. 10	桜の樹の下には	52
1924. 11	城のある町にて	698	1927. 11	器楽的幻覚	61
1925. 6	泥濘	229	1928. 2	蒼穹	58
1925. 9	路上	130	1928. 2	笥の話	50
1925. 10	橡の花	372	1928. 3	冬の蠅	341
1925. 12	過去	88	1928. 6	ある崖上の感情	310
1926. 5	雪後	263	1930. 5	愛撫	96
1926. 7	ある心の風景	239	1930. 9	闇の絵巻	138
1926. 9	Kの昇天	170	1930. 12	交尾	188
1927. 3	冬の日	378	1931. 12	のんきな患者	208
				合計	4210

2. 諸特性

2.1 文の長さ

作品の読みやすさ、歯切れのよさに関わる指標のひとつに文（センテンス）の長さがある。文の長さは、文の初めから句点までの文字数を文長としたが、会話文の場合に、句点が無く右

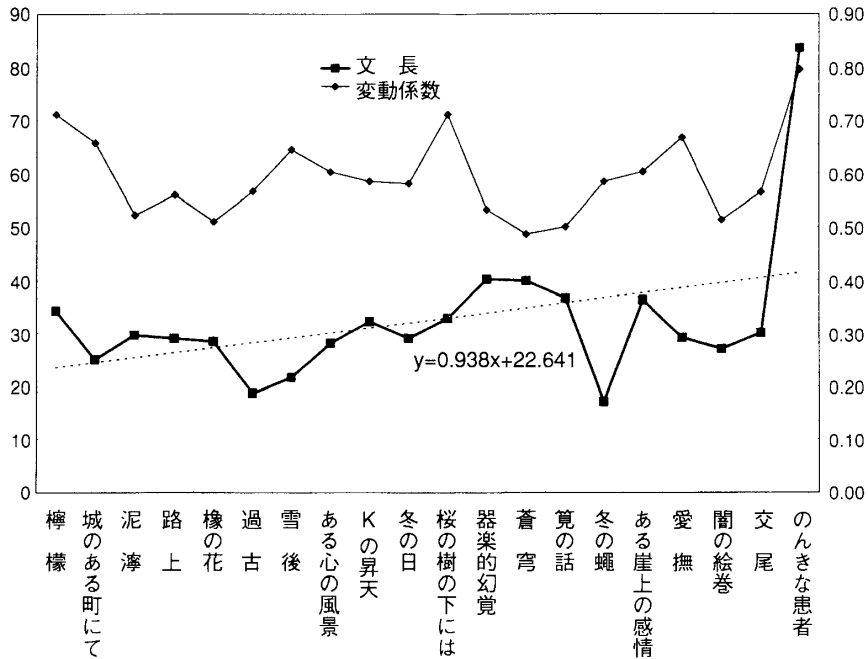


図1 文長変動

括弧で終わっている場合、それを句点とみなして、右括弧までの文字数を文長とした。疑問符(?)と感嘆符(!)も句点とみなした。20作品の平均文長を時系列に並べたものを図1に示す。20作品の平均は、32.5(標準偏差19.8, 変動係数0.6)であるが、各期別平均では一期が27.5, 二期が33.2, 三期が42.5で、三期は一期に比べて約1.5倍になっている。これは最後の作品「のんきな患者」の文長が83.7と非常に長いからである。全体では線形近似で0.9の増加傾向を示している。川端康成の「伊豆の踊り子」(26.1)に比べて、文長は25%ほど長い³⁾。

2.2 「漢字率」と「かな文率」

作品の読みやすさに関わるもうひとつの指標は漢字率である。各作品の漢字率の時系列変化を図2に示す。全体の平均は27.9(標準偏差10.8, 変動係数0.4)であるが、これを発表年順に見ると、最大は「過去」(一期)の33.8, 最小は「ある崖上の感情」(二期)の23.4である。一期の平均は29.2, 二期は27.0, 三期は26.6で、全体としては減少傾向(線形近似で-0.2)を示した。

漢字率と文長の時系列変化を重ね合わせてみると、その増減がまったく逆になっており、「文長が長くなると漢字率が減少する」ことが判る。「過去」と「のんきな患者」の場合が特に顕著である。

漢字率0%即ち漢字が使われていなくて、「かな」ばかりの文が20編中、次の4編(「桜の樹の下には」, 「器乐的幻覚」, 「蒼穹」, 「笈の話」)を除く16編に合計134件(3.2%)現れた。各作品の文の総数に対する「かな」文の割合(かな文率)と会話文の割合(会話文率)を図3に示す。かな文率の最大は、「雪後」の8%であるが、「どっこいしょ, どっこいしょ」

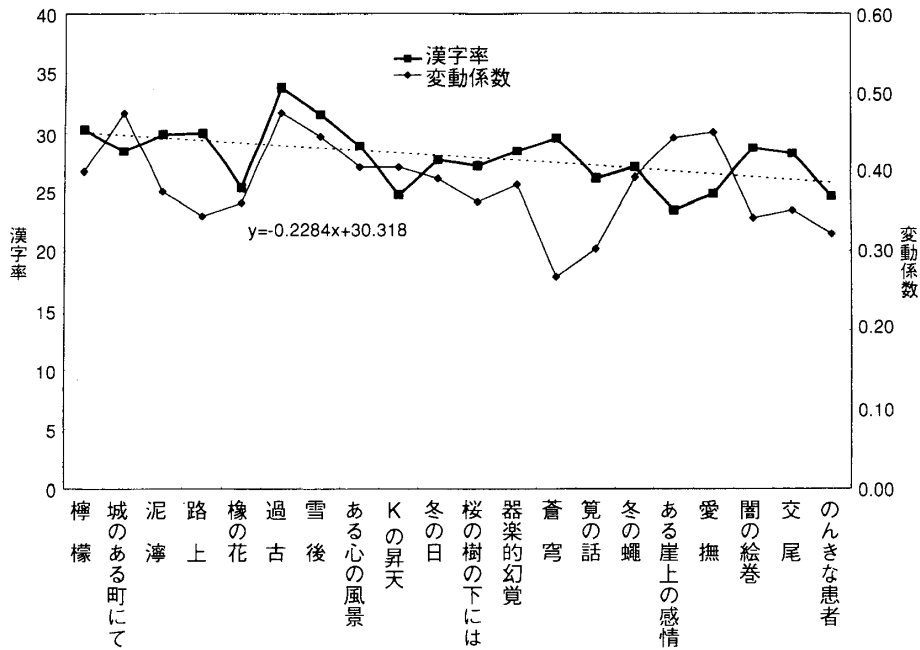


図 2 a 漢字率

とか「どうだったい」、「まあぼちぼちだ」のような会話文が用いられており、会話文率は 25% であった。反面、「のんきな患者」の会話文率が 25% と高いにもかかわらず、かな文率は 1% であった。「私」が客観的、主体的に描かれている「過去」、「桜の樹の下には」、「蒼穹」の 3 作品には会話文を必要としないためか、会話文率は 0% であった。

かな文率、会話文率の 20 作品での平均はそれぞれ 3.2%、8.9% で、ほとんど一定であった。

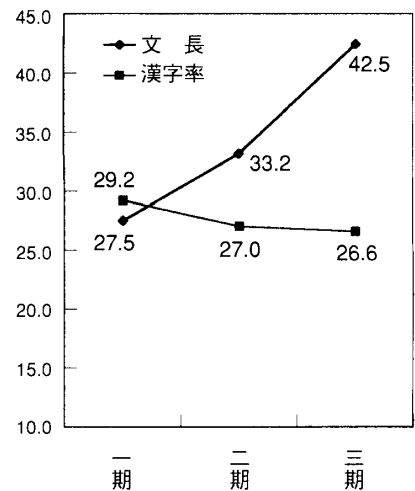


図 2 b 期別変動

2.3 句の長さ

日本語には読点についての定まった規則がないため、書き手の癖がもっとも出やすく、作風の違いを比較する指標と考えられている。読点までの句の長さを用いた作品の分類も試みられている⁴⁾。

読点のみを句の長さの判定の基準として求めた 20 作品の句長を図 4 に示す。二期の「器乐的幻覚」、三期の「のんきな患者」の句長が長いので、線形近似では、増加率 0.2 を示している。20 作品の平均は 16.5 (標準偏差 8.8, 変動係数 0.5) であるが、これを期別での平均で見ると、14.7 (一期)、17.6 (二期)、18.4 (三期) で、三期は一期より 25% 増加している。

読点を伴う文のみの長さの平均と句の長さとの関係を調べた結果を図 5 に示す。直線近似と比較すると、両者とも同じ増加傾向を示すが、文長の増加率は句長の約 6 倍である。読点を含む文長の 20 作品での平均は 48.9 (標準偏差 20.9, 変動係数 0.42) である。期別変動で

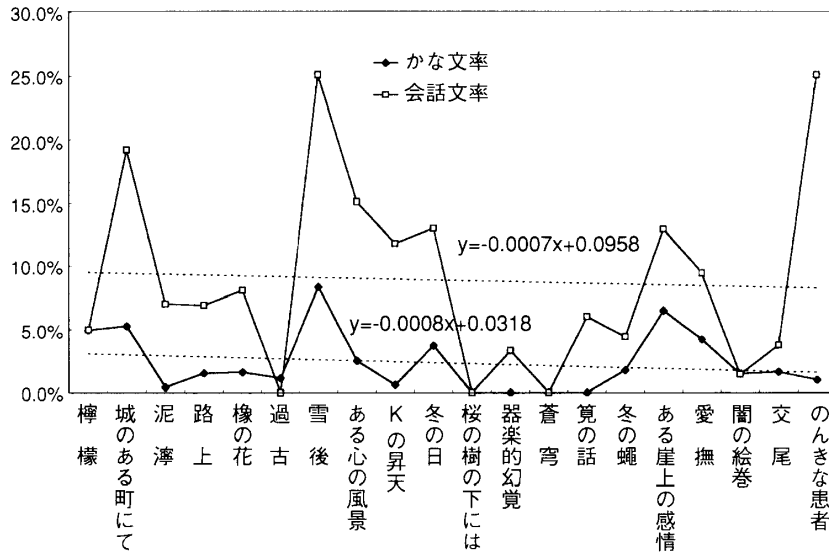


図3 かな文と会話文

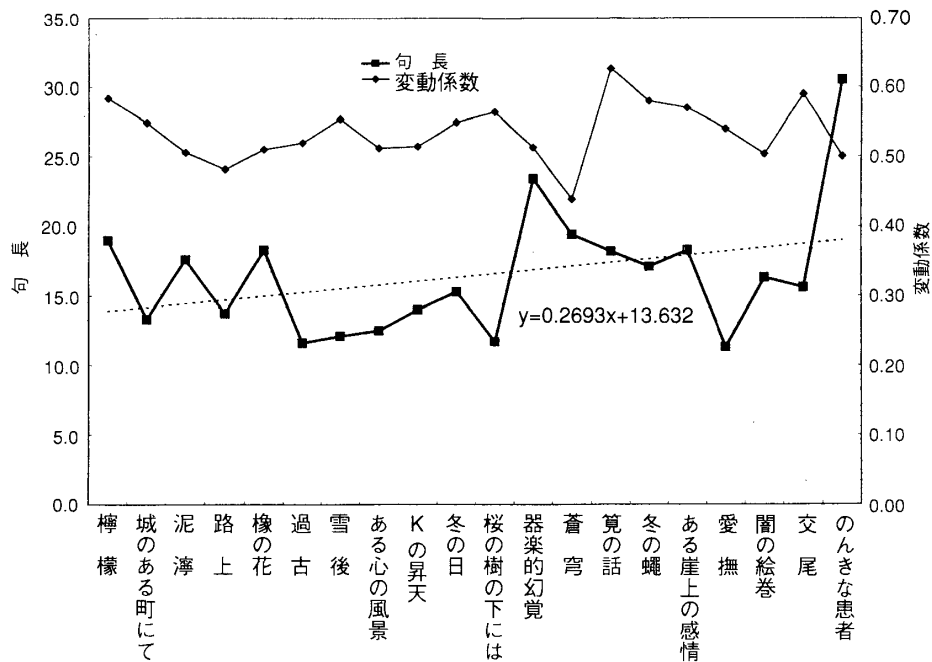


図4 句長変動

は、二期が一期より 20% 増加しているが、二期と三期の変動は少ない。

2.4 無読点文

読点は文を読みやすくし、その内容を理解しやすくするが、逆に読点のない文も作品に与えるテンポのよさとの関係で多用される。文の総数に対するこの読点の無い文の割合（無読点率）と無読点文の長さ（無読点文長）を図6に示す。20作品の平均は、無読点率が62.5%（標準偏差12, 変動係数0.2）、無読点文長は22.7（標準偏差5.5, 変動係数0.2）である。全体的には、無読点率は減少傾向（線形近似で-0.7）を示しているが、無読点文長の時系列変

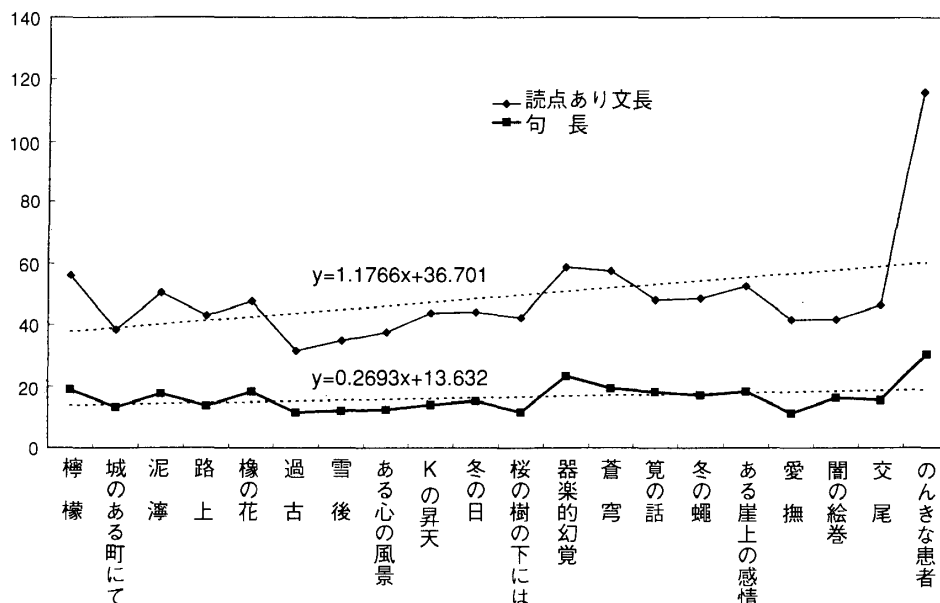


図 5 a 読点を含む文のみの長さ

動からは、緩やかな増加傾向（線形近似で 0.4）が読み取れる。期別の変動でみると二期と三期は変化していない。

無読点率と無読点文長の増減傾向はほぼ同じであるが、「のんきな患者」のみ無読点文長が最大（37.3）であるにもかかわらず、無読点率は 40.9% ときわめて低い。「冬の蠅」にもややこれに似たような傾向が見られた。

読点を伴わない文の長さ（無読点文長）と、読点を伴う文の長さ（読点文長）を 20 作品の平均で比較した場合、1:2 である。また、後者の増加率は前者の増加率の 4 倍である。

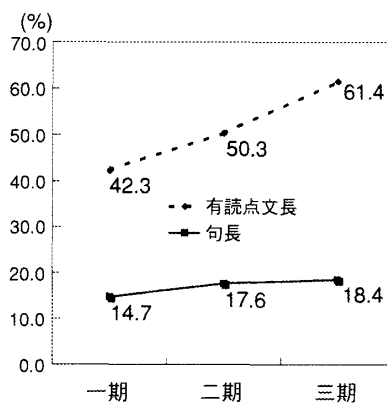


図 5 b 期別変動

2.5 クラスタ分析

三期にまたがる 20 作品に潜在する文体の類似性を見るために、クラスタ分析によるデンドログラム（最近隣法によるユークリッド距離）を求めた。20 作品をケースとし、文長、漢字率、句長、無読点率、無読点文長を変数として求めたデンドログラムを図 7 に示す。「算の話」と「ある崖上の感情」、「泥濘」と「椽の花」、「過去」と「雪後」の 3 組は同一時期の作品であるため、5 変数から見て似た文体から近い距離で結合したといえる。二期の終わりごろの作品「算の話」と「ある崖上の感情」が、最初の作品「檸檬」に近いことは興味深い。「冬の蠅」と最後の作品「のんきな患者」の文体は、他の作品とは特に異なっていることが抽出されている。後者については、5 変数と作品の関係を示す図 8 を見てもこの特殊性が読み取れる。「路上」（一期）と「交尾」（三期）の間隔は約 5 年で、「ある心の風景」（一期）と「愛

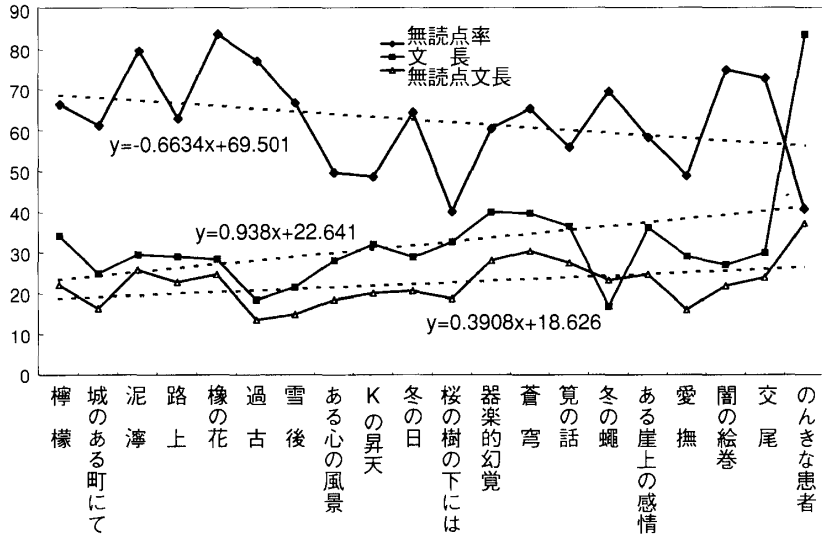


図 6a 無読点率と無読点文長

撫」(三期)の時間間隔は約4年であるが、近い距離での結合は文体の類似性を窺わせている。

2.6 文長変化の周期性

各作品に潜在するある種のリズムを文長変動の周期から読み取るために、自己相関係数を求め、コレログラム⁵⁾を作成した。20作品の文長周期を図9に示す。全体の平均は4.9(標準偏差2.5, 変動係数0.5)である。期別で見ると、一期の平均3.3が、二期では6.7に増え、三期では5に減っているが、全体を直線近似で見た場合、増加率0.21で増えている。二期の

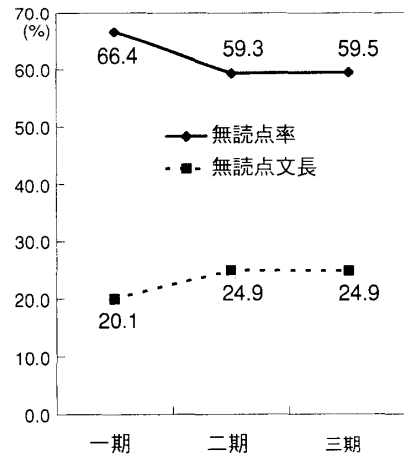


図 6b 期別変動

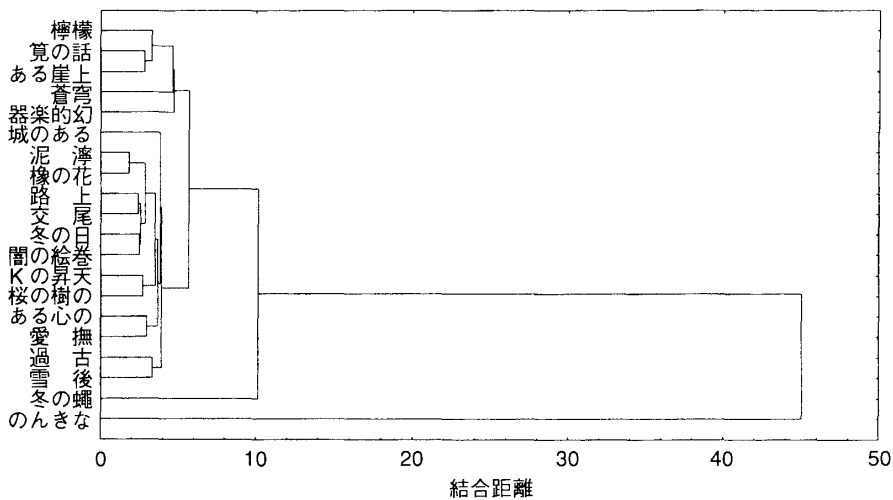


図 7 デンドログラム

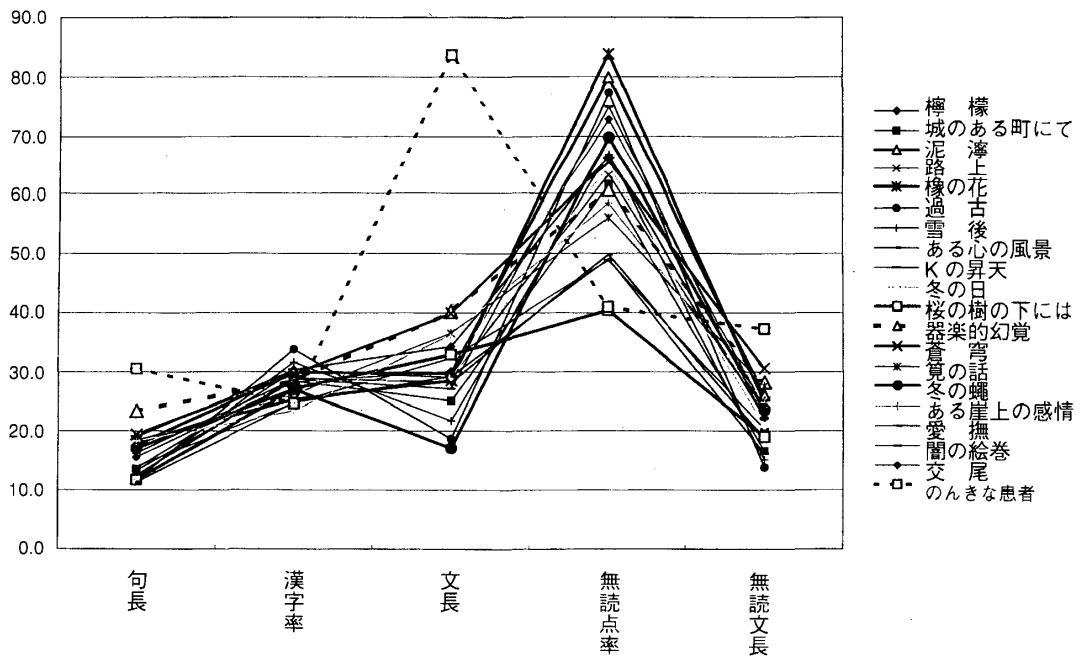


図8 変数から見た20作品

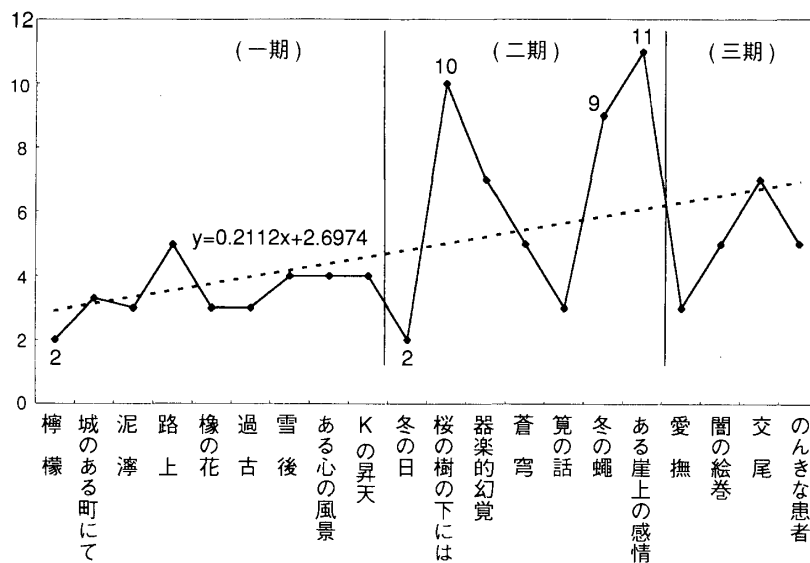


図9 文長周期

「桜の樹の下には」の周期が10, 「冬の蠅」が9, 「ある崖上の感情」が11で, 全体の平均5.2に比べてかなり高く, 最小の「冬の日」2の約5倍である。「冬の蠅」は文長がもっとも短く, 文長周期が3番目に高いのが特徴的である。「檸檬」, 「冬の日」の周期2は, ひとつ前の文長の影響が大きいと考えられる。二期の周期が, 一期と三期に比べて特に変動が大きい。一期と三期には周期が大きくなる山が1回であるが, 二期には山が2回現れた。

3. ま と め

20 作品の平均は、文長 32.5、漢字率 27.9%、かな文率 3.2%、会話文の出現率 8.9%、句長 16.5、文長（読点を含む文のみ）48.9、無読点率 62.5%、無読点文長 22.7、文長周期 4.9 であった。

20 作品の文体について時系列変化を見た場合、文長は増加傾向（0.9）を示しながらも、漢字率は減少（-0.2）しているので、読みやすくやわらかな印象を与えるように作風が変化したと考えられる。「のんきな患者」を除けば、一般的に会話文が多いとかな文率も高くなる傾向を示した。

また句長の増加（0.2）とともに文長も増加するが、増加率は文長が句長の 6 倍であった。読点の無い文の出現率は減少（-0.7）しているが、その文長は増加傾向（0.4）を示した。しかし、読点のある文の長さの約半分であり、読みやすさへの影響は殆ど無いように思える。読点の使われ方に書き手の個性が現れやすいと言われるが、逆に読点を用いない文の分析もまた文体分析の指標になるものと思われる。今後この指標を中心に他の作家の作品分析を進めたい。

5 変数によるクラスタ分析では、4 年乃至 5 年を隔てた作品間の類似性が抽出できたが、これは三期の作風が一期の作風に回帰しながら「平俗さ」を志向したと言われることと関連して興味のあるところである。

コレログラムは、作品に潜在する文長変動の周期性を浮かび上がらせ、文体が持つある種のリズムを示唆してくれる。20 作品の文長変動の周期は増加傾向（0.2）を示したが、特に二期において最大と最小の周期差 9 の変動の大きさは、作風との関連性が窺えて興味深い。平均周期の 2 倍の作品「桜の樹の下には」、「冬の蠅」、「ある崖上の感情」には、ゆったりとした旋律が潜在していると考えられる。他の作品についての変動周期の分析を行い文体分類の指標作りを進めたい。

参考文献

- 1) 磯貝英夫：梶井基次郎論，現代日本文学大事典，p. 237-240, 1965（明治書院）
- 2) 梶井基次郎：檸檬，1998，（新潮社）
- 3) 和田：「伊豆の踊子」における文の長さの規則性，青須我波良（53），p. 137-143, 1997
- 4) 村上征勝：文章分析と統計学，数理科学（389），p. 27-33, 1995
- 5) 藤沢：池内：予測の理論，p. 51-54, 1988，（現代数学社）