

# コンピュータを利用した文学事例研究

浅井邦彦\*

## A case study on comparison of literatures by computer application

Kunihiko ASAI

Recently “Multi-Media” is an emerging technology in our life.

With increasing use of “Multi-Media” in various fields more computers will be put into practical use than ever before.

In this paper a comparison between two authors, Naoya-Shiga and Junichiro-Tanizaki is made as an example.

It is a remarkable point of this paper that FORTRAN is used in this study.

### はじめに

現在、マルチメディアが話題になっている。マルチメディアが本格的に実用化すれば、今まで以上に、家庭や個人生活まで多分野に、コンピュータが浸透することが予想される。

このような背景の中で、コンピュータの新しい使い方の一つとして、文字研究への適用をとりあげてみた。

本稿での具体的事例は、志賀直哉と谷崎潤一郎の作品比較である。もとより、ここで、志賀直哉論や谷崎潤一郎論を展開するものではない。作家作品論そのものよりも、両作家作品の比較をコンピュータの代表的プログラム言語である FORTRAN を道具にしたところにポイントがある。

### 1. 文学研究の統計分析の意義

文章の統計分析に関しては、波多野完治氏（参考資料1）の業績が大きい。波多野完治氏は、作者別、作品別に、品詞の頻度検定や色彩語の出現回数等を調査することで、作者や作品の文章の特徴の客観的裏づけを行っている。波多野完治氏が対象にした作者として、志賀直哉、谷崎潤一郎、芥川龍之介、泉鏡花、薄田泣菫等、多くの異なった作風の書き手を取りあげ、例示しながら分析している。

更に、和歌、俳句、現代詩、新聞の評論等、小説以外でも、俎上にのせた文章の種類は多方面

---

\*電気工学科

に及んでいる。

ところで、一般に、文章は、情報の伝達を行えばよいし、文学作品は味わえればよいのであり統計分析などを加えるのは邪道だという考えがある。確かにそれも一理ある。しかし、作品の見かけ上の価値は、読者との相対的位置づけによることが大きいのも事実である。

筆者の場合を例にとる。17歳の夏、サンテグジュペリ (Antoine de Saint-Exupéry 1900～1944 フランス) の「星の王子様」を読み、すっかり感激した。触発されて、それ風の童話を書きはじめたぐらいである。その感動自身は何年も変わらずに残っていたので、最近、もう一度、この作品を読んでみた。読んだ後、正直なところ、何故、17歳の時に、あんなに感動したのかという疑問である。

従って、特定の読者による特定の作品という限定条件をおいても、読者の諸々の背景が変われば、作品の相対的価値も異なって来るのが普通である。

もとより、ある特定の読者が年齢を重ねても、いつも感動を与える作品もあるだろう。また、その作品が特定の読者だけではなく、性別、年齢、生活環境を越えて多数の読み手の心をひきつけることができれば、それこそ本当の意味で名作と言えるだろう。ただその場合でも感動のさせ方には、多様性があるかも知れない。

作品の統計分析の意義は、読者の目から、ひとまず作品を切り離し、作品自身の独立的な固有性を判断することであり、その結果として、作品を生んだ作者論の一端を担うことになる。

## 2. コンピュータによる作風の比較

現代作家は、志賀直哉か谷崎潤一郎の影響を受けている場合が多いと言われている。このことについて、波多野完治氏は、「文章心理学入門」の中で、深田久弥氏の説として述べている。この説明によると、両作家は、対蹠的な文章を作成すると言っているが、小林秀雄氏はその相違について志賀直哉の文章は観念的であり、谷崎潤一郎のは即物的であると性格わけしている。

さて、本稿では、文学専門家のそのような言葉を踏まえながら、コンピュータ利用による両作家の作品を客観的に分析した。

本来、作品の客観的分析は、波多野完治氏が試みたように、品詞の出現頻度や色彩語の使われ方など多面的に行うべきであるが、本稿では「的」のしぼり方を徹底することで、相違がかえって、鮮明になるのではないかと思い、センテンスの長さにもみ焦点を当てることにした。

最初、志賀直哉の作品は「城の崎にて」、谷崎潤一郎のは「蓼喰ふ蟲」のみをピックアップした。ところが、両作品の統計的分析結果に違いが出たとしても、たまたまそうなったかのかも知れないと思い、両作家からもう一つずつ作品を追加選択した。志賀直哉は「真鶴」、谷崎潤一郎は「春琴抄」である。両作家ともにこの他に多くの作品を書いており、選択に迷うのであるが、それほど深く考えないで、二作品を対象とした。

この四つの作品に関して、センテンスについて、最長、最短、平均、標準偏差を求めたのが、次の「表-1」である。

$n$  このセンテンスの長さを  $a_1, a_2, a_3, \dots, a_n$  とする。 $\bar{a}$  は平均,  $s$  は標準偏差とすると

$$\bar{a} = \frac{1}{n} \sum_{i=1}^n a_i \quad s = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})^2}$$

以下, A系列を志賀直哉, B系列を谷崎潤一郎とする。

A-1: 「城の崎にて」

A-2: 「真鶴」

B-1: 「蓼喰ふ蟲」

B-2: 「春琴抄」

**\*\*SAKUHIN:NAGASA-NO-HIKAKU**

SHIGA-NAOYA  
A-1:KINOSAKI-NITE  
A-2:MANAZURU

TANIZAKI-JUNITIROO  
B-1:TADEKUU-MUSHI  
B-2:SHUNKIN-SHO

	SHIGA-NAOYA		TANIZAKI-JUNITIROO	
	A-1	A-2	B-1	B-2
SAITYO	106.	107.	517.	183.
SAITAN	3.	7.	16.	15.
HEIKIN	25.17	28.90	80.54	76.35
HYOJUN -HENSA	15.38	17.29	76.61	34.25

表-1

この「表-1」から、予想通り、志賀直哉のセンテンスは短く、谷崎潤一郎のは長いことがわかる。なお、この表から、「城の崎にて」の最短センテンスは3文字である。これは、「蝶々だ」である。

言うまでもないが、一つのセンテンスの長さとは。(読点)の次の文字から。(読点)の前

の文字の数であり、漢字も仮名も一文字は一文字である。漢字、仮名の出現頻度と作品の性格とは相関があると思われるが、本稿では触れないことにする。

「表-1」について、更に、次のことがわかる。志賀直哉の作品同志である「城の崎にて」と「真鶴」の最長は、106と107であるのでほとんど同じである。また、谷崎潤一郎の「蓼喰ふ蟲」と「春琴抄」の最短は、16と15で、きわめて近い数字である。

いずれにしても、センテンスの長さに関して、同一作者の別作品を比較した時、類似傾向が出て来る。

そして、本稿のメインテーマである作者の違いによる作品のセンテンスの長さの相違は大きい。(A系列とB系列の比較)

この他、標準偏差値について、志賀直哉は作品間の違いは少なく(15・38と17・29)谷崎潤一郎は必ずしもそうではない。このことは、センテンスの長さのフレに関して、志賀直哉の作品間の違いは小さく、谷崎潤一郎は相対的に大きいことを表わしている。

以上述べたことを明確にするために、次の様に、コンピュータを使用して、「表-2」を作成した。

「表-2」は、両作家の上記四作品について、相対比較した相関数である。

たとえば、最長について、「城の崎にて」(A-1)を1(1.00と表記)にした時、「真鶴」(A-2)は1.01,「蓼喰ふ蟲」(B-1)は4.88,「春琴抄」(B-2)は1.73になる。

「表-2」でも、A系列同志、B系列同志が近く、A系列とB系列とが離れていることがわかる。つまり、センテンスの長さに関しては、作品による相違よりも、作者による違いが大きいことが定量的に言える。

以上述べたことは、志賀直哉と谷崎潤一郎のそれぞれ二つの作品をとりあげたに過ぎず、これをもって、一般論的な作者と作品との関係を論じるつもりはない。あくまで、コンピュータを利用した事例の紹介である。

\*\*SOKAN-1(SAITYO)

	A-1	A-2	B-1	B-2
A-1	1.00	1.01	4.88	1.73
A-2	0.99	1.00	4.83	1.71
B-1	0.21	0.21	1.00	0.35
B-2	0.58	0.58	2.83	1.00

\*\*SOKAN-2(SAITAN)

	A-1	A-2	B-1	B-2
A-1	1.00	2.33	5.33	5.00
A-2	0.43	1.00	2.29	2.14
B-1	0.19	0.44	1.00	0.94
B-2	0.20	0.47	1.07	1.00

\*\*SOKAN-3(HEIKIN)

	A-1	A-2	B-1	B-2
A-1	1.00	1.15	3.20	3.03
A-2	0.87	1.00	2.79	2.64
B-1	0.31	0.36	1.00	0.95
B-2	0.33	0.38	1.05	1.00

\*\*SOKAN-4(HYOJUN)

	A-1	A-2	B-1	B-2
A-1	1.00	1.12	4.98	2.23
A-2	0.89	1.00	4.43	1.98
B-1	0.20	0.23	1.00	0.45
B-2	0.45	0.50	2.24	1.00

表-2

- \*\* SOKAN-1(SAITYO) : センテンスの最長の相関表
- \*\* SOKAN-2(SAITAN) : センテンスの最短の相関表
- \*\* SOKAN-3(HEIKIN) : センテンスの平均の相関表
- \*\* SOKAN-4(HYOJUN) : センテンスの標準偏差の相関表

「表-1」,「表-2」の出力等に使用したコンピュータ資源は次の通りである。

(1)コンピュータシステム

FACOM・M-360

関東学院大学六浦キャンパス

SCC 設置

(2)プログラム言語

FORTRAN77

(3)プログラムの構成

「BUNTYO」と「SOKAN」の二つから構成されている。

- ・BUNTYO : センテンスの最長, 最短, 平均, 標準偏差を求める。
- ・SOKAN : BUNTYO の出力結果から, センテンスの最長, 最短, 平均, 標準偏差について, 4 作品間の相関を求める。

### 3. コンピュータ利用の光と影

マルチメディア時代が来るといふ。光ファイバ網により、情報の大量高速通信が可能となり、文字、画像、音声など媒体の異なるものが有機的に結合されて入出力される。自宅に居ながら、勤務、学習、買物が可能

となり、人々は移動のわずらわしさから解放される。その結果、余った時間を教養、娯楽に十分、ふり向けられる。バラ色の夢として描かれており、部分的には実現されつつある。これは一種の革命とも言えよう。

このような時代に於ける芸術、文化活動を考えてみる。

山梨県のある高校の校歌はコンピュータが作曲したものである。また、コンピュータに小説を創らせることが試みられた。新奇性という意味では面白いが、何となく違和感をおぼえる。

人間とコンピュータとを対比させ、コンピュータが得意とするところをコンピュータにやらせる。人間は人間しかできないこと、人間ならではのことを行う。また、その方向に、人間がエネルギーを費すことが望ましい。コンピュータに向いている対象が何であるかは、コンピュータの特性を考えればよいことであり、このことについては、多くの書物に記述されているので、ここでは述べないこととする。

いずれにしても、本来、人間に向いていることまでもコンピュータに無理にさせるのは問題であろう。

次に、不安になるのは、コンピュータによる情報操作が芸術活動にまでおよんで来ないかというおそれである。コンピュータ時代の到来以前でも、政治や宗教が強く支配している時代や地域には、自由な芸術活動がしにくく、創作のワクが狭い。文学活動の自由について、文芸評論家の中村光夫氏は批判的であり、「主人持ちの文学」を否定している。(参考資料2) また、経済学者のハイエク (Friedrich A Hayek 1899~1992 オーストリア) も芸術活動に於ける自由の必要性を論じている。(参考資料3)

さて、現在、先進国での情報化社会にあって、芸術活動は自由であり、束縛を受けていない。しかし、意図しない「見えない手」が、情報化社会の進展にともなって、芸術活動をある方向に引きずって行くことがあれば恐ろしい。

コンピュータは、雑事からの人間の解放者として歓迎すべきであるが、人間を超えた存在となつてはなるまい。

#### 4. コンピュータによる文学研究の応用例の提案

前章で創作そのものをコンピュータで行うことは、「いかがなものか」と述べた。しかし、コンピュータが文学研究の支援になれば、効果の面でも効率の点に於ても有意義であると思える。

たとえば、2章で論じた志賀直哉と谷崎潤一郎の作品比較をコンピュータを駆使して行うことにより、同一作家であれば作品が異なっても、類似傾向を示すことがわかった。こういう方法を展開していけば、作者不明作品の作者の推定が可能になるかも知れない。

更に、その応用として、古来、種々、言われて来ている「源氏物語」の宇治十帖が紫式部の作かどうかの統計的判定も可能になる。そして、もしも紫式部でないならば、一体、誰なのか、後世の作なのかなどの推定もできよう。

プロレタリア作家の小林多喜二の作品は、境遇の全く異なる志賀直哉の影響を受けていると言

われている。果たして、そうなのかどうかはコンピュータを使うことで客観的な結論が出させようである。

こうすることで、コンピュータによって、ある作家が別の作家の影響を受けている度合を調べることができる。

島崎藤村は、時代と信仰とによって、作風を変えて来たように思える。同一作家であっても、外部の影響度合を時系列的にどのように受けたかを、コンピュータがある程度、判定してくれるだろう。その他、比較文学の研究などにもコンピュータが利用できる場面がありそうである。

志賀直哉と谷崎潤一郎の作風比較をコンピュータを利用して実施した報告を中心に、文学研究に、コンピュータが利用できることを述べた。

### 参 考 資 料

- 1) 波多野完治「文章心理学入門」新潮文庫 1953年 132～144頁
- 2) 中村光夫「日本の現代小説」岩波新書 1989年 35頁
- 3) フリードリヒ・A・ハイエク「隷従への道」(一谷藤一郎, 一谷映理子訳) 創元社 1993年 206頁
- 4) 本稿で調査対象とした志賀直哉と谷崎潤一郎の作品は、河出書房「現代文豪名作全集」に収められているものである。

(平成7年10月14日受理)