

文学作品における文体構造の可視化

宮沢賢治「銀河鉄道の夜」の解析

金田知子（白百合女子大学大学院）、岩崎晴美（法政大学計算科学研究センター）、
斎藤兆古（法政大学工学部電気電子工学科）、宮沢賢治・堀井清之（白百合女子大学文学部）

Visualization of the sentence structures by wavelets transform

Noriko Kaneda, Harumi Iwasaki, Yoshifuru Saito, Kenji Miyazawa, Kiyoshi Horii

ABSTRACT

Previously, we have proposed a new computer oriented methodology for a sentence structure analysis based on a linear space theory. Its result has clarified that any literatures are composed of the controlled flow of the exciting and quiet sentences, sections and chapters. The flow mechanism of these sentences represents not only the identity but also feeling of the authors.

In the present paper, we apply our methodology to the Kenji MIYAZAWA's work focusing on the key words selection. As a result, it is revealed that the story "GINGATETUDOU NO YORU" is characterized by the actions of two boys and the word "2". In addition, It is pointed out that a small variation of the key words may become an important factor in the story.

Keywords: Sentence structure, wavelets analysis, orthogonal space

1. はじめに

文学研究における文体の解析は、作品中に使われる語句の種類や長さの計測をする時点で、研究者個人の感性・考え方など主観的要素が含まれることが免れない。

そこで、コンピュータを使用した数学的手法を使って、文学作品を客観的に可視化する解析の方法論を開発した[1-4]。この解析手法により、作品のイメージを可視化することが可能になったといえる。

本報では、宮沢賢治の「銀河鉄道の夜」に、この解析手法を適用し、登場人物の行動パターンから作品全体の構成と性質の可視化を試み、文学的観点から考察する。

2. 既往研究

本論文と同様の解析方法を用いて、夏目漱石の「虞美人草」と「草枕」の話法に関する解析をした研究がすでに報告されている[2]。この研究では、作品の始まりから終わりまでを時系列で捉えると、二作品はともに円環構造、つまり螺旋構造を描きながら、物語が進行していくことが、四次元空間上に表現された。同様に、夏目漱石の「坊ちゃん」と「三四郎」では、章の長さの違いはあるが、話法については二つの作品が同じ傾向にあることが判明した[3]。これは、作品の解析結果を可視化することで、夏目漱石の思考パターンと内面の揺れが明らかにされたことを意味する。

3. 「銀河鉄道の夜」の解析

3.1. キーワードの選択と計測

「銀河鉄道の夜」(使用テキストは『宮沢賢治全集』7巻 筑摩書房 1985年12月4日)は、主人公の少年ジョバンニが夢の世界で友人のカムパネルラと銀河鉄道に乗って旅を続け、突然カムパネルラが目の前からいなくなり、目が覚めると、現実の世界ではカムパネルラが事故で死んでいたというストーリーである。

本報では、作品の章を経時変化の区切りと捉える点は、既往研究と同様とした。しかし、「話法」という観点からではなく、文章中の動詞に係る「ジョバンニ」「カムパネルラ」と、「二という数」の三つの言葉をキーワードに選択することによって、さらに客観的な結果を導くことを目的としている。ここで、「二」という数字をキーワードにしたのは、少年二人の物語であることから、作品の主題の伏線としてこの数が効果的に使われていると思われるからである(例えば「二つの釘」「二本の電信柱」など)。これらのキーワードの頻出を章ごとに計測し、ウェーブレット変換法を用いてソーテングし、可視化する。

3.2. 解析手順

解析手順は、以下の4プロセスである。

作品中に現れるキーワードを時系列(章ごと)に整理する。

三次元空間座標へ変換する。(グラムシュミットの方法)

抽出の偏りを正規化で取り除く。

ばらつきをウェーブレット多重解像度解析で低減化する。

なお、「銀河鉄道の夜」は、9章から成るが、終章である9章が他の章と比較して長い、このため、本報では節の区切りを加え、12章として計測した。

3.3. 解析結果

表1に作品中のキーワードを分類した頻出度を示す。

表 1. Number of the key words

章	ジョバンニ	カムパネルラ	数字の「二」	場面設定
1	8	2	1	地上
2	9	1	0	地上
3	5	0	1	地上
4	6	0	0	地上
5	6	0	0	天上
6	8	6	1	天上
7	6	7	15	天上
8	10	6	9	天上
9	32	17	12	天上
10	2	2	0	天上
11	1	0	2	天上
12	16	0	1	地上

表1の結果を正規化して図1に示す。

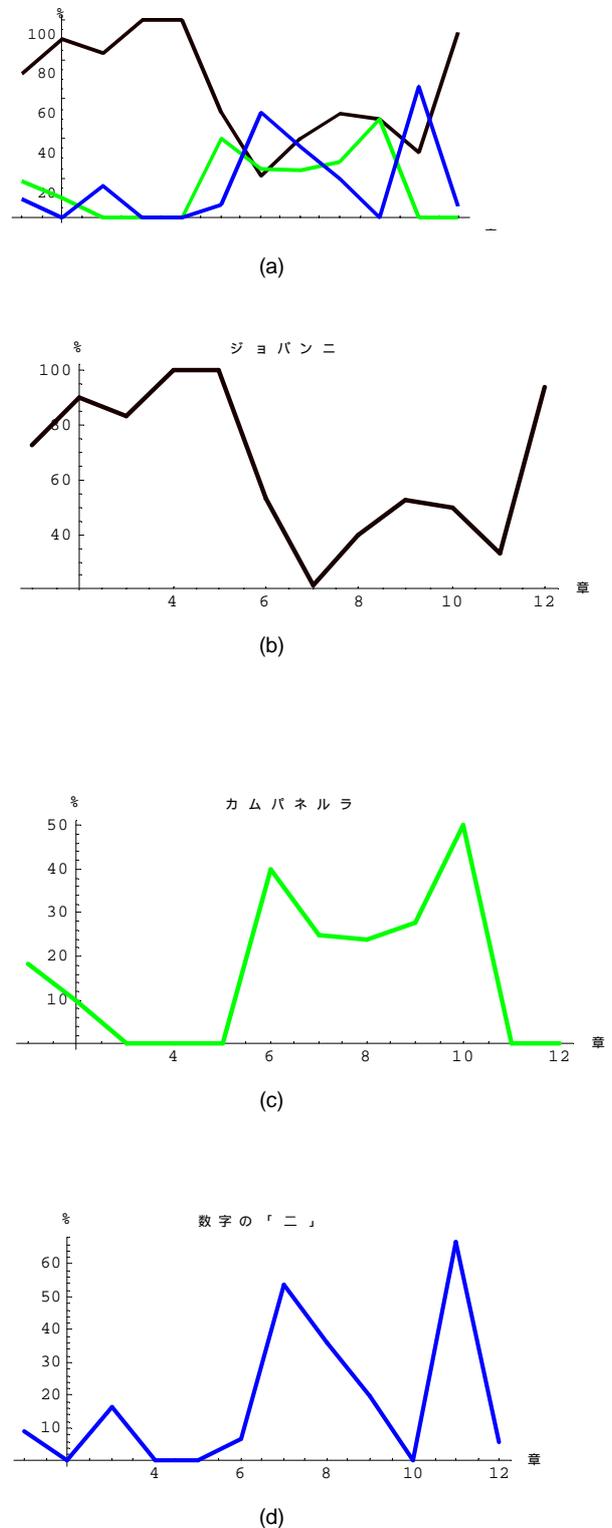
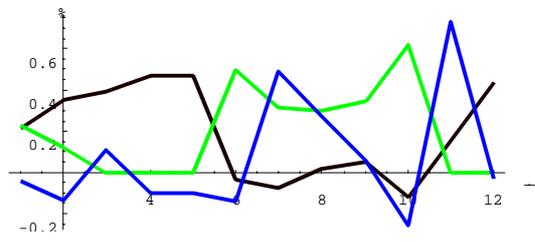
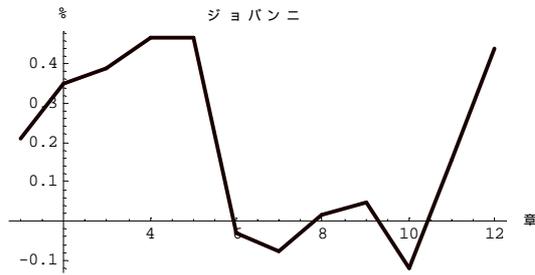


Fig. 1. Normalized number of the key words, Red:Keyword 1, Green:Key word 2, Blue:Key word 3

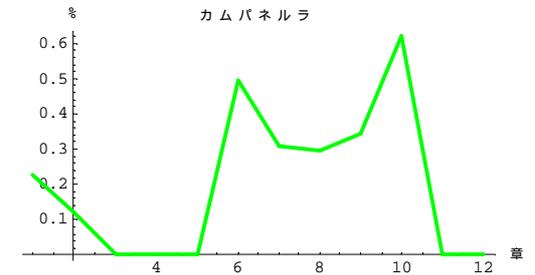
同様に、ウェーブレット変換の多重解像度解析を用いて高次変化分を削除した結果を図2に示す。



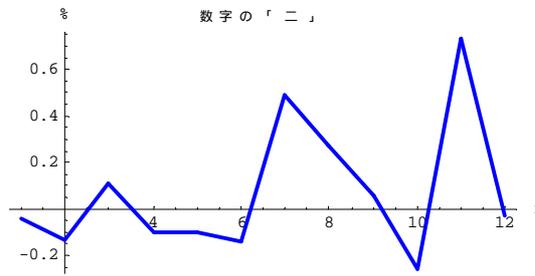
(a)



(b)



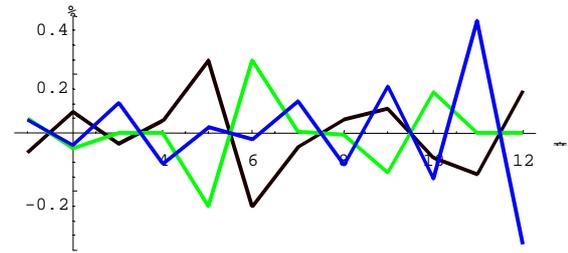
(c)



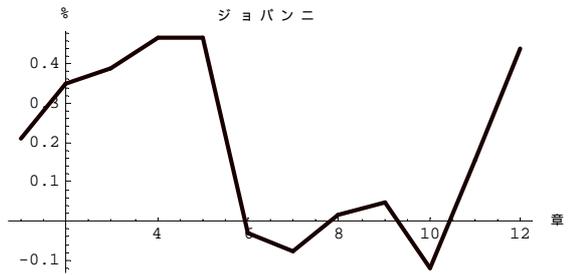
(d)

Fig. 2. Orthogonalized number of the key words,
Red:Key word 1, Green:Key word 2, Blue:Key word 3

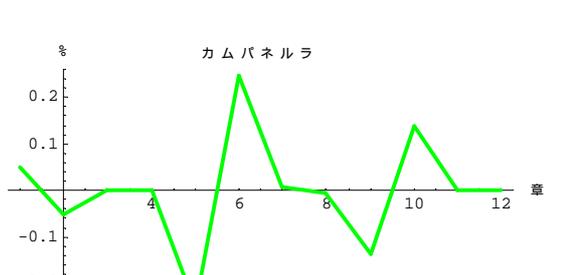
グラム・シュミットの直交化過程で削除された成分を図3に示す。



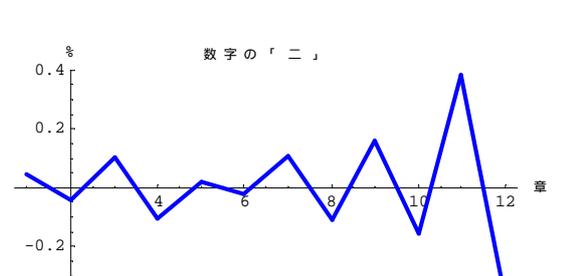
(a)



(b)



(c)



(d)

Fig.3. Deleted components in the orthogonalized process, Red: Key word 1, Green: Key word 2, Blue: Key word 3

図1,2の結果から、ジョバンニの行動は章の前半で積極的であり、中盤で受動的になり、後半で再び積極的になることがわかる。他方、カムパネルラの行動はジョバンニに相補的である。さらに、キーワード数字の「二」は「カムパネルラ」とほぼ同じ傾向だが、物語が盛り上がる中盤と、後半に効果的に使われている。

4. 文学的観点からの考察とまとめ

作者は、登場人物の二人の少年に貧富・片親の不在・クラスの人気者といじめられっ子というように、ジョバンニにマイナスの要素を、カムパネルラにプラスの要素といった正反対の設定を与えている。この解析では、行動の頻度においても、二人に両極性（アンビバレント性）を与えて作品を構成していたことがわかった。

ジョバンニが活発に活動する時は、カムパネルラの行動は静かで、カムパネルラが活発に行動する時には、ジョバンニの行動はその逆である。二人の少年の行動がおとなしい時は、「二」という言葉が代わりに効果的に多用される。特に「二」は、物語後半の山場である9章で最も高いレベルを示している。

また、この解析で正規直交化プロセスによって削減された図3に示す成分は、作者のこだわりであったかもしれないので、軽視できないのではないかという問題点が指摘できる。

参考文献

- [1] 斎藤兆古 著、『ウェーブレット変換の基礎と応用-Mathematicaで学ぶ』（朝倉書店）1998年4月10日
- [2] 岩崎晴美・斎藤兆古・宮沢賢治・堀井清之・金田知子「文学作品の線形空間論による解析」（『法政大学計算科学研究センター研究報告』11巻 1998年
- [3] 岩崎晴美・斎藤兆古・宮沢賢治・堀井清之「「坊ちゃん」と「三四郎」の線形空間論による文体解析」（『法政大学計算科学研究センター研究報告』12巻 1999年
- [4] 宮沢健太郎「文体から賢治を読む 文体論からイメージ文学へ」『日本語学』16号（明治書院）1997年9月