

後期太宰作品は本当に暗いのか！？

一単語感情極性対応表を用いた分析

江戸川学園取手高校 2年 清田葉月 指導教員 苅部達也先生



序論

なぜこの研究をしたのか？→世の中で言われている後期太宰作品が作風が暗いのは本当なのか？

情報技術 (テキストマイニングと単語感情極性対応表) を用いて明らかにしたい！！

太宰治の作風の変遷→前期、中期、後期で作風が異なると言われている (最新国語便覧 2019)

前期&後期→絶望、破滅的で不安定、中期→安定

これらの作風の変遷は太宰自身の私生活の出来事や、時代背景、彼の思想的背景に着目し、作品と照らし合わせながら分析されている→各研究者の印象なのではないか！？そこで、**情報学の手法を用いて数値で客観的に明らかにする**

意義

太宰作品の作風を客観的に深く理解ができる。手法を確立することで様々な文学作品を客観的に分析することに繋がる。

先行研究

- 助詞の使われ方に着目し、前期太宰作品を分析し、『めくら草子』以降作風が異なることを明らかにした(尾城, 金 2017)
 - テキストマイニングの手法を用いて、頻出語により夏目漱石の『こころ』を分析し、頻出語の抽出とマッピングによって物語の流れを明らかにしている。(樋口, 2003)
 - 頻出語、感情分析を用いて、『源氏物語』を分析し、感情分析の章で1巻、2巻のみよりマイナスになっていることを明らかにしている。(Wu, Bai 2019)
- テキストマイニングの手法と感情分析を組み合わせることによってより深い分析ができるのではないか！**

研究対象と方法

対象作品

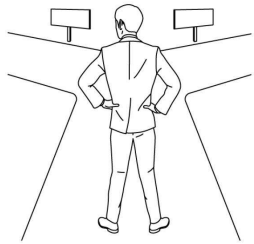
選出方法 作品数が多く、変遷が「下降指向」と呼ばれている中期、後期作品から、新潮社の発行部数ランキングの各時期上位5作品を対象とした。

STEP1 テキストマイニング(Khcoder)を用いて中期 『津軽』『走れメロス』『御伽草子』『惜別』『パンドラの匣』後期『斜陽』『ヴィヨンの妻』『人間失格』『グッド・バイ』『ろまん燈籠』

いて各作品ごとの形容詞、動詞、形容動詞の頻出語を抽出する。(慣用表現による誤差を減らすため)

STEP2 単語感情極性対応表(東工大,高村)を用いて**上位10語の感情極性**を出す。

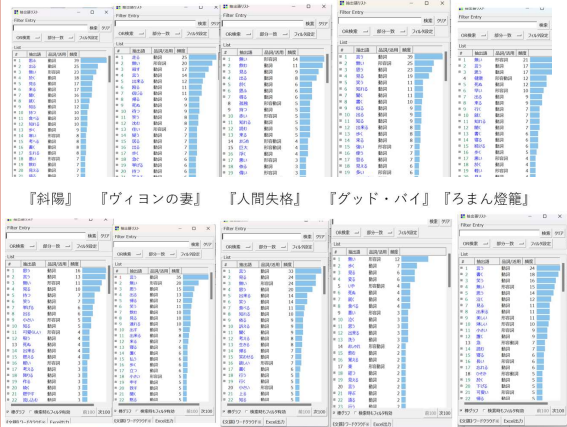
STEP3 各作品ごとの10語の極性の平均値を出し、**中期、後期それぞれの平均を比較**する。



結果

STEP 1

『津軽』『走れメロス』『御伽草子』『惜別』『パンドラの匣』



結果

感情極性を見ると、**中期の平均は -0.5980、後期の平均は -0.7206**となり、**後期のほうが感情値が低い結果となった**→後期の方がマイナスの感情値を持つ単語が多く用いられている。

考察

中期と後期を比較すると後期の単語極性の方がマイナスの数値が高い→後期の作品の方がマイナスの感情極性をもつ表現が多く使われている。したがって、本探求は**既存で言われている中期より後期の方が暗いという作風の違いに関して、情報学の手法を用いても同様の傾向がある**、ということが分かった。

限界と展望

●太宰作品全てを対象として分析出来ていないため、今後は全作品を対象として分析を進めていきたい。

●既存の単語極性対応表に掲載のない単語がみられたため、**多様な表現をする文学作品や80年近く前の作品にも対応できるようなより広範囲な日本語表現に対応した単語極性対応表がない限り、この手法の限界を感じた。**

●「ない」の使い方(間違いはない、と自信がない、が同様に扱われている『斜陽』)により誤差がある可能性があるため、同じくより精度の高い単語極性対応表の開発の必要性を感じた。

●活用の方法として、小説を脚本などに書き換えるなど**異なる文章で同じ作品を表現する際の原作との乖離を防ぐ**ことができるのではないかと感じた。

参考文献 『浜島書店編集部 巻頭増補版 最新国語便覧』 浜島書店 2019 総ページ498 254、255ページ参照

『太宰治』奥野健男文藝文集 第3刷 1973 総ページ373

『太宰治大辞典』志村有弘、渡部芳紀 勉誠出版 2005

『計量文学の射程』村上征勝、金明哲、土山玄、上阪彩香 初版勉誠出版 2016 総ページ249

『各創作時期における太宰治の作風の変遷』匂 健龍 2024 国際日本研究 Vol.16 pp.55-70

『小説の可視化について——与謝野晶子の『源氏物語』の訳本を例として』Wu, B. & Bai, X. 2019 arXiv.org, arXiv:1912.01068v1

『コンピュータ・コーディングの実践: 漱石『こころ』を用いたチュートリアル』樋口耕一 2003 年報人間科学 Vol.24 No.2 pp193-214

『学生によるテキストマイニングの実践』星井進介 長岡工業高等専門学校研究紀要 Vol.59 pp.26-35 2024

『太宰治の前期作品における文体変化の定量的分析』尾城奈緒子、金明哲 2017 日本行動計量学会大会抄録集 Vol.45 pp.168-171

謝辞 単語感情極性対応表(東工大,高村) http://r.pi.titech.ac.jp/~takamura/pubs/pn_ja_dic, KhCoder (樋口耕一) KH Coder: 計量テキスト分析・テキストマイニングのためのソフトウェアを分析に使用させていただきました。株式会社トモノカイから、ポスターの掲載内容に関するフィードバックをいただきました。この場を借りてお礼を申し上げます。

STEP 2 全体的にマイナスに偏りがあった。

作品名	ジャンル	全単語数	1	2	3	4	5	6	7	8	9	10
津軽	1437	1	2	3	4	5	6	7	8	9	10	
走れメロス	2302	1	2	3	4	5	6	7	8	9	10	
御伽草子	2652	1	2	3	4	5	6	7	8	9	10	
惜別	3103	1	2	3	4	5	6	7	8	9	10	
パンドラの匣	2710	1	2	3	4	5	6	7	8	9	10	
斜陽	1437	1	2	3	4	5	6	7	8	9	10	
ヴィヨンの妻	2652	1	2	3	4	5	6	7	8	9	10	
人間失格	3103	1	2	3	4	5	6	7	8	9	10	
グッド・バイ	2710	1	2	3	4	5	6	7	8	9	10	
ろまん燈籠	1437	1	2	3	4	5	6	7	8	9	10	

STEP 3

単語感情極性対応表に掲載のない語は解析できていない

中期の平均は -0.5980、後期の平均は -0.7206