

オンライン小説はどれくらい文学的なのか —一品詞の類似度に注目した予備研究—



福本わかば, 星野雄介 (武蔵野大学 経営学部)

イントロダクション

研究の背景:

- 現代の日本の小説は、近代文学（明治）から始まっている
- オンライン小説は21世紀以降人気になった新しい分野
- 多くの作家は過去の作品に影響を受けている

⇒近代文学とオンライン小説には**異なっている部分と似ている部分**があるはずである

Research Question:

オンライン小説と近代文学は、どこが・どれくらい類似しているのか？

方法論

データ:

- 近代文学の作家7名（133作品）:
 - 青空文庫より
 - Wikipediaより、明治時代に活躍した作家を選出
- オンライン小説23名（23作品）:
 - 異世界転生: 9名、悪役令嬢7名、追放: 5名、その他: 2名
 - 「小説家になろう」「カクヨム」「Arcadia」より
 - いずれもメディアミックスされた作品から選出

形態素解析: Janomeを利用

- 強制抽出・ストップワードは設定なし
- ノイズは時代によって異なるため
- ルビは削除済み

分析: ジャンル間の類似度分析

- 名詞・動詞・副詞・形容詞に絞り、全単語に占める近代文学とオンライン小説の「* 詞平均出現率」を算出
- 文学作品の出現率順に50分割して、分割された単語ごとの近代文学とオンライン小説間のコサイン類似度を集計

作家リスト:

作家名	作品数	文字数	生まれ年 (元号)	生まれ年 (西暦)
森鴎外	16	207,088	文久2年	1862
夏目漱石	19	516,796	慶応3年	1867
幸田露伴	23	417,132	慶応3年	1867
国木田独步	9	86,082	明治4年	1871
田山花袋	7	281,798	明治4年	1872
島崎藤村	16	879,288	明治5年	1872
泉鏡花	24	615,181	明治6年	1873

作家名	ジャンル	作品数	文字数
アネコユサギ	異世界転生	1	101,270
カルロ・ゼン	異世界転生	1	106,868
遠沢天介	異世界転生	1	100,791
江口浩	異世界転生	1	101,295
香月美英	異世界転生	1	103,468
銀色獅/長月達平	異世界転生	1	103,825
伏瀬	異世界転生	1	102,141
理不尽な孫の手	異世界転生	1	101,921
櫻乃ままれ	異世界転生	1	166,229
いのり	悪役令嬢	1	101,895
雨川透子	悪役令嬢	1	101,030

分析イメージ:

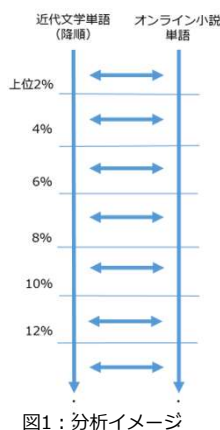


図1: 分析イメージ

先行研究

オンライン小説研究:

頻出単語・流行キーワード・出現単語をベースとしたクラスタリング、推薦システム等の既存研究がある[1][2][3]

文体推移研究: 過去100年の文学作品380編について分析した結果

- 単語レベルの変化は速い
- 構文レベルの変化は緩やか[4]

しかし、オンライン小説を文学史に位置付けて、計量分析した研究はほとんど見当たらない

注: 参考資料リストは次ページ

結果

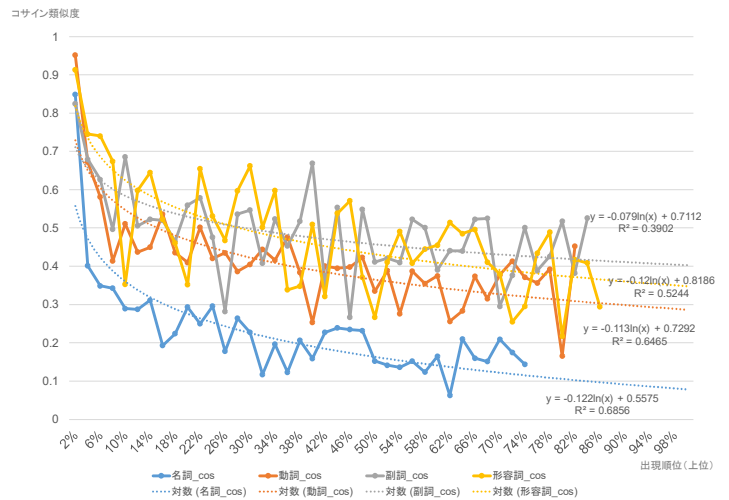
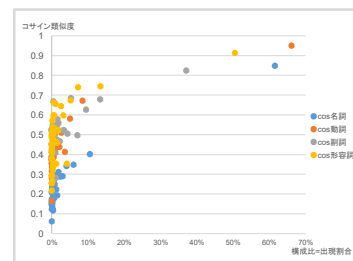


図2: 出現順位ごとの類似度

- トップ2%の類似度は高い
- その後、名詞の類似度の減少は急激だが、動詞・副詞・形容詞の類似度は似たような減少傾向

しかし、出現率の高い単語（≒順位の高い単語）が、実際にどれくらい占めているかが重要



4品詞ともに出現率トップ2%の単語が、品詞ごと総出現数の37%(副詞)から66%(動詞)を占める

⇒構成比の高い単語は近代文学とオンライン小説でかなり類似度が高い
= 構成比の低い単語は類似度が低くなる傾向

図3: 品詞に占める割合と類似度

考察

似ている:

- 動詞・副詞・形容詞は比較的似ている
- 品詞を問わず、出現率の高い単語は似ている

異なっている:

- 名詞は大きく変化する
- 品詞を問わず、出現率の低い単語は変化しやすい

貢献:

単語が変化しやすいという既存研究に対し、どのような単語が変化しやすいのか/しにくいのか、を明らかにした

今後の課題:

- 様々な時代の作品を分析する
- 分析対象のジャンルをそろえる
- 既存研究を参照した多様な分析手法を用いる

参考資料：

- [1] 堺雄之介, & 伊東栄典. (2020). オンライン小説の流行語抽出. *情報処理学会全国大会講演論文集*, 82, 6T-01.
- [2] 本田優也, & 伊東栄典. (2015, September). 小説投稿サイトの小説におけるキーワードの多様性分析. In *電気関係学会九州支部連合大会講演論文集 平成 27 年度電気・情報関係学会九州支部連合大会 (第 68 回連合大会) 講演論文集* (pp. 167-167). 電気・情報関係学会九州支部連合大会委員会.
- [3] 小坂直輝, 小林哲則, 林良彦, 早稲田大学, & 理工学術院. (2019). 隠れた良作を推薦可能な Web 小説レコメンドシステムの提案. *人工知能学会インタラクティブ情報アクセスと可視化マイニング研究会 (第 23 回) 研究発表予稿集*, 1-8.
- [4] 李広微, & 金明哲. (2019, September). CA7-1 過去百年間における小説の文体変容についての定量分析. In *日本行動計量学会大会抄録集 47* (pp. 296-297). 日本行動計量学会.