

令和 2 年 5 月 29 日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2017～2019

課題番号：17K12799

研究課題名(和文)文章の数量分析に基づく西鶴質疑本の著者及び成立年に関する総合的研究

研究課題名(英文) Authorship attribution and quantitative analysis of Saikaku's works

研究代表者

上阪 彩香 (Uesaka, Ayaka)

大阪大学・数理・データ科学教育研究センター・特任助教(常勤)

研究者番号：60780252

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：本研究課題では、日本文学に影響を与えた重要な古典籍として位置づけられている江戸時代前期の俳諧師・作家である井原西鶴(1642～93)の浮世草子作品における種々の著者への疑問に着目し、西鶴と弟子の北条団水(1663～1711)、同時代の浮世草子作家の文章を統計的に分析することで、数理的観点から著者の「特徴・クセ」を把握することに主眼を置き、研究に取り組んだ。

研究成果の学術的意義や社会的意義

研究期間中には、『好色三代男』、『椀久二世の物語』、『眞實伊勢物語』、江島其磧『大尽三ツ盃』、西村市郎右衛門『新撰咄揃』、『小夜衣』のデジタル化を行った。これらのデータとこれまで構築してきた浮世草子作品のデータベースを用いて様々な観点から数理的分析を行い、西鶴の遺稿集4作品に関しては、編集・出版を行った弟子の団水が手を加えた範囲は限定され、文章は西鶴の文章である可能性が高いことを明らかにした。これら二者間の比較に加えて、複数著者間における文章の「特徴・クセ」の把握を試み、各々に文章の「特徴・クセ」が存在することを明らかにし、これらの研究成果を国内外での学会や書籍等で公表した。

研究成果の概要(英文)：This study aims to exploration of the seventeenth century Japanese authors', Saikaku Ihara (c.1642～93), his student Dansui Hojo (1663～1711) and writers of the same period, writing style from a quantitative point of view.

研究分野：計量文献学、テキストマイニング、統計科学、デジタルヒューマニティーズ

キーワード：井原西鶴 統計科学 多変量解析 近世文学作品

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

西鶴は元禄の文豪と評価されており、我が国の文学史における重要性から多くの国文学者によってその思想や記述内容、あるいは成立に関する歴史的考証が続けられてきたが、西鶴が著したのは『好色一代男』のみで他の浮世草子は門弟の北条団水の作、あるいは西鶴と団水の共同執筆であるとする説(森、1955)や、2年間に西鶴の著作が10作品(258,377語)出版されており、この分量を一人で作成するのは難しいことから西鶴には複数の助作者が存在したという西鶴工房説(中村、1969)がある。さらには、弟子の団水によって編集・出版された遺稿集『西鶴置土産』『西鶴織留』『万の文反古』『西鶴俗つれづれ』『西鶴名残の友』にも、西鶴と団水どちらが実質的な著者であるのかという疑問が指摘されてきた(山口、1929)。

申請者は西鶴研究者及び団水研究者による単語の認定、品詞付けに基づいたデータベースを構築した上で、文章の数理的・統計的分析の側面から取り組み、主成分分析をはじめとする多変量解析を西鶴作品に適用し、その特徴の抽出を行うことで西鶴浮世草子の著者問題の究明を目指してきた。

2. 研究の目的

本研究課題ではこれまでに構築したデータベースと本研究課題にて構築したデータベースを用いて、西鶴浮世草子に提出されている著者における種々の疑問に対し、これまで使用してきた教師なし学習に加え、教師あり学習などの数理的手法を用いて検討することで、西鶴遺稿集における弟子の補作・偽作に関する疑問、西鶴工房説に関する疑問、複数著者間における著者の「特徴・クセ」の把握を試みた。

3. 研究の方法

本研究課題では、これまでの検討内容に加え、複数著者間における文章の「特徴・クセ」の把握を試みた。

- (1) 複数著者間において、著者の文章の「特徴・クセ」の抽出が可能であるかを検討するため、データベースへの作品の追加を行った。
- (2) テキストデータを分析に用いるためには、文章を単語に分割し、品詞情報など分析に必要な情報を付与が必要であるため、この分割方法について検討した。
- (3) 西鶴作品のなかで、多作期に描かれている作品に着目し、その前後における文章の特徴の変化について検討した。
- (4) 遺稿集における疑問に関しては、これまで教師なし学習を用いて巻単位での検討を行ってきたが、国文学における先行研究では章単位での検討が行われている。そのため、本研究課題では章を対象として、教師あり学習を適応し、検討を加えた。
- (5) 二者間の比較に留まらず、複数著者間の文章の「特徴・クセ」の把握に取り組んだ。

4. 研究成果

(1) データベースの構築

本研究期間内に『好色三代男』、『椀久二世の物語』、『眞實伊勢物語』、江島其磧『大尽三ツ盃』、西村市郎右衛門『新撰咄揃』、『小夜衣』のデジタル化作業が完了した。

(2) 辞書の検討

これまでの研究では、近世文学研究者の協力を得て作成された形態素解析済みデータを用いて検討を行ってきた。現在では多くの辞書が開発され、さらに古典文に関する辞書には、古文用UniDicSが公開されている。近世浮世草子作品に最も近いものに、洒落本・人情本テキストを主な対象とした近世口語(洒落本)UniDicがあげられる。これらの異なった基準によって文章の分割を行った場合に、どのような分析結果が得られるのかを考察するため、近世文学研究者の協力を得て作成された形態素解析済みデータ、mecab of 0.996を使用しIPA辞書を用いた形態素解析の結果、近世口語(洒落本)UniDicを用いた形態素解析の結果をもとに、著者判別という側面から検討した。

(3) 多作期における文章の変化

西鶴によって1682~1686年に出版された『好色一代男』、『諸艶大鑑』、『西鶴諸国はなし』、『椀久一世の物語』、『好色五人女』、『好色一代女』の6作品と、神保(1990)が代作者・助作者の作品が混入されている可能性を誰も持つ作品として挙げた1686~1688年に出版された『本朝二十不孝』、『男色大鑑』、『懐硯』、『武道伝来記』、『日本永代蔵』、『武家義理物語』、『嵐は無常物語』、『色里三所世帯』、『好色盛衰記』、『新可笑記』の10作品とを主成分分析を用い、比較検討し、1682~1686年に出版された6作品と1686~1688年に出版された10作品には、異なった文章の特徴があることを示した。また、西鶴、団水、『嵐は無常物語』を主成分分析、クラスター分析を用いて比較検討し、『嵐は無常物語』の文章の特徴が西鶴と団水のどちらとも類似していないことを示した。

(4) 遺稿集の検討

バギング法、ランダムフォレスト法、アダブースト法により西鶴、団水の文章の間に差異がみ

られた分析項目を用い、章単位での西鶴遺稿集『西鶴置土産』、『西鶴織留』、『西鶴俗つれづれ』、『西鶴名残の友』のそれぞれの文章が、西鶴と団水のどちらの文章と類似しているのかを検討した結果を公表した。

(5) 複数著者間の文章の「特徴・クセ」の把握

申請者が検討してきた二者間での文章の「特徴・クセ」の把握に加え、これまでに構築したデータベースと本研究課題にて構築したデータベースを用いて、複数著者間における文章の「特徴・クセ」の把握を試みた。西鶴、団水、西村市郎右衛門の3人の浮世草子を単語、助詞(unigram、bigram)、文字(unigram、bigram)を用いて数理的分析を行うことで、これらの3人の浮世草子作家による作品の文章の「特徴・クセ」が著者ごとに異なること、西鶴の文章の「特徴・クセ」は市郎右衛門よりも団水と近接していることを明らかにした。また、西鶴、団水、夜食時分、西沢一風の浮世草子を対象に、これら4作家における文章の特徴把握を試み、助詞の出現率を用いた主成分分析及びクラスター分析を使用して検討したところ、上記の4作家は著者ごとに文章の「特徴・クセ」が異なることを示した。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 上阪彩香	4. 巻 第45巻第2号
2. 論文標題 アンサンブル学習モデルを用いた西鶴遺稿集の著者に関する検討	5. 発行年 2018年
3. 雑誌名 行動計量学	6. 最初と最後の頁 pp.135 ~ 151
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.2333/jbhmk.45.135	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 上阪彩香	4. 巻 424
2. 論文標題 西鶴浮世草子の文章の特徴と出版時期	5. 発行年 2019年
3. 雑誌名 統計数理研究所共同研究リポート	6. 最初と最後の頁 pp.37 ~ 48
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計12件（うち招待講演 2件 / うち国際学会 6件）

1. 発表者名 Ayaka Uesaka
2. 発表標題 Verifying the Authorship of Saikaku Ihara 's Arashi ha Mujiyo Monogatari Using a Quantitative Approach
3. 学会等名 Japanese Association for Digital Humanities2018 (国際学会)
4. 発表年 2018年

1. 発表者名 上阪彩香
2. 発表標題 著者判別分析における形態素解析辞書選択
3. 学会等名 日本行動計量学会
4. 発表年 2018年

1. 発表者名 上阪彩香
2. 発表標題 数量分析を用いた近世浮世草子における著者判別の試み
3. 学会等名 Sixth International Conference at Feng Chia University Exploring the Uncharted Territories: Foreign Languages and Literature Teaching in Taiwan and Southeast Asia (国際学会)
4. 発表年 2018年

1. 発表者名 Ayaka Uesaka
2. 発表標題 Authorship analysis Ihara Saikaku and his environment
3. 学会等名 Digitale Japanologie Frankfurt Digital Approaches: Episode III (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Ayaka Uesaka
2. 発表標題 Verifying the Authorship of Saikaku Ihara's Arashi ha Mujyo Mogatari in Early Modern Japanese Literature: A Quantitative Approach
3. 学会等名 Digital Humanities2017 (国際学会)
4. 発表年 2017年

1. 発表者名 上阪彩香
2. 発表標題 KH Coderチュートリアル
3. 学会等名 人文科学とコンピュータシンポジウム
4. 発表年 2017年

1. 発表者名 上阪彩香
2. 発表標題 テキストマイニングを用いた近世浮世草子の検討
3. 学会等名 東西学術研究所 第23回研究例会 言語接触研究班 KU-ORCAS ユニット1 研究集会 第2部 デジタルアーカイブと言語分析 (招待講演)
4. 発表年 2018年

1. 発表者名 上阪彩香
2. 発表標題 『万の文反古』における版下の系列と文章の特徴
3. 学会等名 言語研究と統計2018
4. 発表年 2018年

1. 発表者名 Bor Hodoscek and Ayaka Uesaka
2. 発表標題 Issues in high- and lowbrow classification of Japanese literature
3. 学会等名 Osaka-Wurzburg Collaborative Workshop: Cross-Linguistic Perspectives on Lexical Diversity in Literary Texts (国際学会)
4. 発表年 2018年

1. 発表者名 Ayaka Uesaka
2. 発表標題 Exploration of the Seventeenth Century Japanese Authors' Writing Style Using a Quantitative Approach
3. 学会等名 Digital Humanities2019: Conference Abstracts. Utrecht: Utrecht University (国際学会)
4. 発表年 2019年

1. 発表者名 田畑智司、上阪彩香、岡部未希
2. 発表標題 デジタルヒューマニティーズと近代日本文学：トピックモデリングによる遠読
3. 学会等名 第4回大阪大学豊中地区研究交流会
4. 発表年 2019年

1. 発表者名 上阪彩香
2. 発表標題 テキストマイニングを用いた近世浮世草子における著者判別研究
3. 学会等名 DuEX若手研究交流会
4. 発表年 2020年

〔図書〕 計2件

1. 著者名 上阪彩香	4. 発行年 2019年
2. 出版社 Department of Foreign Languages and Literature, Feng Chia University	5. 総ページ数 21
3. 書名 「数量分析を用いた近世浮世草子における4作家を対象とした著者判別の試み」, "Innovative Foreign Language Education and the New South-Bound Policy"	

1. 著者名 上阪彩香	4. 発行年 2019年
2. 出版社 勉誠出版	5. 総ページ数 3
3. 書名 「西鶴遺稿集の著者問題」 『文化情報学事典』	

〔産業財産権〕

〔その他〕

HP
<https://ayakau26.wixsite.com/auesaka>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----