

令和 2 年 6 月 5 日現在

機関番号：17102

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K02452

研究課題名(和文) 初期近代英国演劇における「合作」に関する研究 印象批評とコンピューター解析

研究課題名(英文) A Study of "Collaboration" in Early Modern English Drama

研究代表者

太田 一昭 (OTA, Kazuaki)

九州大学・言語文化研究院・学術研究者

研究者番号：10123803

交付決定額(研究期間全体)：(直接経費) 1,500,000円

研究成果の概要(和文)：近年コンピューターを用いた計量文体学的(言語データの統計学的)作者同定手法の導入によって、従来シェイクスピアの共作の可能性が指摘されていた多くの作品について、シェイクスピアの関与の可能性がより確信的に提唱されるようになってきている。本研究は、シェイクスピアの合作に関する近年の作者同定研究を批判的に検証した。計量文体学的作者同定研究は少なくとも現時点では、従来の印象批評による作者判定の代替にはなりえていない。近年の計量文体学的な作者同定研究の多くは、言語データの抽出・収集および処理・分析が恣意的になされていて、その判定結果に信を置くことができない。

研究成果の学術的意義や社会的意義

コンピューターを利用する計量文体学的な作者同定研究を理解するには、コンピューターのプログラムと統計学の知識が多少とも必要である。一般的な英文学研究者はそのような知識を持ち合わせていないから、近年の作者同定研究の成果を適切に評価することは困難である。The New Oxford Shakespeare (2016)は従来執筆者が確定されていない戯曲の作者判定を積極的に推進する編纂本であるが、编者たちの作者判定を支えているのが計量文体学的作者同定研究の成果である。その作者同定研究の可能性と限界とを具体的に提示する本研究は、計量文体学に不案内な文学研究者の作者同定研究の現状理解に裨益する。

研究成果の概要(英文)：In recent years, the results of stylometric authorship identification studies using computers have led to more confident suggestions of Shakespeare's possible involvement in many dramatic works that have previously been considered as possibly written at least in part by him. This study critically examined recent Shakespeare attribution studies, arguing that computer-based stylometric authorship attribution has yet to replace conventional authorship identification based on impressionistic readings of dramatic texts. This is because most of the recent stylometric attribution investigations have been arbitrary and statistically unjustified in their extraction, collection, processing, and analysis of linguistic data, and the results cannot wholly be relied upon.

研究分野：English Literature

キーワード：authorship attribution stylometry collaboration Shakespeare Edward III Double Falsehood Arden of Faversham New Oxford Shakespeare

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。

## 1. 研究開始当初の背景

### (1) シェイクスピアの合作（共作）についての関心の高まり

近年シェイクスピアの合作（collaboration）に対する関心が高まっている。*Shakespeare Survey*, vol.67（2014）は、合作に関する特集号である。World Shakespeare Congress 2016のセミナーの一つは、“Shakespeare, Collaboration, and Co-Creation”であった。シェイクスピアが共同で作品を執筆したというのはもちろん、新説ではない。*Henry VIII*は19世紀以来、ジョン・フレッチャーとの合作だとされてきたし、*Pericles*は、前半はジョージ・ウィルキンズによって書かれたものだという説が提唱されてきた。近年では*Titus Andronicus*はジョージ・ピールとの合作であり、*Timon of Athens*はトマス・ミドルトンとの合作だと主張されている。さらには*Henry VI*第1部もまた合作だという説が従来にもまして強く主張されるようになってきている。四大悲劇の一つ*Macbeth*にはミドルトンの手が入っているというのが通説だが、第1・二折本に収められた*Macbeth*のテキストは、ミドルトンによる「翻案」だと主張する学者もいる。*The Two Noble Kinsmen*にはシェイクスピアの筆が入っていると従来指摘されてきたが、*Edward III*もまたシェイクスピアの合作だという説が近年有力になっている。他に*The Spanish Tragedy*、*Arden of Faversham*、*Cardenio*（*Double Falsehood*）の一部にもシェイクスピアの手が入っているとする研究者が増えている。*Cardenio*は失われた作品であるが、1728年、Lewis Theobaldは、その翻案との触込みで*Double Falsehood*を出版した。この劇は、「アーデン・シェイクスピア」第3シリーズの一冊として2010年に出版された。

### (2) 近年の合作説の特徴、コンピューターによる言語データ解析の進展

20世紀の後半以降、コンピューターによる言語資料の解析技術が飛躍的に発展した。近年の初期近代演劇の合作研究にも、コンピューターを利用した統計学的作者同定の手法が導入されている。コンピューターによる解析は印象批評によって指摘されていた結論と大差ないものも少なくないが、コンピューターによる解析結果に後押しされて、従来シェイクスピアの関与の可能性が指摘されていた多くの作品について、シェイクスピアの合作の可能性がより確信的に提唱されるようになってきている。前述の*Edward III*の「伯爵夫人の場」の作者はシェイクスピアだというのは近年定説化しつつあると言ってよいが、定説化に与って力あったのがコンピューターによる統計学的作者同定研究の進展である。

## 2. 研究の目的

作者同定研究におけるコンピューターによるデータ解析は今日の作者同定研究において、強い発言力をもつようになってきている。研究者の中には、コンピューターによる分析結果をあたかも錦の御旗であるかのように、自説を補強し、権威づけるための根拠として無批判に引用する人もいる。このような姿勢は、「科学的」な態度とは言えない。しかし本研究を含めて文学研究者は一般的にコンピューター解析技術に不案内であり、その解析結果をともしれば、無批判に受け入れがちである。これは好ましい状況ではない。コンピューター解析は作者同定研究において強力な武器になりえるにしても、その妥当性を検証することが必要である。本研究の目的は、シェイクスピアの合作とされる劇作品について、印象批評による作者同定研究と近年のコンピューターのデータ解析による作者同定研究の諸問題とを検証し、シェイクスピアの合作に対して初

期近代英国演劇研究者はどのようにアプローチするのが妥当であるかを示すことである。

### 3. 研究の方法

近年のコンピューターによる計量文体学的(統計学的)解析による作者同定研究を多角的に検証した。本研究では English Drama、Early English Books Online の既存のデータベースに加えて、本研究代表者が編集・構築した 2 種類のデータベースを解析に用いた。1 つは English Drama から抽出した 787 編の(原綴りの)戯曲データベースであり、他は Shakespeare His Contemporaries および Folger Digital Texts のウェブサイトからダウンロードした 523 編の(現代綴りの)データベースである。さらに Pervez Rizvi の初期近代英国演劇連語データベースと作者同定プログラムの有用性を確認し、これを積極的に活用して統計学的文体解析による作者同定研究の諸問題を論究した。

### 4. 研究成果

本研究期間中に発表あるいは執筆した論考(全て単著)は次の通りである。口頭発表の後刊行した論文については、刊行論文の概要のみを記す。

#### (1) “Edward III and Statistical Attribution Studies”(既刊論文)

*Edward III* にシェイクスピアの筆致を感知する学者は多い。*Edward III* の執筆にシェイクスピアが関与しているという所見は近年の作者同定研究によって追認・補強され、今日ではほとんど定説化している。この作者同定研究の趨勢をもたらしたのが、コンピューターによる計量文体解析である。本稿では Thomas Merriam ほか従来の文体統計解析による、*Edward III* の作者同定研究を個別に検証し、近年の計量文体解析による作者同定研究の優れた例として Hugh Craig と Arthur F. Kinney 共編の *Shakespeare, Computers, and the Mystery of Authorship* (2009) の解析方法と本書所収の Timothy Irish Watt の作者同定論文を紹介した。Watt は Zeta テスト(内容語分析) 主成分分析(機能語分析) 希少語分析の総合判定によって、*Edward III* は複数の作家による共作であり、伯爵夫人の場合はシェイクスピアが執筆した可能性が大きいと結論している。

Craig、Watt らのコンピューター解析による作者判定に不安がないわけではない。彼らは膨大な言語データを蓄積しているはずだが、解析の最終結果を提示するだけで、その解析のために収集・蓄積した生データを公開していない。少なくとも内容語や機能語の一部についても、セグメントごとの出現頻度を提示すべきではないだろうか。標準化したデータ、あるいは最終解析数値だけを提示するだけでは不十分である。生データ(たとえば機能語のセグメントごとの粗頻度、出現数)を開示すべきであろう。そのデータがあれば、標本選定の適切・不適切を第三者が判断することができるし、統計量の再検証も可能である。

#### (2) 「ジョナサン・ホープの作者同定研究の功罪」(既刊論文)

Jonathan Hope の *The Authorship of Shakespeare's Plays: A Socio-Linguistic Study* (1994) は、近年の作者同定研究において最もよく引用される文献の一つである。本稿では本書の「功罪」を再検討し、次のように結論した。

*The Authorship of Shakespeare's Plays* は良書である。分析方法は明晰であり、結果の提示方法も全体として適切で、誠実である。Brian Vickers は、Hope の助動詞 do の用法に着目した

分析方法は作者識別の新しい強力なツール ( a new and powerful authorship tool ) と評価しているが、そのような好意的な評価も首肯できる。しかし作者識別ツールとして過信すべきではない。Hope の分析方法はたしかに、作者識別ツールとして十分に使える場合がある。たとえばシェイクスピアとマーロウ、あるいはシェイクスピアと( 世代が異なる、つまりシェイクスピアよりかなり若い ) フレッチャーのコーパスを比較すればかなりの精度で識別が成功すると思われる。シェイクスピアとマーロウは同世代であるが、シェイクスピアと他の同世代の劇作家( リリー、キッド、チャップマン、グリーン、ナッシュなど ) について Hope のやり方でうまく識別できるかは分からない。作品間の識別になると、より難しくなる。より小さい単位の識別、つまり特定の場の識別になると、かなり信頼度は落ちるであろう。Hope の作者同定を援用するのは結構であるが、Hope の解析には限界があることを認識しておく必要がある。

( 3 ) 「『二重の欺瞞』の作者同定と文体統計解析」( 既刊論文 )

*Double Falsehood* をシェイクスピア原作の翻案とする近年の論考の多くは、統計学的手法を用いて作者同定を試みている。本稿では、*Double Falsehood* にシェイクスピアの筆が入っているとする近年の統計学的な作者同定研究に疑義を呈し、次のように主張した。

この劇の作者同定に文体統計解析を導入しても十分な成果は期待できないだろう。というのも『二重の欺瞞』は、たとえ「原作」が存在したとしても、二重、三重に上書きされている可能性が高く、だとすれば適切な言語データを抽出することが非常に難しいからである。演劇テキストはそうでなくとも、脚本の完成から版の印刷に至るプロセスで作者以外の関係者( 劇団、植字工、印刷者その他 ) の干渉によりオリジナル原稿からそれだけ遠ざかる不安定な要素をはらんでいて、執筆者の同定はそれだけ困難である。そこにさらに後世の改作者や劇団関係者によって更なる変更が加えられているかもしれない。しかも上書きの場所も特定することができない。Craig と Kinney 共編の *Shakespeare, Computers, and the Mystery of Authorship* は、コンピューター統計解析による近年の作者同定研究のすぐれた成果の一つである。両編者が解析方法の妥当性を検証するために用いたテキストは、136 編の単独執筆戯曲( シェイクスピア戯曲 27、その他の劇作家作品 109 ) である。彼らはこの 136 作品を 1,300 のセグメント( 1 セグメント 2,000 語 ) に分け、これらを比較対照の基礎資料として、シェイクスピアの共作劇 4 編の比較的長い場( 1,500 語以上の場 ) の執筆者の識別を試みる。解析結果は従来の( 信頼できる ) 執筆者区分と一致しており、それによって彼らの解析方法が有効であることが裏付けられたのであるが、彼らが 1,500 語以上の場についてのみ執筆者を判定したのは、語数の少ない場は解析の精度が落ちると認識しているからである。*Double Falsehood* は全 14 場からなるが、1,500 語以上の場は 4 つにすぎない。これらの比較的長い( したがって、より信頼度の高い解析結果が得られるはずの ) 場にも複数の執筆者が関与しているかもしれない。だとすれば、統計解析によって妥当な結論が導き出せる可能性は絶望的に小さい。今後も統計解析による *Double Falsehood* の作者同定を試みる人は多いだろう。しかしおそらく、説得力ある結論が統計解析によって得られることはないだろうと予測する。

( 4 ) 「Pervez Rizvi の初期近代英国演劇連語データベース」( 既刊論文 )

2017 年 10 月、初期近代英国演劇 527 編( 1602 年出版の *The Spanish Tragedy* の付加台詞

[ additions ] を 1 作品として計上) の作品間連語一致 ( collocation and n-gram match ) データベースがウェブ上に公開された。この連語データベースは、近代初期英国演劇の作者同定研究の基礎資料として作成されたものである。Rizvi は 527 編の戯曲 ( 632 区分 ) 間の連語一致数に基づいて、作品間の言語的遠近関係を数値化している。Rizvi のデータベースのサイズは、たとえば *Hamlet* の連続単語列一致のデータが 13 列約 7 万行、近接単語列一致のそれが 13 列約 100 万行であることから窺えるように、過去の同定研究の基本資料が兎戯のように見えるほど圧倒的である。Rizvi は、彼自身が作成したコンピューター・プログラムを用いてこの膨大なデータベースを構築した。本稿では、このデータおよび解析プログラムの詳細を把握し、近代初期英国演劇の作者同定研究におけるその効用と課題を詳細に検討し、次のように結論した。

単独執筆作品の弁別力はかなり高く、作者同定のツールとして有用である。単独執筆作品の作者判定は、終始デジタル的に行われる。その意味において、データの散布図を利用して作者判定を行う ( したがって、アナログ的・恣意的判断が介在しやすい ) 統計的手法より信頼度が高いと言える。Rizvi の解析資料の構築とその処理方法に瑕疵がないわけではないが、Rizvi は膨大なデータベースを単独で構築し、その生データをその解析プログラムとともに惜しげもなく公開し、全世界の初期近代英国演劇研究者の用に供している。最大級の賛辞を受けるに値する達成である。

( 5 ) 「シェイクスピアは *Arden of Faversham* の作者か 印象批評と計量文体解析」 ( 口頭発表 )

*Arden of Faversham* の作者問題を、従来の印象批評と近年注目が高まっている計量文体解析による作者判定方法の観点から考察した。印象批評によって決着のつかない本作品の作者同定に計量文体解析を適用し、その作者をシェイクスピアだと同定した統計学的判定方法の問題点を指摘した。

( 口頭発表 )

( 6 ) 「計量文体解析と戯曲の作者同定」 ( 既刊論文 )

*The New Oxford Shakespeare: The Complete Works* (2016) の作者同定を支持する研究者たち ( Elliott, Greatley-Hirisch, Jackson ) の統計学的作者判定を批判的に検証し、彼らの解析はサンプル戯曲の選択と解析方法が恣意的・不適切であり、その作者判定は説得力を欠いていると論じた。( 原著は日本語、中国語訳にて出版、訳者は、華東師範大学外国語学部研究生廖柳棉氏 )

( 7 ) 「*The New Oxford Shakespeare* の統計的作者判定方法の限界」 ( 口頭発表 )

NOS の統計学的な作者判定方法の問題点を、Pervez Rizvi の戯曲間連語一致データベースに基づく作者判定結果を傍証として指摘した。NOS の編者・研究者たちは、自分たちの選定した演劇作品から抽出した言語データの解析によって客観的な作者判定結果が得られるかのような議論を展開しているが、データの抽出・処理方法によって判定結果が変動する可能性を無視している。シェイクスピア時代の戯曲は共作が非常に多く、計量文体解析に必要な量の単独執筆テキストを確保することがむずかしく、作者不詳作品の作者を統計学的解析によって確実に同定するのは困難である。( 口頭発表 )

## 5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 太田一昭	4. 巻 42
2. 論文標題 Pervez Rizviの初期近代英国演劇連語データベース	5. 発行年 2019年
3. 雑誌名 言語文化論究	6. 最初と最後の頁 17-38
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 オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 OTA Kazuaki	4. 巻 87
2. 論文標題 Edward III and Statistical Attribution Studies	5. 発行年 2017年
3. 雑誌名 POETICA	6. 最初と最後の頁 63-79
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 有
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 -

1. 著者名 太田一昭	4. 巻 40
2. 論文標題 ジョナサン・ホープの作者同定研究の功罪	5. 発行年 2018年
3. 雑誌名 言語文化論究	6. 最初と最後の頁 1-14
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 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 太田一昭	4. 巻 4 (57)
2. 論文標題 『二重の欺瞞』の作者同定と文体統計解析	5. 発行年 2018年
3. 雑誌名 Shakespeare Journal	6. 最初と最後の頁 11-23
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 有
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 -

1. 著者名 太田一昭	4. 巻 -
2. 論文標題 Edward IIIの印象批評とコンピューター解析	5. 発行年 2016年
3. 雑誌名 日本英文学会第88回大会Proceedings	6. 最初と最後の頁 75-76
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 -

1. 著者名 太田一昭 (著)・廖柳棉 (訳)	4. 巻 44
2. 論文標題 計量文体解析と戯曲の作者同定	5. 発行年 2020年
3. 雑誌名 言語文化論究	6. 最初と最後の頁 67-74
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7件 (うち招待講演 2件 / うち国際学会 1件)

1. 発表者名 太田一昭
2. 発表標題 Jonathan Hopeの著者同定研究の功罪
3. 学会等名 第29回エリザベス朝研究会
4. 発表年 2017年

1. 発表者名 太田一昭
2. 発表標題 Double Falsehoodの「原作者」は同定されたのか
3. 学会等名 第56回シェイクスピア学会
4. 発表年 2017年

1. 発表者名 太田一昭
2. 発表標題 シェイクスピア研究と本文編纂（シンポジウム「Drama, Theater, Performance & Beyond 演劇研究の諸相」）
3. 学会等名 日本アメリカ文学会第56回全国大会
4. 発表年 2017年

1. 発表者名 太田一昭
2. 発表標題 Edward IIIの印象批評とコンピューター解析
3. 学会等名 日本英文学会第88回大会
4. 発表年 2016年

1. 発表者名 太田一昭
2. 発表標題 シェイクスピアはArden of Faversham の作者か 印象批評と計量文体解析
3. 学会等名 「娯楽文化史からとらえるエリザベス朝演劇」第1 回研究会（招待講演）
4. 発表年 2019年

1. 発表者名 太田一昭
2. 発表標題 計量文体解析と戯曲の作者同定
3. 学会等名 「戯単、劇場と20世紀前半の東アジア演劇」学術シンポジウム（招待講演）（国際学会）
4. 発表年 2019年



1. 発表者名 太田一昭
2. 発表標題 The New Oxford Shakespeare の統計的作者判定方法の限界 その傍証としての Pervez Rizvi のデータベース
3. 学会等名 第37回エリザベス朝研究会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
--	---------------------------	-----------------------	----