



Title	Exploring Stylistic Features in Agatha Christie's Works through Stylometric Approaches
Author(s)	土村, 成美
Citation	大阪大学, 2021, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/81941
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

論文内容の要旨

氏名 (土村 成美)

論文題名

Exploring Stylistic Features in Agatha Christie's Works through Stylometric Approaches

(計量的アプローチを使用したアガサ・クリスティー作品の文体の特徴に関する分析)

論文内容の要旨

イギリスの女性ミステリー作家 Agatha Christie は 1920 年に『スタイルズ荘の怪事件』(*The Mysterious Affair at Styles*)でデビューした後、50 年以上に渡って数多くの作品を世に送り出した。彼女の作品の累計売り上げ部数は 20 億部とも言われ、聖書と Shakespeare に次ぐ出版部数を誇っており、今でもなお世界中で広く読まれている。

Christie の作品は何故多くの人々を魅了するのか、本論文では計量的な手法を通して彼女の文体の特徴を解明することを目的とする。本論文では特に、Christie と同時代に活躍した 3 人の女性ミステリー作家の作品との比較分析を行うことにより、彼女の文体の特徴を記述するとともに、50 年余りの期間に渡って著された彼女の膨大な作品群を通時的観点から分析した。

第 1 章では本論文の背景として Christie 作品に対する評価の概要と彼女の生涯について記述し、本論文の構成を示した。

第 2 章では Agatha Christie の作品の文体を扱った先行研究を概観し、文体の計量研究史の中で位置付けた。Christie 作品の文体に対する観点の一つとして、Christie が晩年にアルツハイマー病を発症していた疑いがあることから、彼女の晩年の言語能力の低下に焦点を当てた研究が多く見られる。また、彼女の作品内での言葉の巧みな使い方に関する研究も行われており、彼女は読者を誤誘導しながらも、作家としてのフェアネスを担保していたことを示している。先行研究では分析対象の作品が限定的であることから、本論文では Christie 作品の網羅的な分析を行うことで、従来の研究では得ることの出来なかった知見をもたらすことを意図している。同時に、先行研究では用いられていない機械学習の手法を用いて、Christie 作品の文体の特徴を明らかにすることを目的としている。

第 3 章に本論文の使用データを示した。分析データとして、Christie の長編作品 66 作ならびに、比較対象として彼女と同時代に活躍した 3 人の女性作家(Dorothy L. Sayers, Margery Allingham, Ngaio Marsh)の長編作品を使用した。この 4 作家はイギリスのミステリー黄金時代と呼ばれる 1920 年代から 1930 年代にかけて人気を博した作家であり、「イギリスミステリー黄金時代の 4 大女王」と称されている。ジャンルが文体の差異に及ぼす影響を最小限にするため、分析対象は全てミステリー作品に限定した。

第 4 章では、他作家と比較して Christie の文体にどのような特徴があるのか解明を試みた。分析手法は、教師なしの分析としてクラスター分析と主成分分析を、教師ありの分析として機械学習の一種であるランダムフォレストを使用した。分析には全作家の使用語彙のうち、人名を除外した頻度上位 700 語を説明変数とした。

はじめにクラスター分析を行なった結果、Christie の殆どの作品は他作家の作品と異なるクラスターを形成することが判明した。分析を複眼的なものにするために並行して主成分分析を行なったところ、多少の重なりがあるものの、Christie 作品と他作家の作品が多次元空間において識別可能なクラスターとして布置される結果が得られた。主成分分析において第一主成分と第二主成分における寄与率の高い説明変数に着目すると、代名詞 *she, he* など、女性と男性の対照を反映する語が顕在化した。女性代名詞と男性代名詞の使用実態を吟味するために、それらのコロケーションを調査した。語と語の意味的結合傾向を探るために、統計指標として相互情報量 (MI スコア) に依拠した。Christie 作品においては泣く行為や、恐怖、悲しみを表す語など、作品内でのステレオタイプのな女性性を表象する語が特徴語として浮かび上がってきた。この傾向は他作家の作品における女性代名詞のコロケーションにおいても多少は認められた。他方、伝達動詞 *said* が Christie 作品を強く特徴づけていることがわかった。伝達動詞の頻度は談話の表出、特に直接話法と深く関わっていると考えられるため、Christie と他作家の各作品における直接話法の使用率を集計したところ、Christie 作品は他作家に比べて登場人物間の会話を通してプロットを展開する傾向が強いことが明らかとなった。特に彼女が晩年に執筆した作品ではその傾向が一層高まっていることが分かった。

さらに、頑健性の高い群間比較分析法であるランダムフォレストを実行することで、Christie の文体指標の特定

を試みた。その結果、Christie の特徴語として、*thoughtfully, slowly* など多くの *-ly* 副詞が特定された。Christie と他作家の間で *-ly* 副詞の使用実態にいかなる差異があるか考察するために、コロケーションを詳細に分析した。その結果 Christie は *-ly* 副詞を用いた固定的な表現を多用し、登場人物の言動を表現していることが明らかになった。

第5章では、分析対象を Christie 作品のみに限定し、通時的観点から彼女の文体にアプローチした。Christie は晩年に腕を骨折し、タイプライターによる執筆からディクタホンの使用へと、小説の執筆方法を変更したことが知られている。執筆方法と文体との関係を紐解く試みとして、執筆方法の変化があった時期の前後で文体に変化が認められるのかどうか定量的に調査した。分析手法は4章と共通しているが、使用データは、Christie 作品 66 作のうち、彼女が腕を骨折した 1952 年の前後 5 年の作品を除外した計 54 作品を使用した。腕を骨折した直前・直後の作品を分析から除外することで 2 群のデータ間に統計的な差異が存在するかどうかを鮮明にするため、この 54 作品を前期・後期の 2 群に分類し、分析を行った。

はじめにクラスター分析と主成分分析を行なった結果、女性代名詞と男性代名詞の対比と、一人称代名詞と三人称代名詞の対比を確認することが出来た。前章と同じ原理に基づき、コロケーションを通して *she* と *he* の使用実態を吟味したところ、*she* の共起語は女性登場人物の作中役割を表象する単語に偏っており、女性の描写がきわめてステレオタイプ的ななされている様相が明らかになった。また、*she* のコロケーションには *choked* が含まれているのに対し、*he* のコロケーションには *kills, strangled* が含まれており、そこには殺人事件の被害者は女性、加害者は男性というプロットの典型が反映している。

次にランダムフォレストを使用して前期と後期作品を識別する特徴語を抽出した。その結果、前期の作品の特徴語には動詞の過去形が多く含まれるのに対し、後期の作品の特徴語には動詞の原形が多く含まれることが明らかになった。コンコーダンスの精読を通して特徴語のコロケーションを分析したところ、前期の作品では過去形動詞を使用した固定表現を Christie が繰り返し使用していることが判明した。Christie が執筆方法を変えた後の作品を特徴づける語には直接話法の一部を構成する単語が多く含まれていた。後期の作品の特徴語には曖昧語 (*something, things* など) も含まれており、このことは Christie の言語能力の低下に焦点を当てた先行研究の知見とも一致している。曖昧語の使用頻度の増加は Lancashire & Hirst (2009) 等において言語能力の低下の兆候を示す指標として扱われている。直接話法の引用部における曖昧語の使用については、登場人物の個人語表出の一手段と捉えることもできるが、後期の作品における上記曖昧語は主に語り(narrative)において生起しており、Christie の言語能力低下との関連性を探る手がかりの一つであると考えられる。

第6章ではトピックモデルを使用し、意味的側面から Christie 作品の文体に迫った。トピックモデルは、近接文脈内で共起する語句は同じ話題、トピックを共有しているという前提に立脚し、確率論的アルゴリズムによって共起する語群を特定する文書解析技術である。それにより、従来の手法では扱うことが困難であった意味の問題に焦点を当てた分析が可能となる。はじめに抽出するトピック数を 50 に設定してトピックモデルを行なった結果、他作家と比較して Christie 作品を有意に特徴付けるトピックとして、「家族と人生」、「旅」、「会話表現」、「身体部位と動作」、「不定指示」、「フランス語法」などを表す共起語群が抽出された。「家族と人生」は、*Crooked House* をはじめ、中～後期の作品において頻出するトピックである。「旅」は、Christie 作品の基調をなす題材として言及されていることが読み取れる。「会話表現」には、縮約形も多く含まれている。第4章の会話の含有率グラフからも明らかのように、登場人物間の会話が物語の展開軸として大きな機能を果たす Christie の作風を示す項目といえるであろう。「身体部位と動作」は、*looked at, shook, nodded* など動作を表す動詞に加え、*-ly* 副詞によって構成されている。Christie 作品では、様態や程度を示す *-ly* 副詞による修飾を介して人物の所作の背後にある心理が描かれる顕著な傾向があることが認められた。「不定指示」は、前章で言及した曖昧語によって構成されている。トピックモデルを援用した分析の意義は、ランダムフォレストとは原理が全く異なる分析法を通して共通する特徴が特定されることが示されたことである。「フランス語法」は、Christie 作品の主人公ポワロの作中役割を縮約するトピックであると解釈できる。

第7章に本論文の結論を述べた。以上の分析から、Christie と他作家の文体の違いに関しては、第一に女性代名詞とそのコロケーションを通して観察すると、女性登場人物の作中における役割および女性の描写がステレオタイプの言語表現により体系的になされている様相が明らかになった。第二に Christie は多様な *-ly* 副詞とそれが修飾する動詞を結びつけて、固定的な表現を多用することで読者に人物の心理や類型についてのシグナルを送っていることが示唆された。Christie の文体の通時的な変化に関しては、第一に作品中直接話法で表出される会話の占有率が円熟期の作品に向かって増加する傾向があることを明らかにした。第二に彼女の晩年の作品では曖昧語による不定指示の頻度が高まる傾向にあることを示した。このことは、複数の分析手法を通して、量的な証拠によって先

行研究の知見を裏づけるものである。

Agatha Christie は長編小説にとどまらず、多数の短編小説を出版した作家である。本論文では研究対象を長編作品に限定したため、構成や構造が大きく異なる短編作品においても本論文の知見がそのまま当てはまるのかどうかについては今後の研究課題である。そのためには、トピックモデルを短編小説に最適化し、分析法を精緻化することが必要になる。

論文審査の結果の要旨及び担当者

氏 名 (土村成美)			
		(職)	氏 名
論文審査担当者	主 査	教授	田畑 智司
	副 査	准教授	ホドシチェク ボル
	副 査	名誉教授	木村 茂雄
	副 査	教授	森 祐司
<p>本博士学位請求論文は、計量テキスト分析の手法を通してイギリスの女性ミステリー作家 Agatha Christie の文体の特徴を記述するとともに、50 年余に渡って著された膨大な作品群を通時的観点から分析し、文体の時系列変化にアプローチした研究である。</p> <p>本論文は全 7 章からなる。第 1 章では本論文の背景として Christie 作品に対する評価の概要と彼女の生涯について記述し、本論文の構成を示している。</p> <p>第 2 章は Agatha Christie の作品の文体を扱った先行研究を概観し、本研究の位置付けを行なっている。本論文は、Christie 作品の文体に対する観点として、作品内での巧妙な言語使用、フェアネスを担保しつつも読者を巧みに誤誘導することばの仕掛けを明らかにする系統と、Christie の晩年の言語能力の低下を疑わせる言語特徴に焦点を当てた研究の系譜があることを指摘する一方、分析対象の作品が限定的な既存研究を補うためには、作品群の網羅的な分析により、マクロ的視点からの考察が必要であると論じている。そのために、統計学的視覚化や機械学習の手法を用いて、Christie 作品の文体の特徴を記述する方法論を提示している。</p> <p>第 3 章は、本論文が分析対象とするテキストコーパスについて詳細に記述している。標的コーパスは Christie の長編作品 66 点からなる。比較対象となる統制群コーパスには、彼女と同時代に活躍した 3 人の女性作家(Dorothy L. Sayers, Margery Allingham, Ngaio Marsh)の長編作品計 63 点を収録している。統制群コーパスを構成する 4 人の作家はイギリスのミステリー黄金時代と呼ばれる 1920 年代から 1930 年代にかけて人気を博した「ミステリーの 4 大女王」と称される作家である。本研究は、文体の計量的解析において、分析データの統制を厳密に行う必要があることを重視し、分析対象とするジャンルは全てミステリー作品に限定している。</p> <p>第 4 章では、統制群コーパスの作家と比較して、Christie の文体にどのような特徴があるのか解明を試みている。分析手法は、外的分類基準を参照しない <i>unsupervised analysis</i> として伝統的クラスター分析と主成分分析を、外的分類基準を参照する <i>supervised analysis</i> として機械学習に基づく分類器ランダムフォレスト法を用いている。まず、コーパスにおける高頻度語から、人名を除外した 700 タイプを説明変数とし、クラスター分析を行なった結果、Christie の殆どの作品は統制群の作品と異なるクラスターを形成することを示している。分析を複眼的なものにするために、並行して主成分分析を行なったところ、多少の重なりはあるものの、Christie 作品と統制群の作品が多次元空間において識別可能なクラスターとして布置される結果が得られる共に、Christie の特徴語ならびに統制群の特徴語が顕在化することを指摘している。語と語の意味的結合の強さを測る統計指標、相互情報量 (MI スコア) に依拠して特徴語の典型的な使用傾向をまとめると、Christie 作品においては泣く行為や、恐怖、悲しみを表す語など、作品内でのステレオタイプのな女性性を表象する語が特徴語となっていることを指摘している。他方、伝達動詞 <i>said</i> が Christie 作品を強く特徴づけており、伝達動詞の頻度が談話の表出、特に直接話法と相関するという仮説を検証するために、Christie と統制群の各作品における直接話法の使用率を詳細に計測した結果、Christie 作品は統制群に比べて、登場人物間の会話を通してプロットを展開する傾向が強く、特に晩年に執筆した作品ではその傾向が一層高まっていることを明らかにしている。さらに、頑健性の高い群間比較分析法であるランダムフォレスト法を実行することで、Christie の文体指標の特定を試みた結果、特徴語として、<i>thoughtfully, slowly</i> など数多の <i>-ly</i> 副詞を特定している。コロケーションを詳細に分析することにより、Christie は <i>-ly</i> 副詞を用いた固定的な表現を多用し、登場人物の言動を表現していることを明らかにしている。</p> <p>第 5 章は、分析対象を Christie 作品のみに限定し、通時的観点から作品群を分析している。Christie は晩年に腕を骨折し、タイプライターからディクタホンへと、小説の執筆方法を変更したことが知られている。本論文では、執</p>			

筆方法と文体との関係を紐解く野心的な試みとして、執筆方法の変化があった時期の前後で文体に変化が認められるのかどうかを探っている。まず、クラスター分析と主成分分析の結果、執筆方法変更の前後で、女性代名詞と男性代名詞の対比、一人称代名詞と三人称代名詞の対比を確認している。前章と同様に、コロケーションを通して女性代名詞と男性代名詞の使用実態を吟味することにより、前期作品では *she* の共起語は女性登場人物の作中役割を表象する単語に偏っており、女性の描写がきわめてステレオタイプ的になされている様相を明らかにしている。また、*she* を特徴づける共起語が *choked* であるのと対照的に、*he* を特徴づける共起語は *kills, strangled* などであり、殺人事件の被害者は女性、加害者は男性というプロットの典型が反映していることを指摘している。ランダムフォレストによる特徴語分析の結果、前期作品の特徴語には動詞の過去形が多く含まれるのに対し、後期作品の特徴語には動詞の原形が多く含まれることが明らかになった。本論文では、コンコーダンスの精読を通して、前期作品では過去形動詞を使用した固定表現を Christie が繰り返し使用していることを示している。また、後期作品の特徴語には直接話法の伝達部を構成する単語が多く含まれていることから、後期作品のプロット進行がより会話に依存したものに推移していることを指摘している。後期作品の特徴語には曖昧語 (*something, things* など) も含まれており、このことは Christie の言語能力の低下に焦点を当てた先行研究の知見を裏づけるものである。曖昧語の使用頻度の増加は Lancashire & Hirst (2009) 等において言語能力低下の兆候を示す指標として扱われているが、本論文では後期の作品における曖昧語は主に語り(narrative)において生起しており、Christie の文体変化を探る上で重要な項目の一つであると考えている。

第6章は、自然言語処理技術トピックモデリングの応用により意味的側面から Christie 作品に迫っている。トピックモデリングは、近接文脈内で共起する語句は同じ話題、トピックを共有しているという前提に立脚し、確率論的アルゴリズムによって共起する語群を特定する文書解析技術である。それにより、従来の手法では扱うことが困難であった「意味」の問題に焦点を当てた分析が可能となる。本論文では、トピックモデリングにより、統制群と比較して Christie 作品を有意に特徴付けるトピックとして、「家族と人生」、「旅」、「会話表現」、「身体部位と動作」、「不定指示」、「フランス語法」などを表す共起語群を特定している。「家族と人生」は、*Crooked House* をはじめ、中～後期の作品において頻出している一方、「旅」は、Christie 作品の基調をなす題材として幅広く言及されていることがわかる。「会話表現」は、第4章の分析とも一致し、登場人物間の会話が物語の展開軸として大きな機能を果たす Christie の作風を示す項目であることを裏付けている。「身体部位と動作」は、*looked at, shook, nodded* など動作を描写する動詞に加え、*—ly* 副詞によって構成されている。Christie 作品では、様態や程度を示す *—ly* 副詞による修飾を介して人物の所作の背後にある心理が描かれる、という顕著な傾向を確認している。「不定指示」は、前章で言及した曖昧語によって構成されている。トピックモデルを援用した分析の意義は、ランダムフォレストとは原理が全く異なる分析法を通して共通する特徴が特定されることが示されたことである。「フランス語法」は、Christie 作品の主人公ポワロの作中役割を縮約するトピックであると解釈している。

第7章は本論文の結論である。Christie の作品の特徴として、まず、女性登場人物の作中における役割、および女性の描写がステレオタイプ的な言語表現により体系的になされている様相を明示している。第二に Christie は多様な *—ly* 副詞と共起する動詞により、固定的な表現を多用することで読者に人物の心理や類型についてのシグナルを送っていると述べている。Christie の文体の通時的な変化に関しては、第一に作品中直接話法で表出される会話の占有率が円熟期の作品に向かって増加する傾向があること、第二に彼女の晩年の作品では曖昧語による不定指示の頻度が高まる傾向にあることを、複数の分析手法を通して再現し、量的な証拠によって先行研究の知見を裏づけるものであると言える。

本論文は、原理が異なる複数の分析手法を援用することで、データとしてのテキストを複眼的に考察し、Christie のミステリー作品の特徴を手堅く記述しており、客観性、再現性や反証可能性を担保するための科学的な手順を踏まえた研究となっている。一方で、量的分析から得られた知見が、Christie 作品のいかなる「新たな」読みに寄与するかという踏み込んだ議論を展開するには至っておらず、これは今後の課題であろう。また、論文の英語表現において、明晰さを欠く箇所が若干見受けられることも指摘せねばならない。しかし、これらは本論文の学術的価値を損なうほどの瑕疵ではなく、今後の研究によるさらなる深化発展が期待されるものである。全体として本論文は、計量的テキスト分析の方法論、自然言語処理の解析技法と、コンコーダンスラインの読みを有機的に組み合わせて Agatha Christie のミステリー作品の文体を論じた意義深い研究であり、その学術的貢献は明確である。

以上により、本論文を博士(言語文化学)の学位論文として価値のあるものと認める。

なお、チェックツール iThenticate を使用し、剽窃、引用漏れ、二重投稿等のチェックを終えていることを申し添える。