

平成 29 年 6 月 23 日現在

機関番号：32692

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25370563

研究課題名(和文) The Eighteenth-Century Periodical Essay: A Corpus-based Approach

研究課題名(英文) The Eighteenth-Century Periodical Essay: A Corpus-based Approach

研究代表者

C. P. Brocklebank (Brocklebank, C. P.)

東京工科大学・教養学環・准教授

研究者番号：40386769

交付決定額(研究期間全体)：(直接経費) 900,000円

研究成果の概要(和文)：本報告書では、18世紀の英語のエッセイの文体の解析結果を概説する。解析の対象として、サムエル・ジョンソン、ジョセフ・アディソン及びジョナサン・スウィフトの3名のエッセイストを選んだ。本研究では、コンピュータープログラムを使用し、前記著者のエッセイのセット(コーパス)を分析し、著者のキーワード、及び共通して使用頻度の高いマルチワードストリングを特定した。最も顕著で意味のあるキーワードが属していたセマンティックフィールドは、ジョンソンについては『幸福及び満足』、アディソンについては『言語、スピーチ及び文法』、スウィフトについては『政府及びパブリックドメイン』であった。

研究成果の概要(英文)：This report outlines the results of a stylistic analysis of eighteenth-century English essays. The three essayists chosen for analysis were Samuel Johnson, Joseph Addison and Jonathan Swift. The research used computer programs to analyze sets (corpora) of these authors' essays and to identify the authors' keywords and most commonly used multi-word strings. For Johnson the most prominent and meaningful keywords belonged to the semantic field of 'Happiness and Contentment', for Addison 'Language, Speech and Grammar', and for Swift 'Government and the Public Domain'.

研究分野：言語学

キーワード：文体論 コーパス言語学 18世紀のエッセイ キーワード ジョセフ・アディソン ジョナサン・スウィフト サミュエル・ジョンソン

1. 研究開始当初の背景

過去 10 年間、コーパス文体論の分野は、収集された莫大なテキストを分析し、このテキスト集合体の隠れた文体の特徴を発見するために、コンピューターの能力を利用している。その結果、数多くの発見がなされてきた。しかしながら、これまで大部分のコーパス文体論の研究において焦点が当てられてきたのは、小説や、教師が言語の教材作成のためのデータを提供するための特殊な科学のテキストに留まっていた。ここで概説するプロジェクトで目指したのは、18 世紀の英語の刊行物のエッセイにスポットライトを当て、18 世紀前中期における 3 人の文体の「トレンドセッター」であるサムユエル・ジョンソン、ジョセフ・アディソン及びジョナサン・スウィフトの文体分析を行い、現在のコーパス言語技術を応用し、適用することが可能であるかを確認することであった。

2. 研究の目的

- (1) 調査の目的は、具体的には、コーパス言語学の技術を使い、次の事項を特定することであった。

各著者のエッセイにおけるキーワード、キーワードデータにおけるパターンの発見

3 人の著者の著作物の中で共通して使用される頻度が最も高い単語連鎖(つまり、(1 つの単語ではなく)繰り返される文字列)及び最も重要な単語連鎖、並びにこれらの単語連鎖の分析

- (2) 調査の補足的な目的は、次のとおりである。

著者(ジョンソン及びアディソン)の 2 人を例とする 18 世紀の刊行物エッセイというジャンルの言語と、18 世紀の小説という別のジャンルの言語の比較。この比較もまた、2 つのジャンルタイプに対するキーワードとキーとなる単語連鎖の生成、及び取得データにおけるパターンの発見を含む。

他の文体の領域にこの分析を拡張できるかどうかを確認すること、及び、現在利用可能ないくつかのコーパス言語ツールを試みる。ここで調査する範囲には、エッセイの各章における(文レベル、段落レベル及びエッセイ全体のレベルでの)キーワードの分布の追跡、並びに、キーワードのキーコロケーション(ターゲットの単語の前後で最も頻繁に出現する単語)の特定。

3. 研究の方法

- (1) キーワード取得のため、コーパスツール

の WordSmith のセットを使って 3 つのテキストファイルからワードリストを生成した。このテキストファイルには、ジョンソンの The Rambler、The Adventurer 及び The Idler からのエッセイ、アディソンの Spectator エッセイのすべて、並びにスウィフトの The Tatler、The Examiner、The Spectator 及び The Intelligencer からのエッセイが含まれている。また、エッセイファイルの大きさに比例する各単語の頻度をログ尤度値として計算した。各ファイルの大きさにはかなりの差があったため(ジョンソンのファイルは 434,344 のワード、アディソンのファイルは 323,841 ワード、及びスウィフトのファイルは 88,731 ワードで構成されていた)双方向比較で取得された各上位キーワードについて、平均重要度(keyness)スコアを計算した。例えば、キーワード「BY」の重要度スコアは、ジョンソンとアディソンとの比較及びジョンソンとスウィフトとの比較において、平均で 450.78 であった。ジョンソンについては、合計 92 のキーワード、アディソンについては 64 のキーワード、及び、スウィフトについては 115 のキーワードを分析した。その後、キーワードを共有意味素性ごとのグループに分類し、3 セットのエッセイの構造及び内容における違いを引き出した。

- (2) kfNgram プログラムによってテキストファイルを実行することにより、3 人の著者に最も共通する単語連鎖(n-grams)が得られた。ジョンソン及びアディソンについては、4 単語の連鎖(4-grams)を分析し、3 人の著者すべてと比較した。ただし、スウィフトについては、ファイルの単語数が少ないため(上記のとおり)、意味のある比較をするのに十分な数の 4 単語の連鎖の例が取得できない可能性があった。このため、その代わりとして 2 単語の連鎖について検討することにした。

- (3) 18 世紀の小説との比較において、上記の(1)及び(2)に概説されるステップを行った。大規模な小説のコーパス(750 万ワード超)では、主要な小説の 4 単語の連鎖を、エッセイストのジョンソン及びアディソンの 4 単語の連鎖と比較することができた。

- (4) 単語の分布及びキーコロケーションの調査では若干異なる方法を取ったが、これについては、下記の調査結果の説明で簡潔に取り上げる。

4. 研究成果

- (1) 各著者のキーワードのトップ 10 (降順)を下の表 1 で示す。

ジョンソン	アディソン	スウィフト
BY	IN	MINISTRY
WITHOUT	SEVERAL	PUBLIC
CAN	SELF	CHURCH
AND	VERY	WHIGS
OR	UPON	HATH
HAPPINESS	THO	QUEEN
ALWAYS	READER	PRINCE
LIFE	PARTICULAR	PEOPLE
EVERY	KIND	LATE
NO	POEM	THEIR

表1 ジョンソン、アディソン及びスウィフトのエッセイにおけるキーワードのトップ10

通常、テキストのキーワードは、テキストの内容を反映している。上記の言葉について驚くべきことは、ジョンソン及びアディソンについては、キーワードが内容を示す言葉というよりはむしろ、文法的な要素となっているということである。例えば、ジョンソンのリストには、2つの前置詞(BY、WITHOUT)、2つの接続詞(AND、OR)、2つの限定詞(NO、EVERY)及び1つの助動詞(CAN)が含まれており、「内容」を示す言葉は、HAPPINESS、ALWAYS及びLIFEの3つのみである。このことは、ジョンソンとアディソンのエッセイを特徴付ける文体の差異は、内容だけではなく、言語選択上の傾向も反映していることを示唆している。例えば、ジョンソンのエッセイで「BY」が際立っているのは、彼が記述する際に完全な形の受動体の文章を使う傾向があることが理由の1つとなっており、一方、AND及びORが上位にあるのは、ジョンソンのスタイルの一態様が、エッセイ中多くのパラレリズムであるからである(AND及びORはパラレリズムの要素と関連する。)アディソンに関して「IN」が上位であるのは、彼にINで副詞句を作る傾向があるためである。INは、他のキーワードPARTICULARと共起する。「in this particular」のフレーズは、前の文章で示された概念を、再び言及するために使用される(アディソンはこのフレーズを44回使用し、スウィフトは1回のみ、ジョンソンは全く使用していない。)

ジョンソン及びアディソンの上位キーワードの機能上の偏向のため、さらに、スウィフトのエッセイでより良い比較を引き出すため、研究代表者は、ジョンソン

及びアディソンにおける内容を示す上位キーワードに限定して、次の分析ステップを行うことにした。内容を示すキーワードのトップ10を著者ごとに表2に示す(副詞及び固有名詞も除外している。)

ジョンソン	アディソン
HAPPINESS	READER
LIFE	PARTICULAR
KNOWLEDGE	KIND
ATTENTION	POEM
HOPES	BEAUTIFUL
PLEASURE	ORDINARY
EXCELLENCE	FOLLOWING
FELICITY	POET
POWERS	NATURE
CONFIDENCE	SOUL

表2 ジョンソン及びスウィフトのエッセイにおける内容を示すキーワードのトップ10

次のステップとして、意味論的な類似性に従い内容を示すキーワードを分類した。これを行うため、研究代表者は『UCREL Semantic Analysis System』(USAS)で実行されたカテゴリーを利用した。このシステムは、選択した意味論上のカテゴリーの客観性を確保することにより研究者の先入観を入れないようにしている。

ジョンソンについて確認された主要なグループは、次のとおりであった。(a)『幸福及び満足』HAPPINESS、MERRIMENT、FELICITY、PLEASURE、GRATIFICATIONS、MISERIES及びMISERYから成る。明らかに、ジョンソンのエッセイにおける重要な関心事が「幸福」であることが分かる。(b)『(心理)期待』人間の希望や期待を示す(HOPES、HOPE、EXPECTED、EXPECTATION、EXPECTATIONS)、(c)『(心理)関心、興奮及び積極的』ARDOUR、DILIGENCE、CURIOSITY及びEAGERNESSから成る。(d)『評価』のグループ(SUPERIORITY、EXCELLENCE及びERROURから成る)と、(e)『比較』(EQUALLY、EQUAL、ACCUSTOMED)とは共に、エッセイの評価的な特徴を示す。(f)『性格に関する特徴』(KINDNESS、ENVY及びVANITYから成る。)も顕著である。(g)『頻度を表す副詞』(SOMETIMES、ALWAYS、SELDOM)は、ジョンソンがエッセイの中で何度も一般化を試みて

いることを反映している。

アディソンについては、確認されたグループは、すべて『言語学的なアクション、状態及びプロセス』に関連していた。これはアディソンの *Spectator* の中のエッセイの多くのスペースが、文学的な事項の議論に割かれていることを反映している。この議論には、ミルトンの *Paradise Lost* についての長い議論を含んでいる。USAS システムに従うと、この全体的な言語カテゴリーは、次のように細別することができる。(a) 『言語、スピーチ及び文法』(POEM、POET 及び WORD)、(b) 『メディア』(READER、BOOK 及び WORKS)、(c) 『言語行為』(DESCRIPTION、DESCRIBES)。また、エッセイには、PARTICULAR、KIND、NATURE 及び NATURAL のキーワードなどに見られる『分類』の傾向もある。スウィフトの場合、内容を示すキーワードの 4 つの意味論的グループが存在したが、これは、ジョンソン及びアディソンにおいてより顕著である。これらのグループは次のとおりである。(a) 『政府及びパブリックドメイン』MINISTRY、WHIGS (政党名)、PARTY、PARLIAMENT 及び KINGDOM など、合計で 30 例があった。(b) 『社会的なアクション、状態及びプロセス』PUBLIC、CHURCH、QUEEN、PRINCE 及び PEOPLE を含む 33 のキーワード、(c) 『時間』(「直前」という意味の LATE が最も顕著であった。)、(d) 『金銭及び取引』(CREDIT、TRADE)。ジョンソンの心理一般及び特に幸福への注目、並びに、アディソンの文学的議論の優先と比較すると、スウィフトのエッセイは、政治的事項、政府、君主制及び教会について詳説しており、これは、これは表面化したキーワードの意味論に反映されている。最も顕著な『時間』のキーワード LATE も、その傑出を現在の政治的な運営と直前の (= 'late') 政治的な運営との数多くの比較に負っている。

- (2) キーワードが(通常)原文の内容を反映する一方で、単語連鎖はテキストの構造上の機能をより反映する。4 単語の連鎖を選択した理由は、これより長い単語連鎖は、より制限された統語論的、意味論的及び語用論的な文脈で生じる傾向があるからである。

ジョンソンについては、10 回以上出現した最も顕著な 4 単語の連鎖を、手動で意味論的グループに等級付けし、分類した。存在した 4 つの主要グループは、次のとおりである。(a) 『書簡方式』(to the rambler sir、I am sir &c、to the idler sir、

am sir your humble、sir your humble servant、the rambler sir I) - (現実及びフィクションにおける)書簡の言及は、ジョンソンのエッセイに共通する特徴である。(b) 『範囲』(the greater part of、for the most part、greater part of mankind、the rest of mankind、it is common to)、ジョンソンの一般化傾向の産物である。(c) 『人間』(greater part of mankind、the rest of mankind、of the human mind)、ジョンソンの人間に対する関心を反映した 4 単語連鎖である。(d) 『容易又は困難』(it is not easy、is not easy to)。

対照的に、アディソンのエッセイにおける主要な 4 単語の連鎖は、より明確に談話構造、特に文章の構成のために使用された文言を反映している。副詞について、次の主要 3 タイプが単語連鎖によって実証された。(a) 『接続副詞』先行する部分と節又は文章を関係付ける副詞 (at the same time、and at the same、and by that means、the same time that、in the next place など)。実際、最も共通する 4 単語の連鎖は、「at the same time」であり、これはアディソンが 2 つのアイデアの両立を示すために使用している。(b) 『認識的立場副詞』意見の基礎となる知識に対する著者の評価又は信頼の程度を示す副詞 (there is no question、for my own part、truth of it is、I question not but など)。(c) 『場所の副詞』何かが存在する場所を示す副詞 (in the mind of、in the whole poem、in the first book など)。

スウィフトに関しては、エッセイの数が少ないため、他の 2 人の著者との比較を保証するのに十分な数の 4 単語の連鎖の繰り返しが存在しない。

その代わりに、スウィフトのエッセイでは、最も使われている 32 個の 2 単語の連鎖 (bi-grams) を選択し、その頻度をジョンソン及びアディソンにおける場合と比較した。この 32 個中、最も大きな差が現れたのは(a)『the present』であった。ここには、現 (『the present』) 内閣を前 (『late』) 内閣と比較するスウィフトのプロパガンダ戦略が表面化している。(b) 『such a』 スウィフトが強調のために使用しており、『such a [nominal] as』という形でよく見られる。他の 2 著者と比べ、スウィフトのエッセイでは、bi-gram の『have been』も頻出しており、スウィフトの議論の近接性を反映している。

- (3) エッセイ、小説という 18 世紀の 2 つのジャンル間の文体の差を明らかにするた

めに、限定的な範囲ではあるが、類似の比較を上記(1)に対して行った。小説のコーパスは、1684年(アフラ・ベーンの Love-Letters Between a Nobleman and His Sister)及び1793年(シャーロット・ターナー・スミスの The Old Manor House)以降の小説から得られたバランスの良いサンプルを含んでいる。小説及びジョンソン/アディソンについてのキーワードのトップ10は次の表3のとおりである。

小説	ジョンソン/アディソン
YOU	OF
I	ARE
HER	IS
SHE	OR
ME	THE
HAD	THOSE
MY	HAS
YOUR	WHICH
SAID	PUBLICK
WAS	THEIR

表3 18世紀の小説及びジョンソン/アディソンのコーパスにおけるキーワードのトップ10

小説とエッセイ間との差は顕著である。小説では、1人称(I, ME, MY)と2人称(YOU, YOUR)への言及が多く、女性形の要素(SHE, HER)が、より顕著である。小説では過去形(HAD, SAID, WAS)が、エッセイでは現在形(ARE, IS, HAS)が優位である。エッセイにおける「OF」の優位性は、エッセイでは名詞句がより長く、より複雑になっていることが理由であると思われる。また、同格(OR)又は関係代名詞(WHICH)など、エッセイでは連結のための言葉が多く使用されている。

これらの差の一部は、4単語の連鎖についての分析においても出現した。最もよく使われる4-gramsは、過去時制動詞(that I could not, as if he had) 2人称代名詞(let me tell you, I hope you will)及び1人称目的代名詞(let me tell you, give me leave to)を含む場合が多かった。注目すべきことは、同じ4単語の連鎖『at the same time』は、小説のテキストファイル及びアディソンのテキストファイルに関する両方のリストの1位であったことである。単語連鎖の統語論的な考察を得るために、これらを含むテ

キストファイルを、Wmatrixプログラムによって編集し、タグ付けした。小説における単語連鎖は、エッセイの場合より高い頻度で、冠詞(the rest of the, for the sake of) 『is』(that is to say, is one of the) 及び『of』(a great deal of, in the midst of)を含んでいることが分かった。

- (4) 最後に、ジョンソンのエッセイとアディソンのエッセイにおけるキーワード分布の比較、及び、ジョンソンとアディソン(スウィフトについては十分なデータが存在しなかった)における内容を示すキーワードの主要コロケーション分析について、簡潔に説明する。

ジョンソンとアディソンのエッセイのキーワードが文レベル、段落レベル及びエッセイレベルでどのように分布しているかを確認するため、Rで書かれたプログラムを使用して前述の単位を『Starts』、『Middles』及び『Ends』に分割した。その後、各単位の地位について、WordSmithを利用したキーワード比較を行った(例えば、ジョンソンのキーワード『Sentence Starts』をアディソン『Sentence Starts』と比較する、など)。主要な結果は次のとおりであった。(a)文レベルでは、キーワード「BY」は、ジョンソンにおいては、より頻繁に文の終わりの方に出現し、「TO」と「YET」は文の始めの方に出現する。(b)段落レベルでは、「AND」及び「WITHOUT」は、ジョンソンにおいては、より頻繁に段落の終わりの方に出現し、「YET」は段落の始めの方に出現する。(c)エッセイレベルでは、「BY」及び「RAMBLER」は、ジョンソンにおいては、より頻繁に、エッセイの始めの方に出現し、アディソンにおいては「CAN」がエッセイの終わりの方に出現する。

ジョンソンとアディソンにおける内容を示すキーワードの主要コロケーションを特定するために、2著者のそれぞれにつきキーワードのコンコグダンスを生成し、これらをWordSmithによって他の二人のコンコグダンスと比較した。ジョンソンについては、HAPPINESSの主要コロケーションの詳細分析を、単なるキーワードではなくエッセイの主要概念(『幸福及び満足』)であるとして実行した(上記(2)を参照)。HAPPINESSに関する3つの主要コロケーションは、LIFE、FOUND及びLONGであった。キーワード及び主要コロケーションを含む例を検索し、互いに近接して出現する場合に2つの単語を関連付けるための次の一般化が示された。(a)LIFEは、HAPPINESSの存在する場所である可能性もあるが、

それが欠けていること又はその一時的であることを示すことが多い。(b)HAPPINESSは、探し出すのが難しいか、努力が必要であるがFOUND(発見)する物である。(c)HAPPINESSは、来るべきLONG(長い)時間である場合もある。

- (5) 本研究プロジェクトの結果は、具体的かつ一般的である。具体的に、コーパスソフトウェアの使用なくしては明らかにならなかったであろう、ジョンソン、アディソン及びスウィフトのエッセイにおける語彙及びスタイルの態様を発見できた。例えば、伝統的なテキスト分析方法を使用している研究者が、ジョンソンのエッセイに関するキーワードとして「BY」を選定する可能性は低く、また、3セットのエッセイを綿密に読んだ結果、アディソンが他の著者より頻りに選択した4単語の連鎖として『at the same time』を選択する可能性は低い。さらに、コーパスソフトウェアにより、現在まで信頼されていた主張、例えば、「ジョンソンのエッセイは、パラレリズムを特徴としている」という主張に対する経験的実証を行うことが可能になる。「AND」及び「OR」をキーワードとして確立することにより、この主張を裏付けることができた。上記で概要を示したアプローチが有効となる場合は、より一般的な意味においても、18世紀であるかそれ以降であるかを問わず、採用した方法論が他のエッセイ執筆者にも適用できることは明白である。このアプローチは、そのほか、単一の著者による短いテキストのコレクションに対しても応用することができる。例えば、研究者が複数の著者による書簡、雑誌記事、新聞の論評部分、又は日記の記述を研究する場合にも、上記の比較アプローチを応用できる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 2件)

Brocklebank Paul、Identifying Distributional Patterns in Eighteenth-Century Periodical Essays, Discourse and Interaction, 8:1, 2015, 5-9 (査読あり)

DOI: 10.5817/DI2015-1-5

Brocklebank Paul、Johnson and the Eighteenth-Century Periodical Essay: A Corpus-Based Approach, English Language Overseas Perspectives and Enquiries (ELOPE)、10、2013、21-32 (査読あり)

DOI:10.4312/elope.10.2.21-32.

〔学会発表〕(計 4件)

Brocklebank Paul、A Corpus Linguistic Analysis of Swift's Periodical Output, SDAS 2016: Making It New In English Studies, 2016/09/17、マリボル(スロベニア)
Brocklebank Paul、Key Semantic Domains in Joseph Addison's Spectator Essays, English Studies at the Interface of Disciplines: Research and Practice (ESIDRP), 2016/03/11、スコピエ(マケドニア)
Brocklebank Paul、Language Skills: Working with Text and around Text (2nd International Conference), 2014/09/23、ルブリン(ポーランド)
Brocklebank Paul、Identifying Distributional Patterns in Eighteenth-Century Periodical Essays, Sixth Brno Conference on Linguistics Studies in English: Communication Across Genres and Discourses, 2014/09/12、ブルノ(チェコ)

6. 研究組織

研究代表者

ブロックルバンク C. P. (BROCKLEBANK, Paul)

東京工科大学・教養学環・准教授

研究者番号: 40386769