マクミラン委員会におけるロバートソンの経済観

——テキストマイニングから判明する「実物」と「貨幣」——

仲北浦　淳基
小峯　敦

目　次

Ⅰ．はじめに——本研究の目的
Ⅱ．先行研究——ロバートソンとマクミラン委員会
Ⅲ．テキストマイニングの「接合アプローチ」
Ⅳ．マクミラン委員会におけるロバートソン証言の特色
Ⅴ．おわりに

Ⅰ．はじめに[1]——本研究の目的

本研究の目的は，ケンブリッジ学派の経済学者 D. H. ロバートソン（1890-1963）が 1929 年の「金融および産業に関する委員会 Committee on Finance and Industry」（通称「マクミラン委員会」）に提出した証言資料 statement of evidence をテキストマイニング分析することで，本委員会の議論における彼の言説（特にその経済観）の特徴を探ることである。

ロバートソンは，マーシャル，ピグーを継いで，1944 年から 1957 年までケンブリッジ大学経済学教授を務めた経済学者である。通常，ロバートソンは，経済変動論，銀行論，利子論において大きな影響を及ぼしたと評価される。しかし，「ケンブリッジ学派の星座を占める輝ける星」（菱山 1997 : 161）と賞されるほど有力な経済学者であるにもかかわらず，ロバートソン研究は断片的なものにとどまっており，彼の経済学が体系的・包括的に理解されているとは言いがたい。そこで筆者の 1 名は，ロバートソンの「努力 effort」概念に着目することで，彼の経済学を体系的に把握／再構成しようと試みてきた。その結果，貨幣理論において重要な貢献をしたと考えられてきたロバートソン経済学の根底には，実は「努力」という実物的概念が通底している，と主張するに至っている（仲北浦 2018）。

ところで，マクミラン委員会は，主に，①ケインズの（特に貨幣に関する）思想が幅広い範囲で公表され（岩本 1999；小峯 1999；松川 2013），同時に，イングランド銀行の伝統的な金融政策が白日にさらされた場として（田中 1976；小峯 1998），②中小企業に対する慢性的な資金不足

1）次の研究会における参加者のコメントに感謝する。「戦争と平和の経済思想」および「テキストマイニング」研究会（2018. 9. 11，長崎大学）。また本研究は科学研究費・基盤 B「経済理論の大衆化から経済政策の形成へ：テキストマイニングを応用した実証研究」（研究代表者：下平裕之，課題番号 15 H 03330，2015-19 年度）および，科学研究費・研究活動スタート支援「成長と分配を両立させたマクロ経済分析の可能性－ロバートソン経済学の再検討－」（研究代表者：仲北浦淳基，課題番号 18 H 05709）の支援を受けた。

(「マクミラン・ギャップ」) を解消する専門機関 (後の商工金融公社 ICFC や産業金融公社 FCI) の必要性を訴えた場として (中村 2011 : 4 ; 鈴木 2017 : 67), さらに③金融界・産業界・労働界の利害関係者の見解が顕示された場として (吉田 1985 ; 松永 2003), 研究史上それぞれ重視されてきた。本稿はおおまかには①の流れに属するが, 貨幣をめぐる思想について, ケインズではなくロバートソンの側から眺めるものである。ロバートソンは実物 (「努力」概念) に基づいた貨幣を考えているが, 経済政策思想の形成の場としてのマクミラン委員会も, まさに産業 (実物的現象) と金融 (貨幣的現象) の結節点となっていた。また, ロバートソンは, かつてケインズと蜜月の状態であったが, 『貨幣論』 (1930) の出版前後から, 理論的に極めて冷やかな態度に転じていく。マクミラン委員会は『貨幣論』執筆途中のケインズが, 経済学の非専門家に向けてその神髄を解説している場でもあり, 証言にたったロバートソンとのやり取りに注目することで, 両者の貨幣観・経済観の変容や進化について, 新たな発見もありうるだろう。ゆえに, マクミラン委員会をロバートソン研究の立場から注目することもまた, 十分に意義あるだろう。

さらに, 本稿は通常の経済学史の方法 (原典の研究) にも則るが, なおテキストマイニングという新しい手法を導入する。原典による研究の結果を, 量的・質的に重層的な分析を含むテキストマイニングによって, これまでの研究蓄積を補強すれば (あるいは, さらなる仮説を生み出せれば), なお説得性を増すことができるであろう。逆に, テキストマイニングによっては持論を補強できない可能性も考慮に入れておこう。

本研究は, テキストマイニングを経済学説史・経済思想史の研究に適用する試みの 1 つである。ただし, 本研究は, 小峯・下平 (2017 : 2) で目的に掲げられている「経済学の通俗化」や「大衆を経由した政策形成論」の分析までには踏み込めていない。むしろ, 思想が大衆へと伝播し, 政策が形成されていく以前の段階, すなわち, 経済に関する思想・理論が実践の場で公表される段階を分析する。つまり, ロバートソンの思想・理論が, マクミラン委員会という実践の場でどのような特徴をもっていたのかを分析するのである。

そこで, 本研究では, マクミラン委員会の全体的な論調とロバートソンの主張を量的分析によって比較する。そうすることで, ロバートソンの言説に注目すべき特徴があるのかどうか, もし特徴があるならば, どのような特徴なのかを明らかにする。

以下では, 第 2 節において, マクミラン委員会とロバートソンによる証言に関する先行研究を概観し, 本研究の位置づけと方向性を確認する。第 3 節では, 本研究で用いるテキストマイニングの「接合アプローチ」について概説し, その手法と経済学史研究における従来の方法との統合について検討する。そして, 第 4 節では, 実際に, その手法を用いて『金融および産業に関する委員会報告書』 (1931) と, ロバートソンの証言資料である「世界不況 The World Slumps」 (1930) を分析する。また, 補論においては, テキストマイニング分析を実践する上で直面しうる 2 つの課題に言及する。第 1 は, 小峯・下平 (2017 : 59-60) や喜田 (2018 : 372-373) でも指摘されている, OCR からテキストファイル化までの煩雑な作業についてである。その作業の合理化を目指した一手法を提案する。第 2 は, 著作権の問題である。アナログ文書のデジタル化,

あるいは，デジタル化された文書の解析における著作権の問題について一考する。

Ⅱ．先行研究――ロバートソンとマクミラン委員会

2.1　マクミラン委員会

金融および産業に関する委員会（マクミラン委員会）は，1929年のウォール街における大暴落後，当時の蔵相スノーデンにより大蔵省の委員会として設置された。その主な目的は，英国経済における産業不況，失業，低成長，国際的地位の低下といった長期的停滞の根本的な原因を探り，それに対する改善案を提示することであった（加藤・西村 1985 ; ii）。

議長の H. P. マクミランは法律家であったが，J. M. ケインズを含め，その他の委員の多くは経済・金融の専門家であった。また，証人として合計57名が委員会で証言したが，その中にはイングランド銀行総裁の M. C. ノーマン，同副総裁の E. M. ハーベイの他，諸銀行の頭取たちが含まれる。のみならず，当時の一流の経済学者である，A. C. ピグー，L. C. ロビンズ，R. G. ホートレイも証言にたった。そこにロバートソンも列したのである。

ロバートソンは，委員会の第26日（1930年5月8日）と第27日（同5月9日）に証言した。彼は1930年4月に証言資料を提出していたが，その内容があまりにも難解であったため，2日間にわたって口頭証言をしたようである（Fletcher 2008 : 160／訳244）。また，ケインズは両日ともに出席し，ロバートソンと質疑応答をしている。

2.2　ロバートソンによる証言とその評価

では，マクミラン委員会におけるロバートソンの証言はどのように評価されてきたのであろうか。Presley（1979）と Fletcher（2008）の2つの先行研究を検討しよう。

Presley（1979 : Part. III, Ch. 1）は，ロバートソンの政策論を考察するにあたって，政策を金融政策・財政政策・賃金政策に分類し[2]，それぞれにおいて，マクミラン委員会でのロバートソンの証言に言及している。

金融政策については，ロバートソンが「純粋な貨幣政策では，1920年代の水準で産業活動を維持できないと私は思う」と証言したことを根拠に，市場利子率が投資の増加を刺激するほどには下落しないということに彼が気づいていた，と主張される（同上：243）。財政政策については，ロバートソンがマクミラン委員会の全体的な論調で，不況対策として財政出動を推奨し，「大蔵省見解」を真っ先に批判したことが指摘される（同上：246-247）。そして，賃金政策については，ロバートソンが，賃金の下落は財・サービスへの需要を損なうことを力説したことを根拠に，彼が，賃金水準と財への需要との関係を重視していたことが指摘されている（同上：251-252）。

2）おそらく，ケインズとの対比を念頭に置いた分類であろう。

このように，Presley（1979）は，マクミラン委員会におけるロバートソンの証言を根拠に，ロバートソンの政策論を再構成しようとしている。利子率操作の効果に対する懐疑，財政政策の推奨，そして，有効需要への言及を指摘している点で，この研究は重要であろう。しかし，マクミラン委員会におけるロバートソンの証言を包括的に分析しているわけではなく，むしろ，ケインズ政策との関係において近しい箇所を指摘するといった断片的な言及にとどまっている。

それに対して，Fletcher（2008 : 159-164／訳 243-250）は，マクミラン委員会におけるロバートソンの証言を比較的幅広く検討している。彼はマクミラン委員会とロバートソンの関係について，次のように述べている。

> 「…主な焦点は，銀行や金融・信用の役割や銀行業務に関する国内外における制約条件に合わせられており，さらに銀行が経済活動を促進することに役立つようにする方法も対象となっていた。／これらは，ロバートソンが証言を行うのに十分ふさわしいと思ったであろう問題であった」（Fletcher 2008 : 160／訳 244）。

つまり，マクミラン委員会の金融・信用および銀行業務という主題こそ，ロバートソンにとって証言するに適した論題であった，と言うのである。ゆえに，Fletcher（2008）はロバートソンの証言の中でも，銀行業務のあり方に関するものを特に重視している。また，ロバートソンが「大蔵省見解」を批判し財政出動を支持したこと，賃金の引き下げに後ろ向きだったことに言及している点は，Presley（1979）と同様である。

ただし，Fletcher（2008）は，ロバートソンが人間の欲望には限界があることを認識し，産業不況の原因として「飽和可能性 gluttability」という概念を重視したことを指摘している（同上：161／訳 247-248）。しかし，この指摘に関しても，結局のところ，銀行業がどうあるべきかという議論に結びつけられている。

以上のように，Fletcher（2008）は，Presley（1979）と比べると，マクミラン委員会におけるロバートソンの証言をより広く検討しており，「飽和可能性」というロバートソンの独自性を指摘している。しかし，銀行業のあり方という観点に絞ったことで，ロバートソンの独自的な考え方の意義やその理論上の位置づけまでは捉えきれていないようにみえる。

そこで本稿では，マクミラン委員会とロバートソンの証言を比較検討する。両文献全体を量的に解析することで，それらの議論の全体像を示した上で，ロバートソンの言説の特徴を探ることとしたい。

Ⅲ．テキストマイニングの「接合アプローチ」

テキストマイニングによる分析手法は，いまだ 1 つに定まっているわけではない。どのような研究分野に応用するかによって，様々な手法が並存しているようである。その中でも，最新の包

括的研究である喜田（2018 : 55-66）は，「CRISP_DM」[3]をデータマイニングの最も標準的な分析手法とみなし，その作業手順をテキストマイニング用に再編成した手法を提案している。ただし，この手法は，アンケートの自由記述などの内容分析を重視したものであるため，膨大な言語的資料を扱う経済学史の研究においては適用しにくいかもしれない。

そこで，本研究では，すでに，経済学史（下平・福田 2014；古谷 2014；Matsuyama 2016；小峯・下平 2017；Komine & Shimodaira 2018）のみならず，政策学における議事録分析（増田 2014 ; 2016 ; 2018）や，文学・哲学における思想分析（犬塚 2018）でも活用されている樋口耕一氏の KH coder を用いるため，彼が提案する「接合アプローチ」（樋口 2014）という分析方法を適用する。

本節では，まず 3.1 において，テキストの計量的分析における 2 大アプローチの異同を明らかにする。その上で，3.2 において樋口氏の「接合アプローチ」とその分析の流れを概説する。そして 3.3 において，その手法を経済学史研究に鋳なおす方法を一考する。

3.1　テキストの計量的分析における 2 つのアプローチ[4]

樋口（2014 : 17）は，コンピュータを用いたテキスト型データの計量的分析方法として，従来，異なる 2 つのアプローチがそれぞれ独自の発展を遂げてきたと指摘する。

第 1 は，「分析者が作成した基準（コーディングルール）にしたがって言葉や文書を分類するためにコンピュータを用いるアプローチ」（樋口 2014 : 17）である。これは，Dictionary-based アプローチと呼ばれる。

第 2 は，「同じ文書の中によく一緒に現れる言葉のグループや，あるいは，共通する言葉を多く含む文書のグループを，多変量解析によって自動的に発見・分類するためにコンピュータを用いるアプローチ」である。これは，Correlational アプローチと呼ばれる。

両アプローチは「いずれもコンピュータを利用して言葉や文書を分類し，その結果を計量的に分析するという点では似通っている。しかし分類の基準を分析者が自ら指定するのか，それとも多変量解析に分類を任せてしまうのかという点で大きな違いがある」（同上：強調は引用者による）。つまり，前者のアプローチは，分析者の意図・発想・観点を出発点としてデータを分析するのに対して，後者のアプローチは，生のデータを多変量解析によって機械的に要約するという分析である[5]。

経済学説史・経済思想史の研究においては，分析者の問題設定や独自の意図や観点に基づいてテキストと向き合うことがほとんどであろう。そのため，テキストマイニングを用いるならば，前者の Dictionary-based アプローチの方が適用しやすいのかもしれない。しかし，後者のアプロ

3）Cross-Industry Standard Process for Data Mining. 1996 年にクライスラー社，SPSS 社，NCR 社の代表的な研究者たちが発展させた手法である（喜田 2018 : 55-60）。ただし，これらはビジネスにおけるデータマイニングの活用に主眼が置かれている。

4）3.1 と 3.2 の大部分は樋口（2014 : 17-29）を参照している。

5）牛澤（2018 : 22）は前者を「仮説検証的な分析」，後者を「探索的な分析」と呼んでいる。

ーチの主張によれば，データに対して主観的な意図を混ぜることは，データを「汚染 contaminate」することになるという（Iker & Harway 1969）。つまり，テキストというデータと向き合うときは，できるだけ分析者の意図は含めず，客観的な姿勢に徹さなければならない，と言うのである[6]。確かに，分析者の意図や観点を含めすぎると歪んだ解釈に陥る可能性がある。しかし，情報処理や統計学に基づいた解析だけでは，分析者の問題設定に合わせた柔軟な分析ができなくなるであろう。

このように，Dictionary-based アプローチと Correlational アプローチのいずれにも一長一短がある。そこで樋口氏は，両者のアプローチ，すなわち，分析者の作成した基準に沿ってデータを分析するアプローチと，多変量解析を用いて機械的に分析するアプローチとを併用することで両者を相互に補完する「接合アプローチ」を提案している。

3.2 接合アプローチの利点

樋口（2014 : 18-19）は「接合アプローチ」の利点を以下のように説明している。

Dictionary-based アプローチの長所は，分析者の問題設定に合わせて多方面からテキストデータを分析できることである。しかし，分析者の独自の意図や観点が強調されすぎてしまうと，都合の悪いデータが排除され，都合のよいデータだけで分析してしまい，信頼性が損なわれる可能性がある。

このような短所を，Correlational アプローチによって補う。すなわち，テキスト全体に関するデータを，まずは機械的な多変量解析によって提示する。そして，その手順を終えてから，Dictionary-based アプローチによる 2 段階目の分析を進める。このように，まず機械的な解析結果を提示することで，どの段階で分析者の意図・発想・観点が入ったのかを第 3 者が確認することができ，ひいては信頼性の確保につながるであろう。

Correlational アプローチの長所は大量のデータを機械的に扱う点にあるのだが，その反面，分析結果に対して分析者が抱く様々な発想や関心のすべてには対応することができない。分析者の意図や観点の入り込む余地が非常に小さいからである。そこで，Dictionary-based アプローチによる探索的な分析手法を組み合わせることが，この短所を補うであろう。

以上のように，樋口氏は，両アプローチの峻別および接合の必要性を主張し，次のような 2 段階からなる「接合アプローチ」を提案している（樋口 2014 : 19, 25）。

段階 1　Correlational アプローチにならい，多変量解析を用いることで，分析者のもつ理論や問題意識の影響を極力受けない形で，データを要約・提示する。

段階 2　Dictionary-based アプローチにならい，コーディングルールを作成することで，明示的に理論仮設の検証や問題意識の追求を行う。

6）ただし，Correlational アプローチにおいても，無意識のうちに分析者の理論や予断をデータに押し付けてしまうことがあるようだ（樋口 2014 : 28）。

つまり，「接合アプローチ」の段階1では機械的な解析結果の提示に徹し，段階2において分析者の問題設定を加味した分析を行うのである。樋口氏が製作・改良を続けているテキストマイニング用ソフトウェア「KH Coder」は，このような「接合アプローチ」のアイデアに基づいて構築されている。

3.3 「接合アプローチ」の経済学史研究への統合

以上のような樋口氏の「接合アプローチ」を，経済学史研究にどのように応用すればよいだろうか。本稿では，4段階からなる分析方法を提案する。以下で，その方法について説明を加えよう。

まず，第1段階は「柔らかな作業仮説」である。この段階では，当該テーマに関する先行研究をふまえて，分析者自身の問題意識に基づいた仮説を明らかにする。これは，従来の経済学史研究の方法でも用いられてきたことであり，かつ，テキストマイニングにおいても重視される作業である（小峯・下平 2017 : 15）。

続いて，第2段階は「機械的な量的分析」[7]，第3段階は「質的カテゴリーを加えた量的分析」である。これらの両段階は，それぞれ，樋口氏の「接合アプローチ」を適用する。「機械的な量的分析」では，分析者の問題意識を極力加えず，ただコンピュータによってテキストを解析する。「質的カテゴリーを加えた量的分析」では，コーディングルールという形で分析者の問題意識を交え，その上でテキストを解析する。

そして，第4段階は「原典による引証」である。ここでは，第2段階と第3段階の解析結果をふまえて，原典の内容分析を行い，分析者の主張を根拠づける。この作業は，まさに従来の経済学史の研究方法である。

このような分析の流れであれば，従来の経済学史研究の方法とテキストマイニングの方法を結びつけることができる。さらに，これら4つの段階を踏むことによって，分析者の主張がより説得的になるという意味で，従来の経済学史研究の方法をより強化できるのではないだろうか。

次節では，以上のような分析方法を実際に用いて，マクミラン委員会におけるロバートソンの言説を分析する。

Ⅳ．マクミラン委員会におけるロバートソン証言の特色

4.1 柔らかな作業仮説

マクミラン委員会におけるロバートソンの言説に関する先行研究は，本稿第2節で示したとお

7）「質的」と「量的」の定義は見かけ以上に困難である。樋口（2014 : i）は「数値化されていない文章のようなデータ」を質的データと呼ぶ。他方，大量のデータを無差別に解析することを量的分析，少数のデータに特別に注目する分析を質的分析と呼ぶこともある。尺度では，便宜的に付けられた数値や順序のみに意味がある数値を質的データとも呼ぶ。なお，本稿では，数値化・構造化したデータを分析することを「量的分析」と呼んでいる。

りである。すなわち，Presley（1979）は，ケインズ政策との関連性という問題設定の上でロバートソンの証言を検討しており，また，Fletcher（2007）は，主にロバートソンの銀行論を補強するために彼の証言を検討していた。つまり，両者とも先に問題設定があり，それに基づいてロバートソンの言説を分析するという通常の経済学史の方法論に則っている。しかし，ここでは先行研究の消化や当該原典の熟読を前提としながら，なおいったん分析者の意図や関心から切り離された形の，つまりテキストの量的解析を挟み込むという方法論を採ってみよう。

筆者の1人のこれまでの主張は，次のとおりである。

ロバートソンによれば，経済学は人間の活動を研究する学問である（仲北浦 2017 c）。人間は自発的な意思決定により，自らの「努力」量を決定して経済活動（生産・消費）を行っている。そして，「努力」というすべての人間に共通する心身的な苦痛こそ価値の源泉であり，それが財を構成する。つまり，「努力」の増減が，直接，財の生産量・消費量に影響を及ぼす。よって，経済は，決して貨幣を中心として動いているのではなく，実物としての人間の意思決定とそれに基づく活動そのものを根本とするのである。このような思想ゆえに，ロバートソン経済学は，貨幣理論を展開する場合においても，実物理論が通底しているのである（仲北浦 2018）。

ゆえに，マクミラン委員会におけるロバートソンの証言には，次のような特徴があると予想できる。

（1）ロバートソンは，金融の議論においても，実物理論を重視した。
（2）ロバートソンは，経済現象の説明において，経済主体の意思決定とそれに基づく経済活動を重視した。

このような柔らかな作業仮説を念頭に置いた上で，以下の分析に取りかかる。

4.2 機械的な量的分析（「接合アプローチ」段階1）

本節では，第2段階として，『金融および産業に関する委員会報告書』（以下『報告書』と省略する）とロバートソンの証言資料である「世界不況」のテキストデータを多変量解析する。この段階では，できるかぎり分析者の意図，観点，問題意識を混入させないように留意する。

この段階における具体的な分析内容は，テキストデータ全体において多く出現した語を確認する「頻度分析」，語と語の結びつきや中心的な語を探る「階層的クラスター分析」および「共起ネットワーク分析」，そして，テキストの部分（部や章など）ごとの特徴語を探る「対応分析」である。これらの各分析を「KH Coder[8)]」を用いて行う。このソフトウェアでは，上記の分析に加えて「複合語の検出」，「多次元尺度構成法」，「自己組織化マップ」も可能であるが，本稿では用いない。使用しない理由はそれほど自明ではないが，紙幅や時間の関係という消極的な理由だ

8）本研究では2018年8月時点での最新バージョンである「KH Coder 3」（Windows版）を用いている。http : //khcoder.net/dl3.html にてダウンロードできる。

けでなく，とくにここで取り上げる分析が，視覚化に最も優れていて，経済学史の研究にもより適合的だという積極的な理由もある。なお，抽出された語が元のテキストでどのように用いられているかを知るために，原文に戻れるコマンドである「KWIC コンコーダンス」[9]も適宜，用いている。

一般に，データマイニングでは，①データを概観する，②相関関係を探る，③回帰分析を行う，という3段階がある。第1は，平均値・分散・中央値・最頻値などの基本的指標をまず押さえ，頻度の分布などをヒストグラムとして描くことである。第2は，特に2つの変数について，互いに関連があるかどうかを分析することである。第3は，相関関係だけでなく因果関係が予想される場合に，因果関係の入力（説明変数）と出力（目的変数）の関係を関数の形で求めることである（山内 2017 : 63, 70, 73）。本稿ではデータを概観し，相関関係を探る段階を分析している。

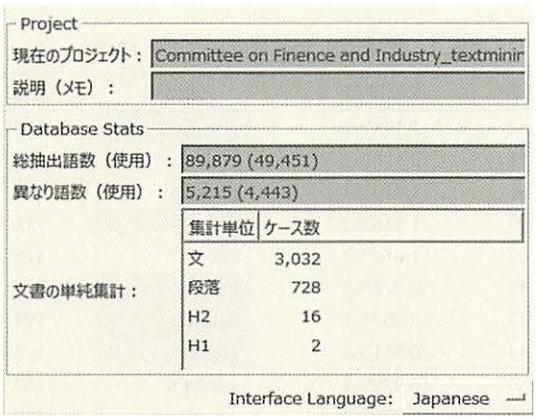
なお，本研究では，各品詞の中で最も理念・概念・思想を象徴する「名詞」のみを扱って分析する（喜田 2018 : 210；山内 2017 : 140）。また，以下では書籍の PDF ファイル化および OCR からテキストファイル化までが適切に済まされていることを前提としている。テキストファイル化までの手順については補論（1）を参照されたい。

4.2.1 Database Stats

まず，『報告書』と「世界不況」という2種類のテキストの総語数を確認しておく。KH Coder においては，「前処理」を終えた段階で，「Database Stats」を見ることができる（図表1および図表2）。「総抽出語数」とは，対象テキスト内に出現した語の延べ数であり，「異なり語数」とは，出現した語の種類数である。また，「(使用)」とは，分析に使用された語数を意味する。分析の対象として選択できるのは，名詞，固有名詞，外来語，代名詞，形容詞，副詞，動詞，疑問詞である。本研究では名詞のみを抽出しているので，「(使用)」内の数値は，名詞の語数である。

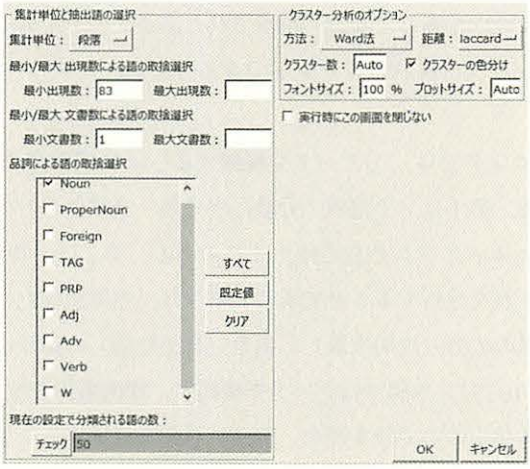
以下に，『報告書』と「世界不況」の Database Stats を示す。

図表1　『報告書』の Database Stats



9）Key Words in Context Concordance。コンコーダンスとは用語索引のことである。

図表 2 「世界不況」の Database Stats



『報告書』の総抽出語数は 89,879 語であり，そのうち分析に使用された語数は 49,451 語である。他方，ロバートソンの「世界不況」はそれぞれ 8,512 語と 4,704 語である。このように，両テキストの分量には 10 倍以上の差がある。ゆえに，以下の分析では，語の絶対数ではなく相対数（割合，%）に注目することとする。

4.2.2 頻度分析

まず，頻度分析を行う。対象テキストにおいて，どのような名詞がどれくらいの頻度で出現しているかを確認することで，そのテキストの全体像をとらえることができる。以下では，マクミラン委員会の『報告書』とロバートソンの「世界不況」における名詞の出現頻度（上位 50 語）を提示する。そして，両者に〈共通している語〉と〈共通していない語〉を分類し，両テキストにおける中心的な論題の異同を検討する。

以下に，『報告書』の名詞の出現頻度（名詞 1706 語中上位 50 語）と，ロバートソンの「世界不況」における名詞の出現頻度（名詞 590 語中上位 48 語）を示す。

図表 3 『報告書』の頻度分析（上位 50 語）

抽出語	語数	出現割合	抽出語	語数	出現割合
bank	462	(2.388%)	policy	125	(0.646%)
gold	421	(2.176%)	power	125	(0.646%)
country	356	(1.840%)	position	122	(0.631%)
price	299	(1.546%)	part	118	(0.610%)
credit	204	(1.054%)	effect	115	(0.594%)
time	188	(0.972%)	loan	114	(0.589%)
issue	187	(0.967%)	balance	104	(0.538%)
rate	184	(0.951%)	capital	103	(0.532%)
industry	180	(0.930%)	standard	102	(0.527%)
world	179	(0.925%)	condition	100	(0.517%)
market	176	(0.910%)	business	99	(0.512%)
deposit	166	(0.858%)	cost	99	(0.512%)

note	157	(0.812%)	fact	98	(0.507%)
year	156	(0.806%)	currency	96	(0.496%)
money	154	(0.796%)	security	96	(0.496%)
figure	151	(0.781%)	stock	96	(0.496%)
system	151	(0.781%)	increase	92	(0.476%)
reserve	148	(0.765%)	cash	91	(0.470%)
volume	148	(0.765%)	result	90	(0.465%)
amount	145	(0.750%)	banking	89	(0.460%)
case	143	(0.739%)	bill	89	(0.460%)
percent	135	(0.698%)	million	88	(0.455%)
level	128	(0.662%)	interest	87	(0.450%)
change	127	(0.656%)	value	85	(0.439%)
period	126	(0.651%)	way	84	(0.434%)

図表4「世界不況」の頻度分析（上位48語）

抽出語	語数	出現割合	抽出語	語数	出現割合
banking	33	(1.875%)	proportion	13	(0.739%)
system	30	(1.705%)	case	12	(0.682%)
bank	27	(1.534%)	change	12	(0.682%)
trade	27	(1.534%)	depression	12	(0.682%)
gold	22	(1.250%)	result	12	(0.682%)
country	21	(1.193%)	authority	11	(0.625%)
deposit	21	(1.193%)	industry	11	(0.625%)
goods	20	(1.136%)	public	11	(0.625%)
money	20	(1.136%)	asset	10	(0.568%)
part	18	(1.023%)	cause	10	(0.568%)
investment	17	(0.966%)	difficulty	10	(0.568%)
production	17	(0.966%)	reason	10	(0.568%)
time	17	(0.966%)	reserve	10	(0.568%)
volume	17	(0.966%)	term	10	(0.568%)
capital	16	(0.909%)	thing	10	(0.568%)
export	16	(0.909%)	balance	9	(0.511%)
price	15	(0.852%)	circumstance	9	(0.511%)
year	15	(0.852%)	consumption	9	(0.511%)
income	14	(0.795%)	need	9	(0.511%)
loan	14	(0.795%)	policy	9	(0.511%)
wages	14	(0.795%)	view	9	(0.511%)
world	14	(0.795%)	way	9	(0.511%)
demand	13	(0.739%)	work	9	(0.511%)
level	13	(0.739%)			
percent	13	(0.739%)			

図表内の網掛けは，『報告書』（上位50語）と「世界不況」（上位48語）で〈共通している語〉を示している。高頻度の名詞のうち，およそ半分の25語が共通している。〈共通している語〉である balance, bank, capital, deposit, gold, industry, loan, money, reserve から，マクミラン委員会では「金融および産業に関する委員会」という名のとおり，やはり金融と産業が中心的な話題となっており，その点ではロバートソンの議論もその線に沿っていたことがうかがえる[10]。また，共

通して高頻度の語 balance, bank, capital, deposit, gold, industry, loan, money, reserve のうち，balance, deposit, reserve は，貨幣全体のうち，市場に流通しない部分を示している。ロバートソンは，金融論の中でも，これらの語を重視していたことがうかがえるが，それは保蔵の議論とも関係しているからかもしれない。

では，〈共通していない語〉はどのようになっているのだろうか。以下に，『報告書』では高頻度だが「世界不況」では低頻度の語（図表 5）と，「世界不況」では高頻度だが『報告書』では低頻度の語（図表 6）を示す。その際，出現割合に注目することとし，両者の出現頻度に比較的大きな差がある語（出現割合が 5 倍以上）には網掛けを施す。

図表 5 『報告書』（高頻度）・「世界不況」（低頻度）

	『報告書』		「世界不況」		差
credit	204	(1.054%)	4	(0.227%)	4.640
issue	187	(0.967%)	2	(0.114%)	8.506
rate	184	(0.951%)	5	(0.284%)	3.348
market	176	(0.910%)	2	(0.114%)	8.006
note	157	(0.812%)	2	(0.114%)	7.142
figure	151	(0.781%)	5	(0.284%)	2.747
amount	145	(0.750%)	2	(0.114%)	6.596
period	126	(0.651%)	1	(0.057%)	11.463
power	125	(0.646%)	7	(0.398%)	1.625
position	122	(0.631%)	3	(0.170%)	3.700
effect	115	(0.594%)	4	(0.227%)	2.616
standard	102	(0.527%)	7	(0.398%)	1.326
condition	100	(0.517%)	8	(0.455%)	1.137
business	99	(0.512%)	2	(0.114%)	4.503
cost	99	(0.512%)	6	(0.341%)	1.501
fact	98	(0.507%)	6	(0.341%)	1.486
currency	96	(0.496%)	0	(0.000%)	−
security	96	(0.496%)	2	(0.114%)	4.367
stock	96	(0.496%)	1	(0.057%)	8.734
increase	92	(0.476%)	6	(0.341%)	1.395
cash	91	(0.470%)	1	(0.057%)	8.279
bill	89	(0.460%)	0	(0.000%)	−
million	88	(0.455%)	1	(0.057%)	8.006
interest	87	(0.450%)	4	(0.227%)	1.979
value	85	(0.439%)	2	(0.114%)	3.866

まず，『報告書』では高頻度だが「世界不況」では低頻度だった語のうち，出現割合の差が大きいものは，差が大きい順に period, stock, issue, cash, market, million, note, amount であり，currency, bill に関しては「世界不況」では出現していない。stock, issue, cash, note, currency, bill は

10) このような分析結果は当然と言うべきかもしれない。しかし，量的分析によって，その当然の結果が得られたこともまた，重要な発見と捉えるべきであろう。目を引くような結果のみを追求してしまう「出版バイアス publication bias」はなるべく避けなければならない（cf. 伊藤 2017 : 249-252）。

金融問題において重要な語と考えられるが，ロバートソンは，それらの語をほとんど用いていないのである。「世界不況」で money という語が多用されていることを考慮すると，ロバートソンは cash, note, currency, bill という貨幣に関する語を money で統一している可能性がある。

図表 6 「世界不況」（高頻度）・『報告書』（低頻度）

	「世界不況」		『報告書』		差
trade	27	(1.534%)	66	(0.341%)	4.497
goods	20	(1.136%)	27	(0.140%)	8.142
investment	17	(0.966%)	80	(0.414%)	2.336
production	17	(0.966%)	81	(0.419%)	2.307
export	16	(0.909%)	63	(0.326%)	2.792
income	14	(0.795%)	50	(0.258%)	3.078
wages	14	(0.795%)	32	(0.165%)	4.809
demand	13	(0.739%)	77	(0.398%)	1.856
proportion	13	(0.739%)	51	(0.264%)	2.802
depression	12	(0.682%)	46	(0.238%)	2.867
authority	11	(0.625%)	40	(0.207%)	3.023
public	11	(0.625%)	30	(0.155%)	4.030
cause	10	(0.568%)	56	(0.289%)	1.963
difficulty	10	(0.568%)	63	(0.326%)	1.745
reason	10	(0.568%)	55	(0.284%)	1.999
term	10	(0.568%)	51	(0.264%)	2.155
thing	10	(0.568%)	10	(0.052%)	10.992
circumstance	9	(0.511%)	61	(0.315%)	1.622
consumption	9	(0.511%)	8	(0.041%)	12.366
need	9	(0.511%)	24	(0.124%)	4.122
view	9	(0.511%)	74	(0.383%)	1.337
work	9	(0.511%)	25	(0.129%)	3.957

他方，「世界不況」では高頻度だが『報告書』では低頻度だった語のうち，出現割合の差が大きいものは，差が大きい順に consumption, thing, goods であった。ここから，ロバートソンは，貨幣という論題において，財市場における goods, thing, あるいは，その消費活動を意味する consumption を強調したことが分かる。

以上のように，上位 50 語（名詞）の頻度分析だけでも，マクミラン委員会とロバートソンの証言における中心的な論題をだいたい把握することができたであろう。さらに，出現割合に大きな差がある語に注目することで，ロバートソンは，goods や thing といった実物の財，あるいは，その消費活動としての consumption といった実体経済を重視していたことが垣間見られた。

4.2.3 階層的クラスター分析[11)]

次に，階層的クラスター分析を行う。クラスター分析とは，個々のデータの類似度をある種の距離として表現し，距離の近いデータ同士をまとめて房（クラスター）を作ることで，全体をい

11) クラスター分析については，Romesburg (1984)（西田・佐藤訳）を主に参照した。本著作では，クラスター分析について，できるかぎり数式ではなく言葉での説明が試みられている。

くつかの部分にグループ化することである。類似度が近い房を統合していき，最後に1つのクラスターになるまで繰り返す分析を階層的クラスター分析と呼び，その結果は，樹形図 dendrogram として表現される（小峯・下平 2017 : 38-39）。

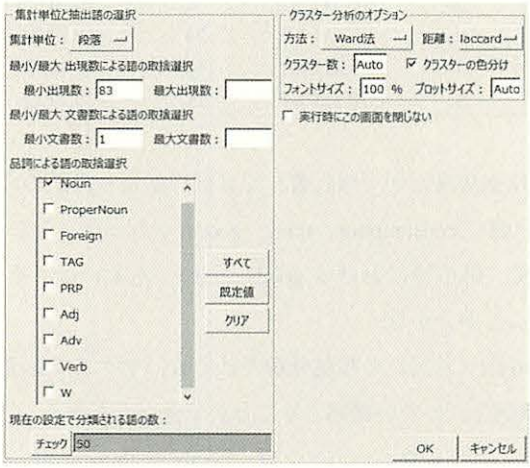
先の頻度分析で用いた上位 50 語において，出現パターンが類似している語群をいくつかのクラスターに分ける。ここで，集計単位は「段落」にしている。そうすることで，同じ段落の中でよく一緒に出現した語の組み合わせほど，出現パターンが似ているとみなされるようになる（樋口 2014 : 37）。また，ここでは分析する語数を 50 語に調整するため，図表7左上の「最小出現数」を「83」に設定した。図表7右側の「クラスター分析のオプション」においては，クラスター化法を「Ward（ウォード）法」[12]に設定し，距離係数の算出法を「Jaccard（ジャッカール）」[13]に設定している。

クラスター数について1つだけ留意すべき点を述べておく。クラスター数は通常，図表7右上の「クラスター数」にあるように「Auto」で実行すればよい。しかし，Auto によるクラスター数が最適というわけではない。適当なクラスター数を決定するためには，「併合水準」を参照する必要がある。

クラスター数を「Auto」にして分析すると，図表8のようにクラスター数は7つになる。ただし，このクラスター数が最適というわけではない。図表8左下にある「併合水準」コマンドを実行することで，クラスター分析を行った際の「併合水準」を確認することができる（図表9）。

では，「併合水準」とは何か。階層的クラスター分析において，最初のクラスター数は，分析

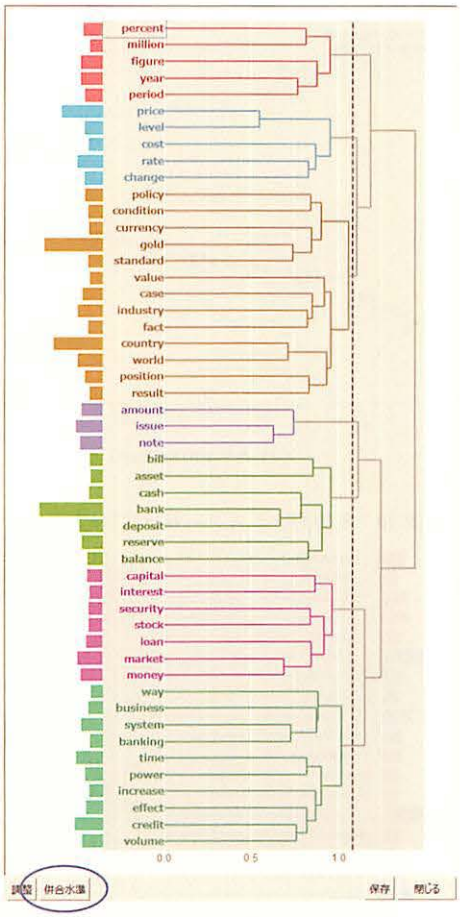
図表7　階層的クラスター分析：オプション画面



12)「Ward 法」が最も広く用いられているが，その理由は，その方法が最適だからではなく，1967 年という早い段階から，その方法のためのコンピュータプログラムが利用できたからであるようだ（Romesburg 1984 : 135／訳 170）。KH Coder では他にも「群平均法」（UPGMA）と「最遠隣法」（CLINK）を選択することができる。これらの計算方法の相違については，Romesburg（1984 : 119-129／訳 151-160）を参照のこと。

13) KH Coder では他にも「Cosine」と「Euclid」を選択することができる。これらの計算方法の相違については，Romesburg（1984 : 96-104／訳 122-131）を参照のこと。石田・金編（2012 : 12）や樋口（2014 : 112）は Jaccard 係数を勧めているが，後藤（2016 : 3. 3 節）は Eucild 係数を勧めている。

図表8 「併合水準」コマンド

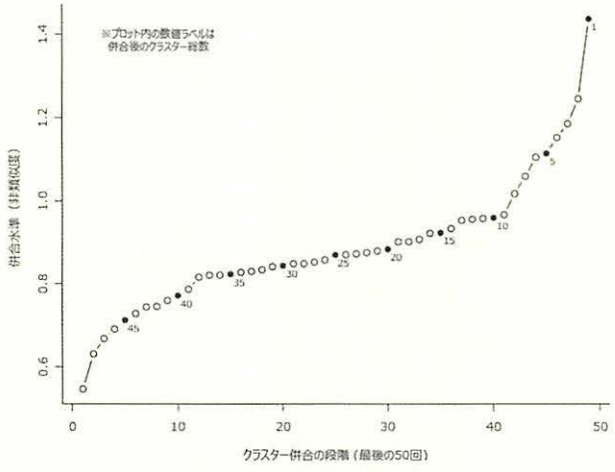


する語の数だけ存在する。つまり，分析対象が50語ならば，最初のクラスター数は50である。これを起点として出現パターンが類似する語を1つずつ併合していくことで，クラスター数が減っていき，果てにはクラスター数は1つになる（Romesburg 1984 : 14-23／訳 18-28）。そのため，併合が進むにつれて（クラスター数が減っていくにつれて），しだいに併合は難しくなっていく。このような，併合の難度，言い換えれば，併合するときにどれだけ無理をしているか，これを数値化した指標こそ「併合水準」である。この数値が高ければ高いほど，併合の際に無理をしていることになる。

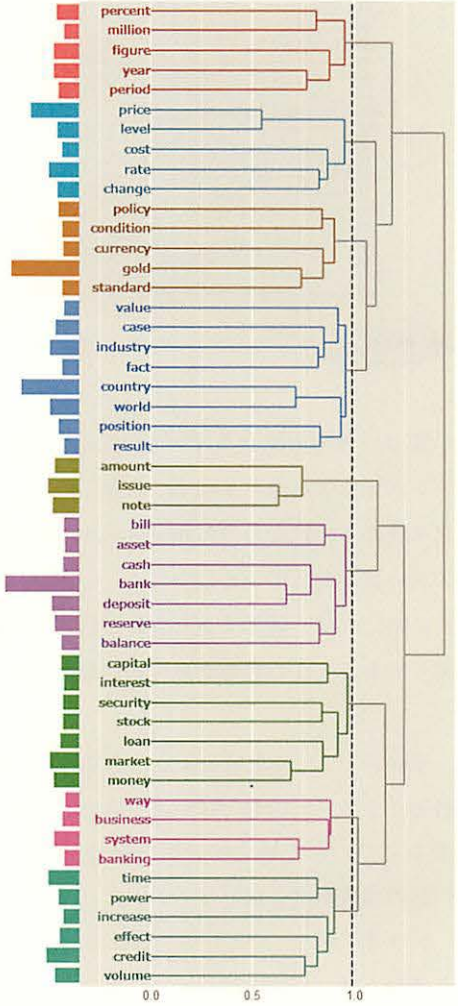
よって，図表9を見れば，Auto（7つ）の水準よりも，クラスター数が9つのときの方が併合水準は低いため，今回の階層的クラスター分析では，クラスター数は9つにした方がよい，ということになる。クラスター数を9つにした場合の分析結果は，図表10のとおりである。なお，図表左側の棒グラフは，その語の出現頻度を表している。

上から順にクラスター1，クラスター2，…として，各クラスターの解釈を行う。ただし，ここで行う解釈は，あくまでも予測を伴った暫定的なものであり，必ずしも適切なものとは言えな

図表 9　併合水準



図表 10　階層的クラスター分析（『報告書』）

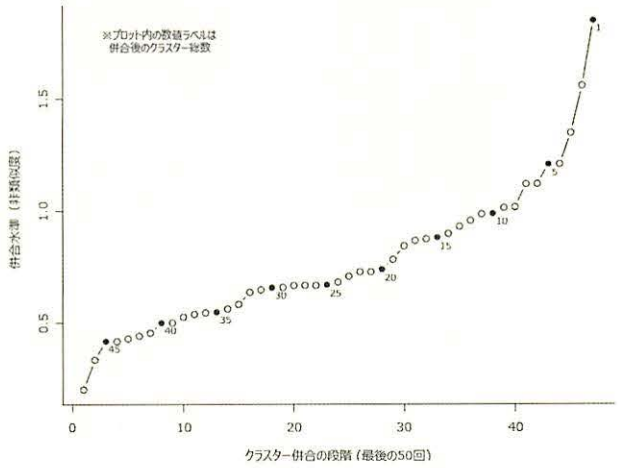


い。深読みをしないように留意しなければならない（樋口 2014, 38）。

クラスター 1：統計的数値（percent, million, figure, year, period）
クラスター 2：物価水準と生産費用（price, level, cost, rate, change）
クラスター 3：金本位制と通貨政策（policy, condition, currency, gold, standard）
クラスター 4：産業の国際的地位（value, case, industry, fact, county, world, position, result）
クラスター 5：銀行券の発行量（amount, issue, note）
クラスター 6：銀行準備と資産残高（bill, asset, cash, bank, deposit, reserve, balance）
クラスター 7：貨幣市場と資産市場（capital, interest, security, stock, loan, market, money）
クラスター 8：銀行組織による貸付（way, business, system, banking）
クラスター 9：信用創造とその効果（time, power, increase, effect, credit, volume）

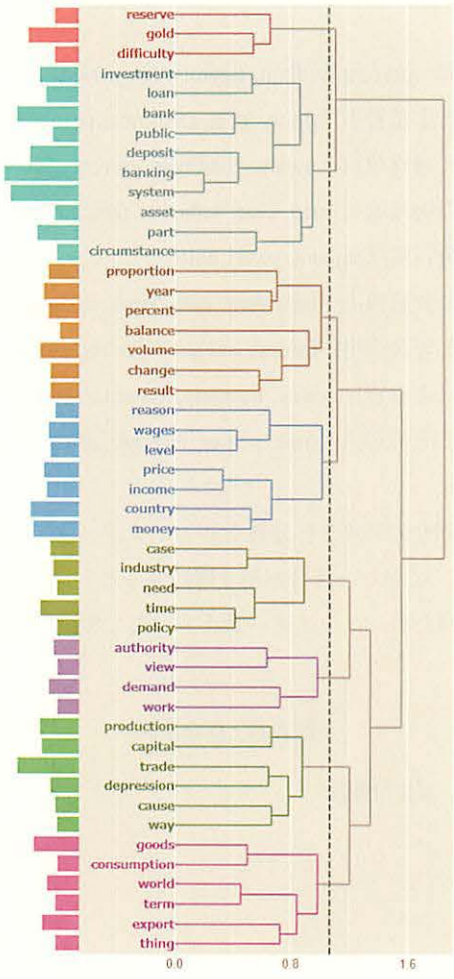
次に，「世界不況」の階層的クラスター分析を行う。クラスター数を Auto に設定すると，クラスター数は 7 つになった。しかし，併合水準を確認すると，クラスター数は 8 つの方がよいことが分かる（図表 11）。そのため，クラスター数を 8 つとして階層的クラスター分析を行った。以下に解析結果を示す。

図表 11　併合水準



クラスター 1：金準備率（reserve, gold, difficulty）
クラスター 2：預金・融資・投資（investment, loan, bank, public, deposit, banking, system, asset, part, circumstance）
クラスター 3：残高の状況（proportion, year, percent, balance, volume, change, result）
クラスター 4：所得と物価水準（reason, wages, level, price, income, country, money）
クラスター 5：産業の貨幣需要と融資方針（case, industry, need, time, policy）
クラスター 6：中央当局による需要管理（authority, view, demand, work）

図表 12　階層的クラスター分析（「世界不況」）



クラスター 7：固定資本と産業不況（production, capital, trade, depression, cause, way）

クラスター 8：財の国内消費と輸出（goods, consumption, world, term, export, thing）

『報告書』と「世界不況」の階層的クラスター分析の結果から，両者を比較検討しよう。頻度分析でも明らかになったように，金融が中心的な議題であることは両テキストに共通した特徴であった。しかし，階層的クラスター分析では，金本位制，銀行準備，貨幣発行，資産残高，信用創造など，金融という広い領域をさらに細分化した形で論題を解釈することできる。こうして，テキストの全体像がより一層明瞭になってくるのである。

次に，両テキストに特徴的なクラスターを検討しよう。

『報告書』では，ほとんどすべてのクラスターで金融に関連する語が並んでいる。しかし，クラスター 4 では「産業の国際的地位（value, case, industry, fact, county, world, position, result）」について議論されているようだ。本稿 2.1 でも触れたように，マクミラン委員会は，英国産業の国

際的な地位低下を克服することを目的の1つとしていた。そのため，世界経済における英国の位置づけは，十分妥当な論題であろう。このように，『報告書』には，やはりマクミラン委員会の当初の目的に沿った論題が並んでいると考えられる。

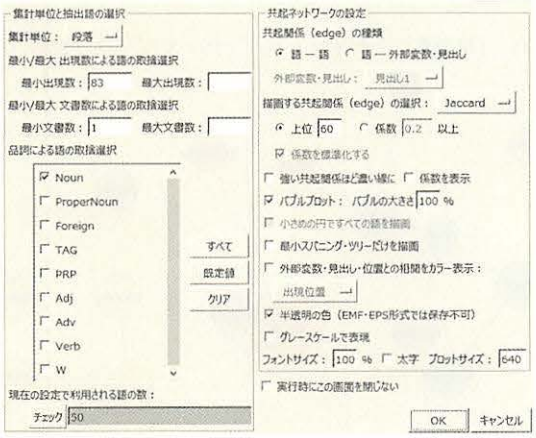
それに対して，ロバートソンの「世界不況」は特徴的なクラスターが複数ある。クラスター1～5は金融に関する論題であるが，クラスター6～8ではdemand, production, consumptionといった経済活動そのものに関する語が並んでいる。本文を精査しなければ確証を得られないが，少なくとも『報告書』では現れていない論題だということは確実である。また，頻度分析でもロバートソンの特徴として指摘したgoods, thing, consumptionの語が1つのクラスターを形成している（クラスター8）。やはり，実物的な財とその消費について，一塊の議論がなされていると推測される。

4.2.4　共起ネットワーク分析

次に，共起ネットワーク分析を行う（小峯・下平2017：43）。先の階層的クラスター分析は，類似の出現パターンによって語群をグループ分けした。だが，共起ネットワーク分析では，出現パターンの類似性（共起性）のみならず，各語のまとまりどうしの関係の強さや，テキスト全体において中心的な役割を果たしている語を視覚的に確認できる。

まず，『報告書』の共起ネットワーク分析の手順を以下に示す。

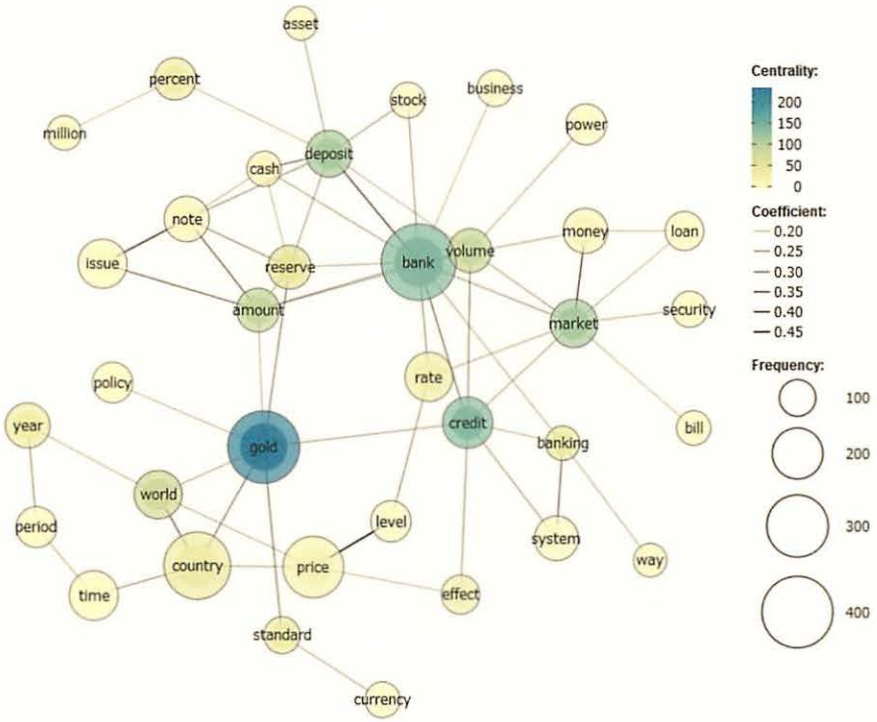
図表13　『報告書』共起ネットワーク：オプション画面



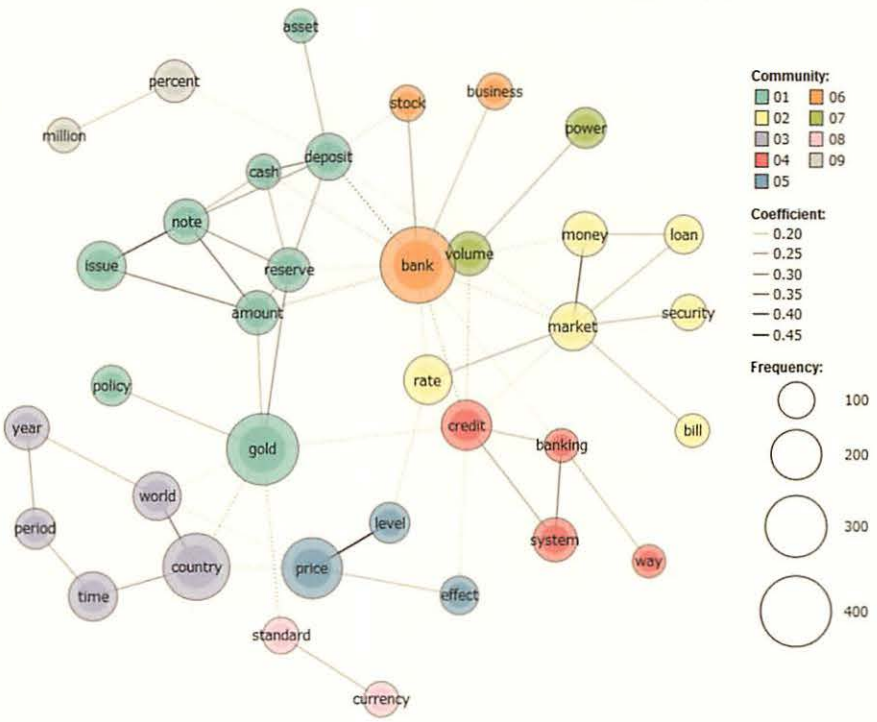
これまでと同様に，名詞50語を分析対象としている。また，図表13右側「共起ネットワークの設定」において，「強い共起関係ほど濃い線に」にチェックを入れる（図表13はチェック前）。

共起ネットワーク分析には，「中心性」と「サブグラフ検出」の2種類の図示方法がある。中心性は，テキストにおいて中心的な役割を担っている語を示す方法であり，色が濃いほど中心的な語であることを意味する。他方，サブグラフ検出では，グループ分けされた各要素間の関連の強さを示すことができる。よって，色が異なる語どうしが実線で結ばれている場合，それぞれのグループ間に何らかの強い関連があることを示している。

図表 14 『報告書』共起ネットワーク図：中心性（媒介）

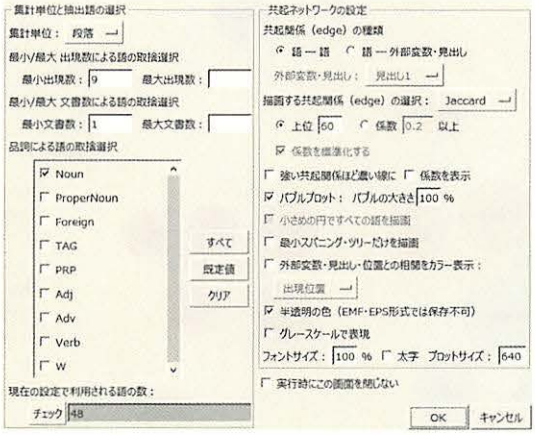


図表 15 『報告書』共起ネットワーク図：サブグラフ検出（媒介）

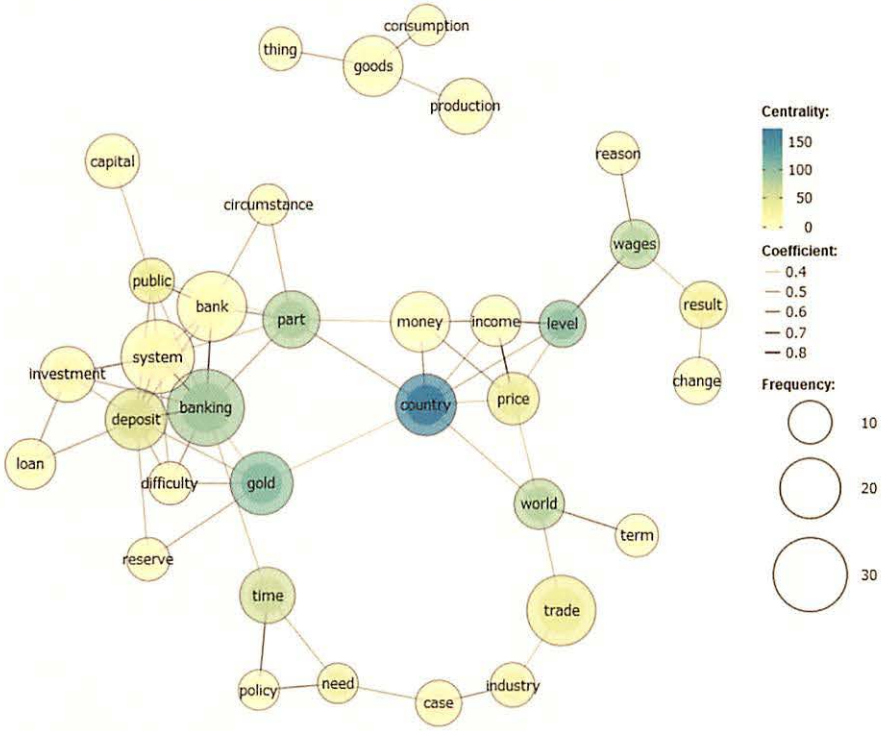


次に，「世界不況」の共起ネットワーク分析の結果を示す。

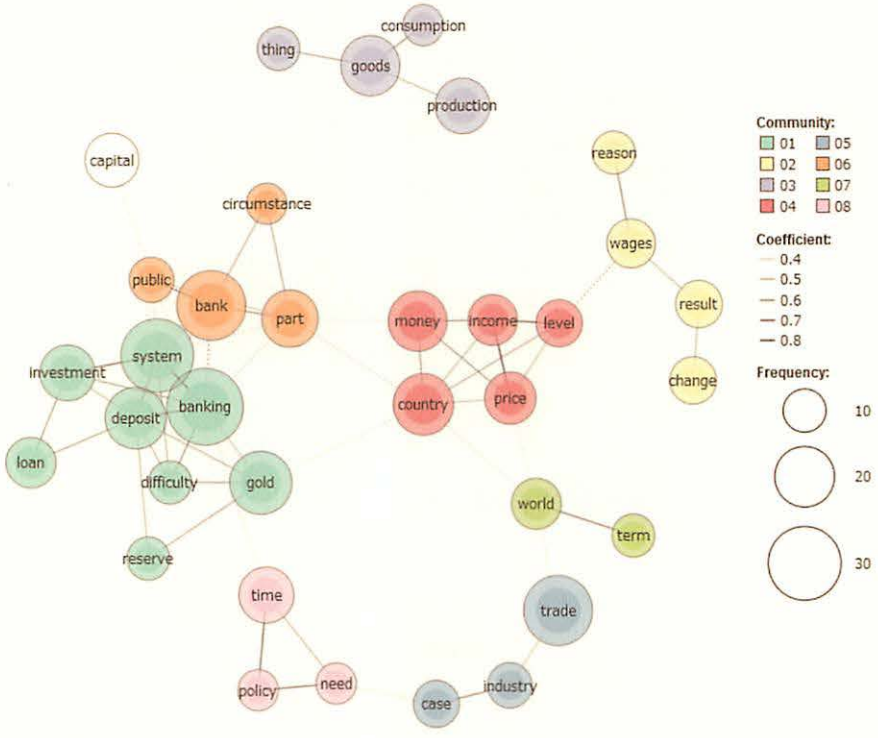
図表 16 「世界不況」共起ネットワーク：オプション画面



図表 17 「世界不況」共起ネットワーク図：中心性（媒介）



図表 18 「世界不況」共起ネットワーク図：サブグラフ検出（媒介）



両者の解析結果を比較することで，ロバートソンの「世界不況」について以下のように言えそうである。

『報告書』においては，図表 14 のとおり，最も中心的な語は gold であり，次に credit，bank，deposit，market，amount，world と続く。それに対して，「世界不況」においては，図表 17 のとおり，最も中心的な語は country[14]であり，次に gold，level，banking，wages と続く。確かに，ロバートソンの証言においても，gold や banking のように金融の問題が中心的なテーマとなっている。しかし，それに加えて，wages と level[15]が中心語に現れていることから，国民生活という実体経済にも主眼が置かれていたことがうかがえる。

また，「世界不況」のサブグラフ検出の図を見ても，その多くが金融と関連したコミュニティを形成し，相互に連関しあっている。しかし，図表 18 上部の goods，thing，consumption，produc-

14) この解析結果では，'country' がどの 'country' を指しているのかが分からない。そのため，KH Coder の「KWIC コンコーダンス」機能で 'country' という語を検索することで，その周辺の文を確認することができる。'country' を検索すると，'a country' など一般論としての「国家」を示している場合が大半を占めた（21 例中 17 例）。このことから，ロバートソンが国家を中心として経済を論じていたことが推測される。なお，他の用例は，海外の国々を表す 'European counties' や 'other countries'，英国を表す 'this country' であった。

15) 本稿脚注 14 と同様に，「KWIC コンコーダンス」にて 'level' を検索すると，計 13 例中，level of wages と level of income など，'level' が 'wage' と 'income' に関連づけられた使用が 9 例であったことから，'level' は主に賃金・所得の 'level' を指していることが分かった。

tion のコミュニティは，他のコミュニティとは独立して布置されている。頻度分析と階層的クラスター分析でも指摘したように，やはり goods, thing, consumption, production といった実体経済への言及はロバートソンの特徴とみなせそうである。

4.2.5　対応分析

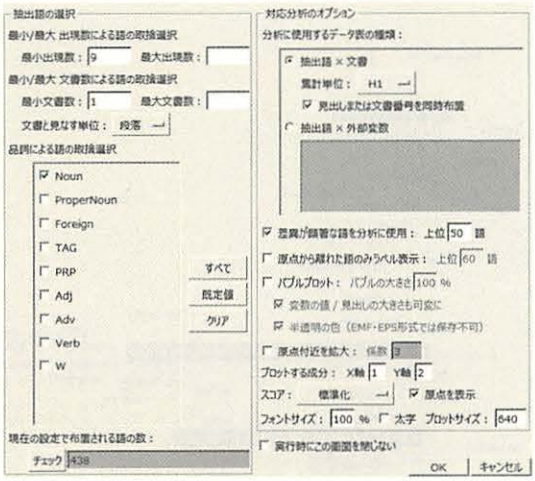
最後に対応（コレスポンデンス）分析を行う。対応分析とはクロス集計を視覚化する方策である。クロス集計とは，2つの項目を掛け合わせて集計し，関連性を分析することである（牛澤 2018 : 57）。この分析は，単語同士の類似関係や差異などのパターンを分類する面で優れている[16]。とくにテキストマイニングの対応分析は，テキスト内の部や章といったグループにおける特徴語を抽出するものである。KH Coder では，テキストファイルに〈h 1〉や〈h 2〉といったタグをつけることによって，テキスト内の文章グループを分類することができる。

本研究では，マクミラン委員会におけるロバートソンの証言の特徴に注目しているので，ここでは，ロバートソンの「世界不況」における特徴語を抽出する。

その際，まず『報告書』と「世界不況」のテキストファイルを1つのファイルに統合した。『報告書』の第Ⅰ部と第Ⅱ部にそれぞれ〈h 1〉のタグをつけ，「世界不況」のテキスト全体にも〈h 1〉のタグをつけることによって，統合ファイルを3分割した。

対応分析においては，図表 19 のとおり，分析対象とする語の「最小出現数」を設定することができる。今回は，『報告書』と「世界不況」の総語数が 10 倍近く異なっているので，最小出現数を大きく設定すると，「世界不況」の語が分析されなくなる。よって，図表 19 のとおり，最小出現数を「9」とした。ここまでの分析と同様に，「世界不況」に出現する名詞の上位 48 語を分析対象とするためである。

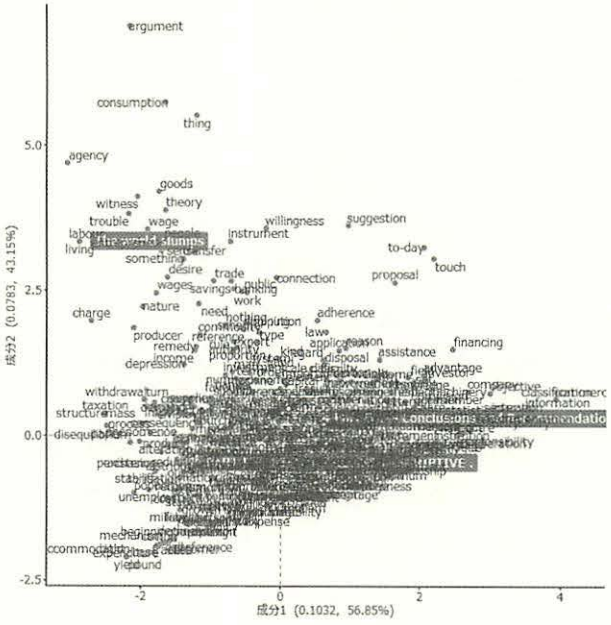
図表 19　対応分析：オプション画面



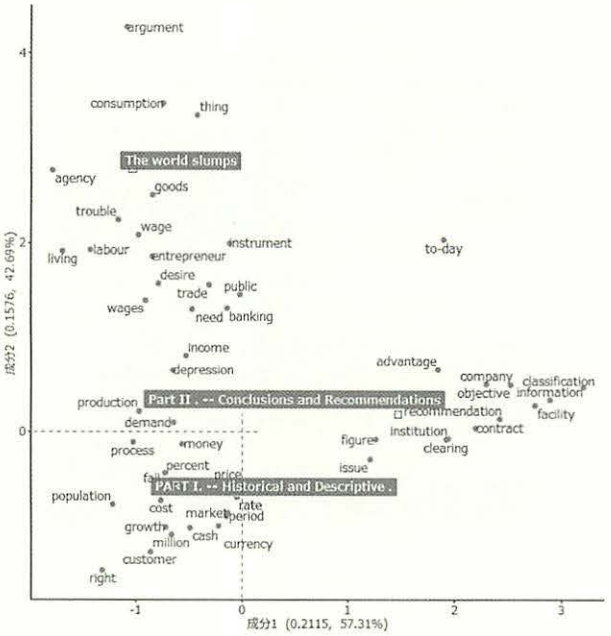
16）対照的に，主成分分析はデータを圧縮して関係性を見る（データの構造を再現する）場合に優れている（水本 2009 : 54, 63）。

ただし，最小出現数を「9」とすると，かえって『報告書』の分析対象が多くなる。ゆえに，図表 19 左下にある「現在の設定で布置される語の数」が 438 語となっている。このままでは大量の語が布置され散布図が真っ黒になってしまう（図表 20）。よって，図表 19 右中の「差異が顕著な語を分析に使用」にチェックを入れることで対処する。

図表 20　対応分析（失敗例）



図表 21　対応分析（差異上位 50 語）



この図において，原点（0, 0）に近く布置された語は，テキスト全体において広く出現した一般語である。逆に，原点から遠く布置された語ほど，そのテキストにおいて特徴的であることを意味する。互いに関連性の強い項目は，原点から見て同一方向に配置されている。

図表21左上を見ると，ロバートソンの「世界不況」においては，agency, consumption, desire, entrepreneur, goods, instrument, labour, living, thing, trouble, wage などが特徴語として出現したことが分かる。consumption, goods, thing はここまでの分析において繰り返し言及してきた。他の語に注目すると，agency, entrepreneur, labour といった人間（経済主体）を表す語，さらには，living, wage, desire といった経済生活に関する語が特徴語して現れていることが分かる。

さて，第2段階である「機械的な量的分析」の結果を，ロバートソンの特徴に注目してまとめておこう。

まず，すべての解析結果で共通していたのは，ロバートソンの「世界不況」では，goods, thing, consumption,（production）という実体経済に関する語が特徴的だということである。金融問題の議論では，資産市場・貨幣市場が中心的な話題となりそうだが，ロバートソンは，金融問題においても財市場を併せて考慮していたことが予想される。頻度分析では，ロバートソンが，貨幣に関連する語を money に統一している可能性があることと，彼の保蔵理論と関係すると考えられる balance, deposit, reserve を多用していたことが分かった。さらに，対応分析では，ロバートソンが実体経済のみならず，それを形成する人間（経済主体）とその活動や経済生活をも議論の俎上にのせている可能性があることが分かった。

4.3　質的カテゴリーを加えた量的分析（「接合アプローチ」段階2）

次の第3段階においては，分析者の観点をより前面に出した上で，テキストデータを分析する。すなわち，「コーディングルール」という質的なカテゴリーを作成し，報告者の問題設定に合わせてテキストマイニング分析を行うのである。

以下では，4.3.1 においてコーディングルールの作成手順を示し，4.3.2 において，本研究の問題設定である，マクミラン委員会におけるロバートソンの言説の特徴を検証する。

4.3.1　コーディングルールの作成

4.2 で行った第2段階の分析においても，次の2点が導き出せた。そしてこの結果は，4.1 で想定していた「柔らかな作業仮説」とも一致していることになる。

第1点は，ロバートソンが貨幣理論においても，やはり実物理論を重視していたのではないか，という点である。この論点は，第2段階で示した goods, thing という特徴語から着想を得た。ロバートソンは金融に関する議論においても財市場を併せて考慮していたのではないだろうか。

第2点は，ロバートソンが人間の動機とそれに基づく経済活動の分析を重視していたのではないか，という点である。これは，第2段階においてロバートソンの特徴語と予測された consumption, production といった経済活動に関する語や agency, entrepreneur, labour といった経済主

体に関する語から想起した。

このような結果を踏まえて，さらに第3段階「質的カテゴリーを加えた量的分析」を行う。すなわち，(1) 実物理論，(2) 経済主体，(3) 動機・活動，に関連する語をまとめた質的なカテゴリーを作成する。このように，同じ属性をもつと考えられる語を1つのカテゴリーにまとめることをコーディングといい，その基準のことを「コーディングルール」という（樋口 2014 : 7）。このような処理を行うことで，低頻度の語であっても分析対象として掬いとることができるのである。

ただし，コーディングルールの作成には，分析者の独自の意図・発想・観点が多分に含まれている。よって，可能な限りどのような判断に基づいてルールを設定したかを示しておくのがよいであろう[17]。なお，ここでは，コーディングの対象となる語を決定する際に『報告書』と「世界不況」との両方のテキストデータを参照して検討した[18]。

(1) 実物理論（real theory）
- 実物的な財：goods, thing, commodity, instrument, machinery
- 財の生産：product, output, yield

(2) 経済主体（economic agent）
- 人間一般：human, person, people, man, public
- 生産関連：entrepreneur, industrialist, labour, manufacturer, producer, worker, employer
- 消費関連：buyer, consumer, customer
- 流通関連：dealer, importer, exporter, seller, trader
- 金融関連：banker, borrower, broker, creditor, debtor, financier, investor, rentier

(3) 動機・活動
- 動機：desire, want, will, willingness, wish
- 活動：banking, consumption, financing, investment, production, saving

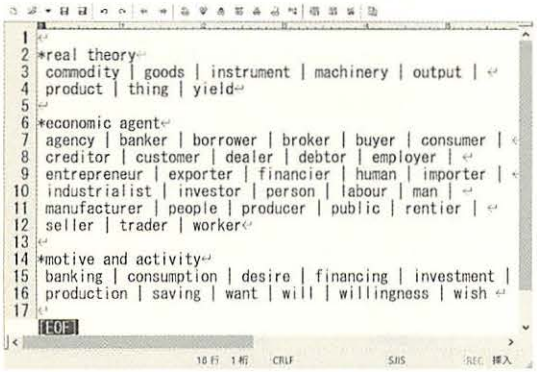
これらのコーディングルールは，新たなテキストファイルに当該の語を入力することで作成する（図表22）。図表22において，「*」を付しているのがコード名である。すなわち，今回は，「*real theory」，「*economic agent」，「*motive and behaviour」という3つのコードを作成した。なお，「｜」の入力は「or」を意味する[19]。

17) 喜田（2018 : 116-129）も，分析対象とする言葉のカテゴリー設定を行っている。ここでも客観性や信頼性という問題が言及されているが，共同研究者3名で「議論を尽くす」ことで，「カテゴリーの妥当性と信頼性を確保」したと主張されている。3者での合意が信頼性を担保するのかどうか，また，この方法が経済学史研究においても適用可能なのかについては議論の余地があろう。

18)「世界不況」のテキストデータだけを用いて検討すると，データに偏りが出てしまうであろう。そのため，『報告書』と「世界不況」の頻度分析において作成されたExcelファイルを用いて，それぞれのカテゴリーに分類できる語を目視で探索した。

19) orと入力しても構わない。また，andは「&」，notは「！」を入力する。

図表 22　コーディングルール作成

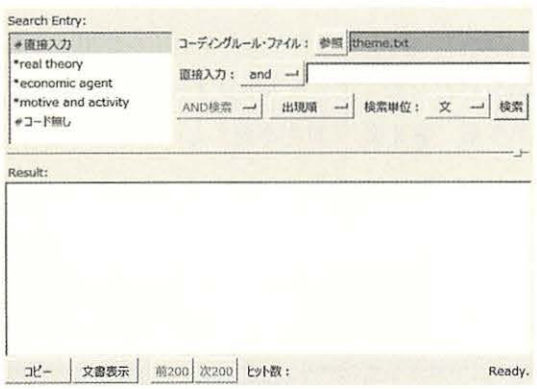


4.3.2　コーディングルールに基づく集計

このようにして作成したコーディングルールを用いてテキストマイニング分析を行う。

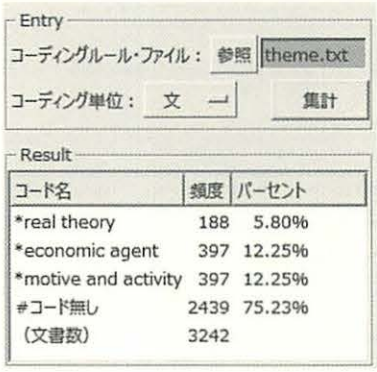
まず，KH Coder の「ツール」→「文書」→「文書検索」内の「コーディングルール・ファイル」に，作成したコーディングルールのテキストファイルをインポートする（図表 23）。ここでは図表 23 右側の「検索単位」を「段落」ではなく「文」に設定している。「段落」にすると，ロバートソンの「世界不況」では分析対象となる段落が少ないためである。分析対象の数をできるだけ大きくするために，今回は「文」に設定した。

図表 23　「文書検索」にてコーディングルールをインポート



次に，「ツール」→「コーディング」から，「単純集計」と「クロス集計」を行う。

図表 24 単純集計

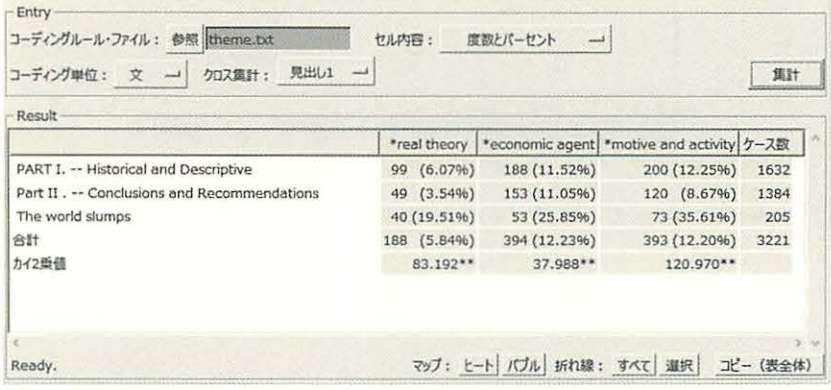


コード名	頻度	パーセント
*real theory	188	5.80%
*economic agent	397	12.25%
*motive and activity	397	12.25%
#コード無し	2439	75.23%
（文書数）	3242	

「単純集計」では，各コードに含まれる語が，全文書（3242 文）のうち，何文において出現したかを集計する。図表 24 を見ると，「*real theory」に分類された語は 188 文（全体の 5.8%）で出現し，「*economic agent」は 397 文（12.25%），「*motive and activity」は 397 文（12.25%）で出現したことが分かる。「#コード無し」が 75.23% であることから，上記 3 つのコードでは文書全体の 25% ほどしか要約していない。少ないコードで文書全体をカバーするのではなく，特徴あるコードを抜き出しているという理由から，この数値を正当化しておこう[20]。これらは全文書中での集計なので，テキストの各部分の特徴などを捉えることはできないが，この集計結果が次の「クロス集計」において活かされる。

「クロス集計」では，テキストの部分ごとにおける出現頻度を知ることができる。今回の場合，『報告書』と「世界不況」を統合した 1 つのテキストファイルを用いており，かつ，『報告書』を部（Part）によって 2 つの分けたため，全テキストは 3 分割されている。よって，クロス集計においては，『報告書』の第Ⅰ部，第Ⅱ部，「世界不況」という 3 つの部分における各コードの出現

図表 25 クロス集計



	*real theory	*economic agent	*motive and activity	ケース数
PART I. -- Historical and Descriptive	99 (6.07%)	188 (11.52%)	200 (12.25%)	1632
Part II . -- Conclusions and Recommendations	49 (3.54%)	153 (11.05%)	120 (8.67%)	1384
The world slumps	40 (19.51%)	53 (25.85%)	73 (35.61%)	205
合計	188 (5.84%)	394 (12.23%)	393 (12.20%)	3221
カイ2乗値	83.192**	37.988**	120.970**	

20) コーディングによってテキスト全体を要約する場合は，「#コードなし」の割合をできるだけ小さくしなければならない（牛澤 2018 : 96）。なお，3 つのコードとも統計的に有意である（図表 25：カイ 2 乗検定）。

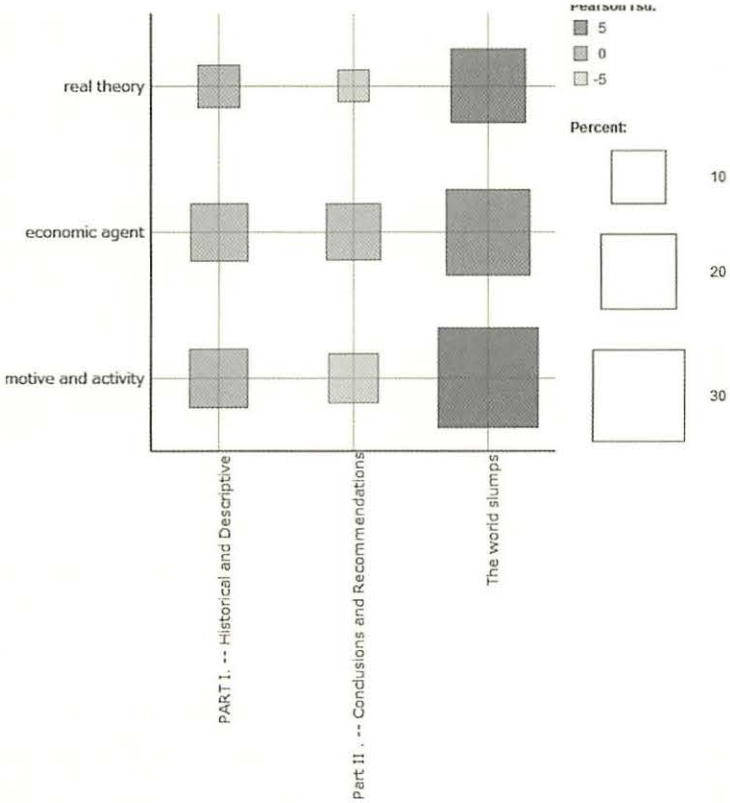
頻度が示される。分析結果は次のとおりである。

図表25には，左の列から，文書の部分（『報告書』の第Ⅰ部，第Ⅱ部，「世界不況」），次に各コード「*real theory」，「*economic agent」，「*motive and activity」，そして，「ケース数」（出現した文の数）が並ぶ。記入された数値は，各コードに属する語が出現した文の数であり，（　）内の数値は，当該部分における，その文章の出現割合である。例えば，図表25の2行2列目の「99（6.07%）」は，「*real theory」コードに属する語が，『報告書』第Ⅰ部の全1632文のうち99文（6.07%＝99/1632）で出現した，ということである。

数値をこのまま検討してもよいが，KH Coderではこの結果を図示して視覚的に理解することができる。以下に「マップ（バブル）」[21]と「折れ線」を示す。

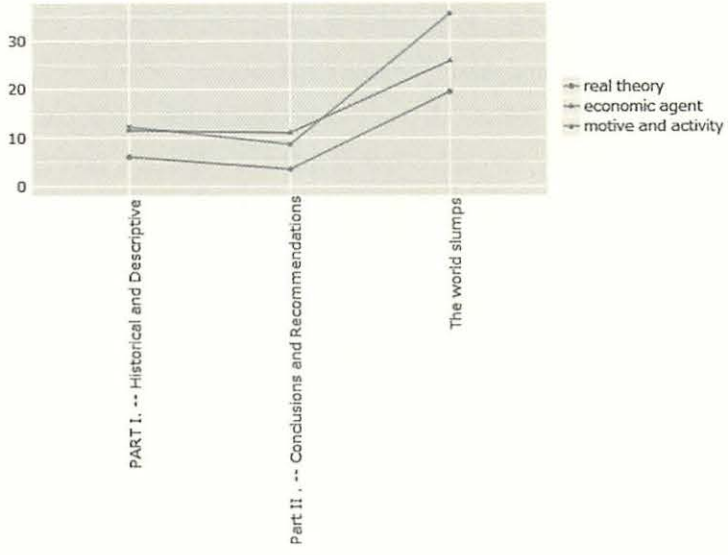
バブル図では，2つ情報が示される。1つ目は，出現割合の大きさであり，これは四角形の大きさによって表現される。2つ目は，Pearson残差であり，これは四角形の色の濃さによって表される。言い換えれば，ある部分での出現割合が他の部分の出現割合よりも相対的に大きければ，四角形の色は濃くなる。

図表26　クロス集計：マップ（バブル）



21）「バブル」の他に「ヒート」という図を選択することもできる。これは色の濃さによって，出現頻度の多さを示した図である。モノクロ印刷の場合は見えにくくなるかもしれない。

図表 27 クロス集計：折れ線



よって，図表 26 から判断すると，いずれのコードにおいても，ロバートソンの「世界不況」における出現割合が最も大きく，かつ，他の部分の出現割合と比較しても，多く出現していることが分かる。また，図表 27 の「折れ線」を見ても，出現割合の違いを確認することができる。冒頭で作成したコードは，やはりロバートソンの特徴とみなしてよさそうである。

以上のように，マクミラン委員会におけるロバートソンの言説の特徴として，金融問題においても実物理論を重視したこと，そして，経済主体としての人間の意思決定とそれに基づく活動を強調したことを挙げることができるであろう。

4.4 原典による引証

ここまで展開してきたテキストマイニング分析によって，次のような仮説が機械的な量的分析でも質的カテゴリーに基づいた分析でも同様に示された。

(1) ロバートソンは，金融の議論においても，財市場（実物理論）を重視した。
(2) ロバートソンは，経済現象の説明において，経済主体の意思決定とそれに基づく経済活動を重視した。

いずれも，全体の文書である『報告書』と比較して著しい特徴とみなせる。

本節では，最終的な第 4 段階として，これらの仮説を対象となる文献の精査（原典による引証）によって改めて確認する。以下では，とくに，‘goods’ と ‘commodity’（財市場），‘desire’ と ‘willingness’（経済主体の意思決定）の語が含まれた文（文章）を抜き出す。

4.4.1 金融問題と実物理論

まず，経済主体の意思決定と貨幣の関係ついての記述を引用することで，ロバートソンの主張を検討しよう。なお，強調の傍点と〔　〕内は引用者によるものである。

> 「銀行組織という媒介を通じて貯蓄をしようとする公衆の願望が増大していることは，…所得が費やされたそれらの（新たな道具財を含む）財の「一般物価水準」の下落として表れるであろう」（Robertson 1930 : 117），

とロバートソンが述べているように，彼は，公衆の貯蓄欲求からデフレを説明している。貯蓄欲求の増大は，消費欲求の縮小を意味する。消費欲求が縮小すると，貨幣の流通速度が減退するため，物価が下落するのである[22]。この点については，次の言説からも明らかであろう。

> 「彼ら〔銀行家〕が示す必要があったのは，…彼らの貸付の圧倒的大部分の結果として創造された貨幣が，諸商品…と容易かつ活発に接触するような種類のものだ，ということであった」（Robertson 1930 : 119）。

つまり，銀行家が自ら行った貸付の正当性を主張するためには，生産的な活動に融資したということに訴えるのではなく，その生産的な活動が最終的に取引によって消費まで行きついたのかを検討しなければならないのである。生産された財が交換されねば，経済は減退するのであり，その貸付は正当化しえないのである。

このように，ロバートソンは，財と貨幣の交換（取引）における変化が物価に影響を及ぼすと考える。ゆえに，彼は，銀行がデフレ・インフレをコントロールすることに関して悲観的な立場をとる。

> 「銀行組織は，結果として生じる預金が商品と効果的に接触させることを保証しない形でしかその資産を拡大することができないかもしれない。それゆえ，何らかの助けがなければ，商品の価格水準の下落を止めたり，その上昇を画策したりはできないかもしれない」（Robertson 1930. : 121-122）。

以上のように，ロバートソンは，消費・貯蓄をしようとする個人の意思決定，あるいは，財と貨幣の交換（取引）を起点として物価変動を考察していることが分かる。

そして，産業不況の原因に話が及ぶと，彼はこのように述べる。

22）ロバートソンは物価について議論する際，「マーシャルの K」すなわち公衆の保蔵欲求の変化に注目していた（Robertson 1931 : 410）。詳細は，仲北浦の博士論文（同志社大学）p.104，脚注 97 を参照されたい。

「産業不況の原因の中でも，人間のある特定の欲求の大きな集合における一時的な飽和可能性 gluttability が，一番重要であると私は思う」（Robertson 1930 : 122）。

しかも，「人口と世界貿易の成長率が低減しており，道具〔財への需要〕の飽和は，昔よりも長く続く傾向にある」（Robertson 1930 : 124），と。つまり，ロバートソンは，経済主体の消費欲求が減退するという需要側の原因を強調しているのである。

ただし，彼は供給側の原因についても指摘している。

「貯蓄を銀行に託す願望の増大が，主に確信と事業の減退を原因とするならば，運転資本を用意するための前払いへの反拡大に対する抵抗は段違いである。新たな事業を始めようとする意志やこれまでの事業を拡大しようとする意志は欠けているかもしれない」（Robertson 1930 : 118-119）。

確信の欠如の結果として貯蓄が増大した場合，銀行はその増大した預金をどこに融資するのか。基本的には，運転資本か固定資本に融資することになるが，確信が欠如しているということは，生産活動の拡大も見込めない。ゆえに，増大した貯蓄を活用できず，融資が滞るのである。

以上のように，ロバートソンは，産業不況の原因として，金融問題を中心的には扱っていない。むしろ，経済主体の意思決定とそれに基づく経済活動の変動を中心に据えて考察しているのである。

4.4.2 ロバートソンの経済政策

以上のような特徴をもつロバートソンの言説は，どのような政策提言に行きつくのか。本節では政策に関する彼の証言を検討する。

ロバートソンは，物価は財と貨幣の交換（取引）によって決定され，実際に取引が行われなければ経済活動は活性化しないと考える。そのため，貨幣を単に移転するような政策には慎重な姿勢をとる。

「銀行組織という媒体を通じて貨幣をばらまく dish out ことは，そうして得られた購買力がある特定の諸財に費やされるのではなく，全般的な財に対する潜在的な支配権という形で保持〔保蔵〕されるならば，そのばらまきは不適当な救済策であるかもしれない」（Robertson 1930 : 122）。

つまり，財への需要がすでにあり確実に取引が活性化するという状況でない限り，景気を向上させる政策として所得移転を行うことは適切ではない。あくまでも，移転された貨幣が財市場で費やされねば（消費活動が行われねば）不況から回復できないのである。

また，需要の飽和を不況の第1の原因とするロバートソンは，その飽和を緩和する政策につい

て，消費財と道具財（資本財）に分けて論じている。

「消費財に関して，その主要な救済策は，新たな欲求を絶え間なく刺激するという，むしろ魅力のないものであるようにみえる」（Robertson 1930 : 122）。

「道具財と建設財――鉄道車両，鉄塔，トイレなど――への通常の取引需要が一時的な飽和状態にあるとき，道具財と建設財への共同的な需要を組織化して絞り出すために，中央・地方の公共団体が，鉄道会社や電力委員会などのような半公共団体と同じように介入すべきだということが，最も純粋な共通感覚であると私はつねに思ってきた」（Robertson 1930 : 124）。

消費財については，単に需要を喚起し続けるというやや消極的な政策にとどまっているが，資本財については，政府が代りに需要を拡大する公共事業を支持している。この点こそ，Presley（1979）や Fletcher（2008）でも指摘されていた「大蔵省見解」との対立点である。

最後に，ロバートソンは消費財と資本財の生産・消費のバランスを調整することもまた政策として必要だと主張する。

「道具財産業の不況が永続的には継続しないであろうという仮定の下では，消費財の流れの増大は，道具財市場が稼働するときに，その市場を維持するのに必要であろう」（Robertson 1930 : 123）。

資本財産業の生産活動を維持するためには，その産業従事者のための生活資料として消費財が必要である。そのため，好景気に入れば，資本財の生産が活発になるため，消費財が不足がちになる。消費財の不足は，消費財の価格上昇（インフレ）として表れ，それは公衆への負担をともなうであろう。よって，消費財が不足しないように，不況期のころから消費財の量を維持しておかねばならない。ロバートソンは，その管理を政府がすべきだと考え，次のように主張するのである。

「建設財産業の活動が新たに噴出する機が熟したときに，その活動に従事する人たちの生活資料は，インフレと強制的な節欲[23]を通じてそのときの消費者から搾り取る必要もなく利用できるようにするために，技術的に可能な消費可能財を蓄えておくという政策に，何ら本質的に愚かなところはない」（Robertson 1930 : 123）。

23）ここで言及されている「強制的な節欲」こそ，かの「自動的ラッキング」である。この概念もまた，ロバートソンの実物重視の特徴（「努力」概念をコアとする経済学体系）を表している。詳しくは仲北浦（2017 a）を参照されたい。

以上のように，本研究では，テキストマイニングの「接合アプローチ」という手法を応用して，マクミラン委員会におけるロバートソンの証言の特徴を探った。さらに，この「接合アプローチ」の前後に，それぞれ「柔らかな作業仮説」と「原典による引証」という，伝統的な経済学史研究の蓄積および方法を配置するという重層的なアプローチをかけた。

このような新しいスタイルでも同様に，ロバートソンが，金融に関する議論において，人間の意思決定とそれに基づく経済活動という実体経済を重んじていたことが明らかになった。

V．おわりに

本稿では，ロバートソンの貨幣理論における実物理論について主に論じた。しかし，本稿4.2.2でも指摘したように，彼の貨幣理論を特徴づける必要がある。例えば，貨幣を表すmoneyという語の使用法や，保蔵に関する議論などである。これらの分析をさらに深めるためには，今回分析対象とした「世界不況」のみならず，他の著作を広く検討する必要があるだろう。

また，本稿では，ロバートソンがマクミラン委員会に提出した証言資料だけを分析した。しかし，ロバートソンの特徴をさらに理解するためには，証言資料のみならず，彼の口頭証言（委員との質疑応答）をも分析対象に含める必要があるであろう。そこにはケインズとのやり取りも多く含まれているため，両者の言説を同時に解析することで，同じ文脈（マクミラン委員会）で共通する用語と共通しない用語など，外部変数を用いたテキストマイニングが可能になるであろう。このような分析によって，ケインズが『貨幣論』を出版する直前における，ロバートソンとケインズの思想の異同を解明しうる。これらの論点に関しては，今後の課題としたい。

補論（1） PDFファイルのOCRからテキストファイル化までの作業の合理化

小峯・下平（2017：59-60）が指摘しているように，テキストマイニングには「テキストの前処理自体に非常に時間がかかり，しかもOCR（光学的文字認識）から整形したテキストでもまだ誤植が発見されるなどの問題が残っている」。また，喜田（2018）でも同様の問題が指摘されているが，その解決策としてはソフトウェア改善という長期的な展望が挙げられているのみである（喜田2018：372-373）。

前処理作業には，このような時間的コストの問題があるものの，その作業自体はテキストマイニングにおいて最も重要な前提の1つである。というのは，前処理の結果によって，テキストマイニングの解析結果に大きな揺れが生じるからである。前処理が不十分だと，正しい解析は行われない。また，分析の途中で前処理の不備に気づいたとすれば，その分析は初めからやり直さなければならない。よって，前処理はテキストマイニングで最も正確に行われなければならない作業と言える。

このように，テキストマイニングには，前処理の時間的コストと正確性という課題がある。そこで，補論では，筆者がテキストマイニングにおいて試行錯誤してきた前処理作業の合理化の方

法を紹介する。現在の技術では，PDF ファイルを読み取り可能なテキストファイルに変換する際，必ずノイズが発生してしまう。しかし，これを修正していくには，現在のところ目視しかない。ゆえに，「目視でノイズを修正する」ということを前提として，その作業の効率化と精度向上を図っていかねばならない。そうすることで，テキストマイニングの作業負担が少なからず軽減され，ひいては，分析規模のさらなる拡大，テキストマイニング導入者の増加を見込むことができるかもしれない。

以下では，「FineReader 14 Standard」[24]（OCR 用ソフトウェア），「Microsoft Word」，「Microsoft Excel」，「サクラエディタ」[25]（テキストエディタ）そして「KH Coder 3」を用いて，効率的かつ高精度と考えられるテキストファイル化の手法を紹介する。

①「FineReader」による OCR および「検証」

日本での流通状況を考えると，PDF ファイルの OCR には，Adobe 社の「Acrobat Professional」が広く利用されているかもしれない。しかし，先述のとおり，現在の技術では，OCR の過程で多くのノイズが生じてしまう。そのため，PDF ファイルから Word ファイルへと変換し，Word の標準機能である「校閲」から，1 語ずつ目視して確認していく他ないのである。この作業は多大な時間と労力を要する。

そこで，目視による確認作業をより効率的に行う方法を提案する。それは，OCR に ABBYY 社の「FineReader」を用いる方法である。このソフトウェアの利点を以下に示す。

FineReader には「OCR エディタ」という機能がある[26]。この機能は，PDF ファイルと Word ファイルに齟齬（ノイズ）がないかの確認を支援してくれる。本機能では，図表 28 のように，画面上に PDF ファイル（左側）と Word ファイル（右側）が隣り合わせで表示される。さらに，信頼性の低い（ノイズかもしれない）箇所はハイライトで示される[27]。これだけでも，非常に大きな時間の節約になる。

次に，ハイライトで示された信頼性の低い箇所は，「検証」ツールを用いることで，PDF と Word を照らし合わせて，誤りの有無を確認しながら修正することができる（図表 29）。その際，当該箇所を目視で探す必要はない。というのは，新たに開いたウィンドウ内に，その該当箇所が自動的に示されるからである。そして，そのウィンドウ内で Word ファイルの修正をすることができる（図表 30）。

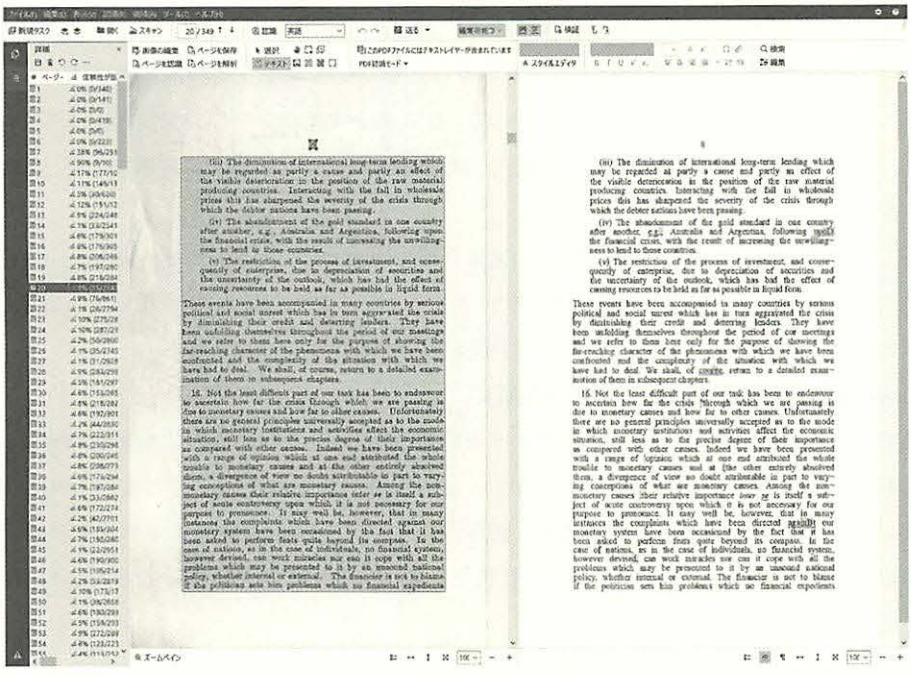
24）本研究では Windows 版「FineReader 14 Standard」（JPY 20,000）を用いた。なお，Mac 版「FineReader Pro for Mac」では同様の作業ができないようなので注意が必要である。詳細は ABBYY 社 HP（https : // www.abbyy.com/ja-jp/）を参照のこと。

25）筆者はテキストエディアに「サクラエディタ」というフリーソフトを利用している。Windows の標準機能である「メモ帳」でも十分であるが，「サクラエディタ」の方が，検索と置換の面で優れている。https : //sakura-editor.github.io/ で入手できる。

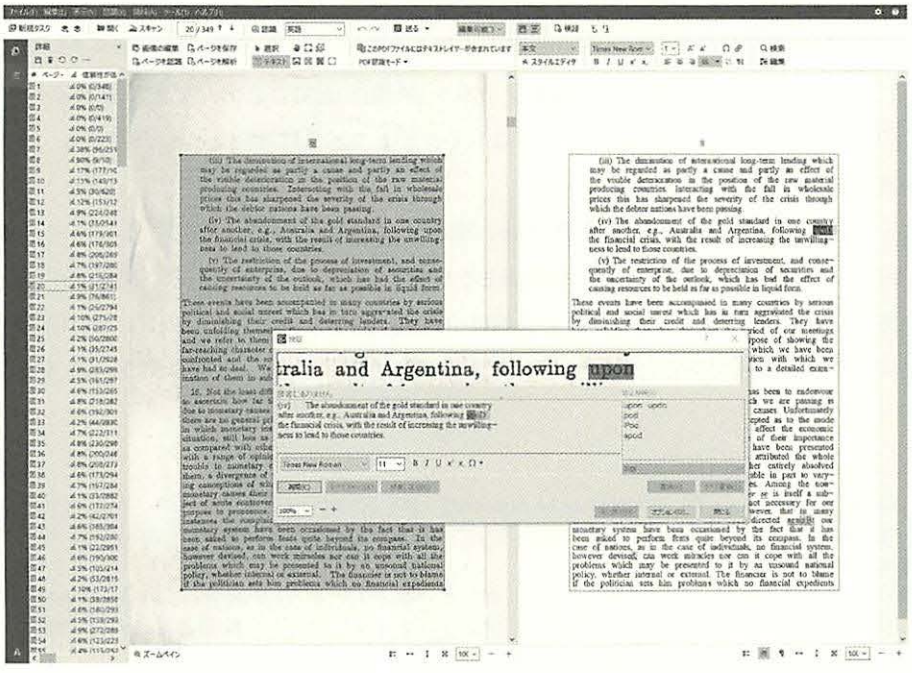
26）ソフトウェアを起動し，「ツール」→「OCR エディタ」を選択する。

27）Presley（1979）の書籍で実際に検証した結果，信頼性の低い箇所は本文全体に対して約 2.3% であった。しかし，信頼性が低いと認定されたもののうち，そのほとんどが誤っていなかったので，実際のノイズは 1% 程度だろうと思われる（1, l, I, i などの識別が苦手のようである）。

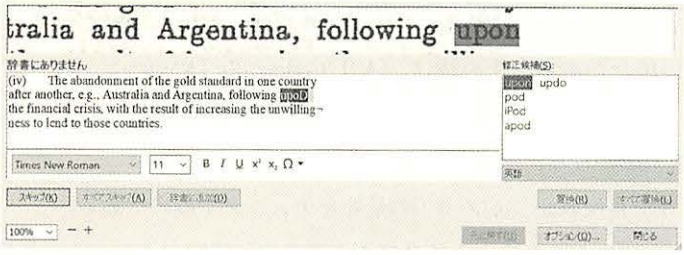
図表 28　PDF ファイルと Word ファイル



図表 29 「検証」ウィンドウの表示



図表30 「検証」ウィンドウ内での Word 編集



以上のように，「FineReader」は，「目視でノイズを修正する」という煩雑な作業を軽減する。これによって，作業者自身がノイズを目で追って探すという手間が格段に削減される。のみならず，ソフトウェアの補助があることで精度の面でも有利に働くであろう。

②「Microsoft Word」による「校閲」

以上のようにして作成した Word ファイルを「Microsoft Word」の「校閲」機能でさらに洗練していく。この作業は，目視・手作業となるが，①の手順を経ているため，多くの時間はかからない。

「校閲」が終われば，Word ファイルをテキストファイル化するのだが，この際に1つの留意点がある。それは，Word ファイルにおける「’」（シングルクオーテーション）」と「”」（ダブルクオーテーション）」と「－（エムダッシュ）」が，テキストファイルでは，それぞれ「e」と「h」と「？」に変換されてしまうことである。このような自動的な変換を避けるため，「Microsoft Word」の「検索と置換」（Ctrl＋h）で，あらかじめ消去しておくと良いであろう。ただし，「置換」の機能を使用する際は，細心の注意を払うべきである。というのは，一括での置換をすることで，意図しない変換が行われるかもしれないからである[28]。そのような操作上の誤りは，テキストマイニングの解析結果に大きな影響を及ぼしかねない。しかも，そのような誤りはすぐには気づかず，後になって気づくことが多い。よって，一括置換などの大きな編集を加える場合には，別のファイルとして新たに保存するなどの工夫をしておくのが良いであろう。

③「KH Coder」による最終チェック

①と②の手順を経れば，ほとんどのノイズは修正されているであろう。Word ファイルから作成されたテキストファイルを「KH Coder」に読み込ませ「前処理」を実行する。ただし，この作業はテキストマイニングのためのものではない。「KH Coder」の各機能を使って，ノイズの有無を最終確認するのである。

まず，「ツール」→「抽出語」→「抽出語リスト（Excel 出力）」にて「品詞別」にチェックを入れて「Microsoft Excel」を開く（自動）。ここでも「校閲」を行って，再度ノイズの有無を確

28）例えば，シングルクオーテーション「’」を一括で消去したときに，「it’s」も「its」に置換されてしまうという事例が考えられる。

認し，修正する[29]。

だが，ときに，抽出語リスト内でノイズらしき語を発見しても，PDF ファイル内ではその語が見つけられない場合がある。その時は，「KH Coder」の「KWIC コンコーダンス」を活用する。「ツール」→「抽出語」→「KWIC コンコーダンス」内の「抽出語」に，当該のノイズらしき語を打ち込み検索をかける。すると，その語の出現箇所が表示される。その情報を手掛かりとして PDF ファイルを再度確認し，ノイズの是非をチェックするのである。

以上の手順によって，おそらく，ノイズ修正の効率化と精度向上が，ある程度は実現できるであろう[30]。もちろん，この方法が最善というわけではない。OCR や情報処理技術の向上によって，あるいは，さらなる試行錯誤の結果として，より効率的な方法を模索していくべきであろう。

補論（2） テキストマイニングと著作権

第 2 の問題として，著作権の問題がある。テキストマイニングは，①アナログ文書をデジタル化（PDF 化およびテキストファイル化）する過程と，②デジタル化された文書を解析できる形に整形（前処理）し解析する過程に大別される。これらの過程において，生じうる著作権の問題について一考する[31]。

まず，著作権の保護期間について概説しておこう。著作権の保護期間は，ベルヌ条約（第 7 条(1)）に基づいて，日本では著作権法第 51 条第 2 項で「著作者の死後 50 年」と定められている。海外の著作物であっても日本では「死後 50 年」の原則が適用される。ただし，サンフランシスコ平和条約第 15 条（および国内法「戦時加算特例法」）の定めにより，日本は連合国の国民に対して著作権保護期間の戦時加算が義務づけられている。例えば，アメリカ・イギリス・フランスにおける戦前（1941 年 12 月 8 日より前）の著作物の保護期間は，通常の「死後 50 年」に 3794 日（約 10 年 5 か月）を加えた期間となる[32]。すなわち，戦前の著作物であれば，著者の死後から約 60 年間が著作権の保護期間である。

なお，2018 年 12 月 30 日に発効する環太平洋パートナーシップ（TPP）協定では，保護期間は，「死後 70 年」となる。また，「戦時加算」については，撤廃の方向性が各国と話し合われている。

29）とくに接頭辞や接尾辞には注意が必要である。例えば「con-」や「-able」は，Word や Excel では独立の単語として認識されてしまい，「校閲」によっては修正することができない場合がある。

30）以上の手順で作業を行った場合，343 ページある著作（Presley 1979）のノイズ処理にかかった時間は約 3 時間であった。

31）小林（2017 a : 18）はアメリカ・韓国におけるフェアユース fair use という概念を紹介している。これは個人による非営利の研究目的の場合，著作物の複写利用が可能になる考えである。

32）戦中（1941 年 12 月 8 日から 1952 年 4 月 27 日まで）の著作物の場合，著作権を取得した日から終戦日（1952 年 4 月 27 日）までの日数を，「死後 50 年」に加算する。戦後の著作物に関しては，戦時加算の必要がないため，「死後 50 年」が適用される。ここで，サンフランシスコ平和条約の締結日までが「戦時中」と見なされている。

以上をふまえて，冒頭で示したテキストマイニングの過程における，著作権の問題を一考する。

まず，①アナログ文書をデジタル化（PDF 化およびテキストファイル化）する過程について。著作権法第 30 条で定められている「私的使用のための複製」であれば，著作権の保護期間に関係なく，デジタル化することに問題はないようである。例えば，数人の研究者グループ内で，著作権の切れていない著作物を PDF 化して共有することは基本的に許されるようだ。ただし，不特定多数の人がアクセスできる状態にしてはいけないため，その点については注意が必要である。

次に，②デジタル化された文書を解析できる形に整形し解析する過程について。この点に関しては，2018 年の著作権法改正（施行 2019 年 1 月 1 日）が注目に値する。この改正は，近年のデジタル化・ネットワーク化の進展に対応するために，柔軟な権利制限規定を整備するためのものである。この改正によって，ビッグデータを活用した情報解析のための著作物利用が大幅に認められることになった[33]。とくに，第 30 条の 4 の改正は，直接，テキストマイニングに関係するものと思われる。この改正において，「情報解析（多数の著作物その他の大量の情報から，当該情報を構成する言語，音，影像その他の要素に係る情報を抽出し，比較，分類その他の解析を行うことをいう。……）の用に供する場合」，「著作権者の利益を不当に害すること」にならない範囲であれば，「いずれの方法によるかを問わず，利用することができる」。つまり，著作権で保護されている著作物であれ，その著作権者の利益を不当に害さないかぎり，テキストマイニング分析をすることは許される，と解してよいであろう。ただし，文部科学省による「改正の概要」でも指摘されているように，この規定は「ある程度抽象的に定めた」ものであるため，随時，判例などの状況を把握しておく必要があるだろう。

例えば，朝日新聞の記事データベース「聞蔵」では，その利用規定（2018 年 4 月現在）の 4 に，「本サービスは，日本語解析のために利用したり，プログラムによる自動検索をしたりすることはできません」とある。また日本経済新聞の記事データベース「日経テレコン」では，「データベースからアウトプットされた検索結果を報告書や書籍など外部に使用する場合は著作権者の許諾を得る必要があります」とある。これらの規定は，2018 年の著作権改正と両立しないかもしれない。今後の対応に注目する必要があるだろう。他方，オープンアクセス化[34]の世界的な規模での進展を背景に，例えば PLOS 社はテキストマイニングによる利用を念頭に，20 万点以上の論文データを XML フォーマットで提供すると発表している（2017 年 11 月）[35]。また Gale 社も自社のデータベースをデータマイニング可能になるような配慮をしている[36]。

以上のような著作権の問題に加えて，テキストマイニング分析において配慮すべきと考えられ

33）逆に言えば，情報解析の利用において，著作権の保護範囲が縮小した。

34）学者自身や学術出版社などが一定の条件下で，その論文を制限なしで閲覧・ダウンロード可能にすること。学術の成果は万人にアクセス可能であるという理念による。

35）https : //blogs.plos.org/plos/2017/11/unrestricted-text-and-data-mining-with-allofplos/（2018. 11. 3 アクセス）

36）https : //www.cengage.com/search/showresults.do?N = 197 + 4294891079（2018.11.3 アクセス）

る点に言及しておこう。それは，①と②の過程（操作）における責任の所在を明示化し，かつ，作成したテキストデータを開示できる状態にしておくことである。つまり，文書のデジタル化をしたのは誰か，デジタル化された文書を解析可能な形に前処理したのは誰か，その解析をしたのは誰かについて，それぞれ明示する必要がある。というのは，テキストマイニングでは，データ上の軽微なノイズであっても，その解析結果に大きな影響を及ぼすからである。そのため，どのデータを用いたのかという点が，研究上において重要な情報となるのである。

例えば，A 氏が作成したデータによる解析結果と，B 氏が作成したデータによる解析結果があるとしよう。これらの解析結果の間で齟齬が見られる場合，両方のデータを検証してどちらが正確かを確かめることによって，どちらの解析結果を認定するかを判断することになるだろう。その際，そのデータ作成における責任者が，A 氏なのか B 氏なのかが不明なままだと，その判断ができなくなる。それぞれの操作における責任の所在を明示化することによって，より正確な解析を目指すことができる。また，このような方式を用いることによって，テキストマイニング分析の透明性と再現性[37]が担保されるであろう。テキストマイニングは，より説得的な証拠の提示を目指すものであるため，他の研究領域における手法を取り入れていくという姿勢が必要になると考えられる。

参 考 文 献

（テキストマイニング分析の対象文献）

Report : Committee on Finance & Industry : presented to Parliament by the Financial Secretary to the Treasury by command of His Majesty, June, 1931, London : His Majesty's Stationery Office, 1931.

Robertson, D. H.（1930）'The World Slumps,' in Robertson, D. H.（1931）*Economic Essays and Addresses,* with A. C. Pigou, London : P. S. King & Son. Ltd., pp.116-138.

（英文文献）

Iker, H. P. and Harway, N. Y.（1969）"A Computer System Approach toward the Recognition and Analysis of Content," in Gerbner, G. A. et al. eds., *The Analysis of Communication Content : Developments in Scientific Theories and Computer techniques,* New York : Wiley & Sons, pp.381-486.

Fletcher, G.（2008）*Dennis Robertson,* Basingstoke : Palgrave Macmillan（下平裕之訳（2015）『デニス・ロバートソン』勁草書房）.

Komine, A. and H. Shimodaira（2018）"Keynesian Elements in Beveridge's *Free Society*（1944）: A Text Mining Approach to the History of Economic Thought', Discussion Paper Series, No.18-01, Faculty of Economics, Ryukoku University, September.

Matsuyama, N.（2016）"A Study of Text mining for Research into the History of Economic Thought : The case of Alfred Marshall's *Principles of Economics*（1890）," Discussion Paper, University of Hyogo, No.89.

Presley, J. R.（1979）*Robertsonian Economics : An examination of the work of Sir D. H. Robertson on industrial fluctuation,* London : Macmillan.

Robertson, D. H.（1915）*A Study of Industrial Fluctuation : An Enquiry into the Character and Causes of the So*

37）つまり，誰が解析しても結果は同じになる，ということである。

-called *Cyclical Movements of Trade*, London : P. S. King & Son.

――― (1926) *Banking Policy and the Price Level : An Essays in the Theory of the Trade Cycle*, London : P. S. King (高田博訳 (1955)『銀行政策と価格水準――景気循環理論に関する一試論――』巌松堂書店).

――― (1931) "Mr. Keynes' Theory of Money," *Economic Journal*, Vol.41, No.163, pp.395-411.

――― (1933) "Saving and Hoarding," *Economic Journal*, Vol.43, No.171, pp.399-413.

Romesburg, H. C., (1984) *Cluster Analysis for Researchers*, Belmont : Lifetime Learning Publications (西田英郎・佐藤嗣二訳 (1992)『実例クラスター分析』内田老鶴圃).

(邦文文献)

伊藤公一朗 (2017)『データ分析の力――因果関係に迫る思考法』光文社新書。

石田基広・金明哲編 (2012)『コーパスとテキストマイニング』共立出版。

犬塚元 (2018)「データフィケーションの時代における思想・哲学研究」日本イギリス哲学会 2017 年度研究大会。

岩本武和 (1999)『ケインズと世界経済』岩波書店。

牛澤賢二 (2018)『やってみようテキストマイニング――自由回答アンケートの分析に挑戦！』朝倉書店。

岡嶋祐史 (2006)『数式を使わないデータマイニング入門――隠れた法則を発見する』光文社新書。

加藤三郎・西村閑也 (1985)『マクミラン委員会報告書』日本経済評論社。

喜田昌樹 (2018)『新テキストマイニング入門――経営研究での「非構造化データ」の扱い方』白桃書房。

小林雄一郎 (2017 a)『R によるやさしいテキストマイニング』オーム社。

――― (2017 b)『R によるやさしいテキストマイニング［機械学習編］』オーム社。

小峯敦 (1998)「ノーマン総裁とケインズ～公定歩合の効果と合理化問題」『経済学史学会年報』第 36 巻, 64-76 頁。

――― (1999)「政策におけるケインズ革命～失業対策から管理経済へ」西沢保・服部正治・栗田啓子『経済政策思想史』有斐閣, 181-198 頁。

小峯敦・下平裕之 (2017)「ペヴァリッジ『自由社会における完全雇用』のケインズ的要素～テキストマイニングを加味した量的・質的分析～」Discussion Paper Series (龍谷大学) No.17-01。

後藤和智 (2016)『Text Mining Maniax――フリーソフトで始める日本語計量テキスト分析』後藤和智事務所 OffLine, 2016. 8. 14 開設, on-line 版のみ。

後藤和智 (2018)『Text Mining Maniax Advance――R, KH Coder, Excel による計量テキスト分析の拡張』後藤和智事務所 OffLine。

下平裕之・福田進治 (2014)「古典派経済学の普及過程に関するテキストマイニング分析：リカード, ミル, マーティノーを中心に」『人文社会論叢　社会科学編』(弘前大学), 第 31 巻, 51-66 頁。

鈴木茂 (2017)「ケンブリッジ・サイエンスパーク」『熊本学園大学経済論集』第 23 巻第 1-4 号, 53-71 頁。

竹村彰通 (2018)『データサイエンス入門』岩波新書。

田中生夫 (1976)「公定歩合政策に関するノーマン総裁の証言――イギリス再建金本位制の金融政策研究序説」『金融経済』(金融経済研究所) 第 158 号, 1-27 頁。

菱山泉 (1997)『近代経済学の歴史』講談社学術文庫。

古谷豊 (2014)「テキストマイニングを用いたスミス『国富論』普及の分析」DSSR Discussion Papers, 第 29 号, 1-17 頁。

増田正 (2014)「群馬県下における主要 3 市議会会議録に関するテキストマイニング分析」『地域政策研究』(高崎経済大学) 第 17 巻第 1 号, 1-17 頁。

――― (2016)「北関東地方における政策課題と地方議会改革：主要 7 市議会会議録のテキストマイニング分析」『地域政策研究』(高崎経済大学) 第 18 巻第 2/3 号, 33-49 頁。
――― (2018)「我が国地方議会における政治・行政関係の計量テキスト分析」『地域政策研究』(高崎経済大学) 第 20 巻第 3 号, 1-19 頁。
松川周二 (2013)「ケインズの「マクミラン委員会」証言と『マクミラン報告書』」『立命館経済学』第 62 巻第 3 号, 211-233 頁。
松永友有 (2003)「マクミラン委員会証言録分析」『群馬大学教育学部紀要　人文・社会科学編』第 52 巻, 291-308 頁。
水本篤 (2009)「コーパス言語学研究における多変量解析手法の比較――『主成分分析 vs. コレスポンデンス分析』『統計数理研究所　共同研究レポート』232：53-64。
村上征勝 (2004)『シェークスピアは誰ですか？――計量文献学の世界』文藝春秋。
仲北浦淳基 (2016 a)「ロバートソンにおけるリスクとコントロールの経済学――『産業変動の研究』と『産業のコントロール』との関連性の考察――」『経済学論叢』(同志社大学) 第 68 巻第 1 号, 141-170 頁。
――― (2016 b)「ロバートソンの最初期の経済学研究と『努力』概念」『経済学論叢』(同志社大学), 第 68 巻第 2 号, 249-274 頁。
――― (2017 a)「ロバートソンの実物理論における消費・投資・貯蓄――『努力』概念から『ラッキング』概念への展開――」『経済学論叢』(同志社大学) 第 69 巻第 1 号, 183-208 頁。
――― (2017 b)「ロバートソンにおける実業家と銀行家」『経済学論叢』(同志社大学) 第 69 巻第 2 号, 39-57 頁。
――― (2017 c)「ロバートソンにおけるケンブリッジの伝統と革新――『努力』概念の系譜と新たな意義――」『経済学論叢』(同志社大学) 第 69 巻第 3 号, 89-115 頁。
――― (2018)「D. H. ロバートソンの『実物』的経済変動論――『産業変動の研究』における『努力』概念――」『経済学史研究』第 59 巻第 2 号, 35-55 頁。
――― (博士論文)「ロバートソンにおける英国正統派経済学の伝統と革新――「努力」概念による体系化――」(同志社大学) 2018 年 3 月。
中村宙正 (2011)「英国の新規株式公開市場に関する諸学説の再検討」Discussion Paper, Series B (北海道大学) 第 94 巻, 1-35 頁。
西村閑也 (1985)『マクミラン委員会証言録 (抜粋)』日本経済評論社。
樋口耕一 (2014)『社会調査のための計量テキスト分析――内容分析の継承と発展を目指して』ナカニシヤ出版。
山内長承 (2017)『Python によるテキストマイニング入門』オーム社。
吉田正広 (1985)「マクミラン委員会における産業利害関係者の金融政策批判」『歴史と経済』第 27 巻第 4 号, 51-61 頁。

(受付　2018 年 11 月 14 日)