様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 12 日現在

機関番号：32711
研究種目：基盤研究（C）
研究期間：2010～2012
課題番号：22500597
研究課題名（和文）　コーパス言語学的アプローチによるクーベルタン・オリンピズムの受容史研究
研究課題名（英文）　The Corpus Linguistics' Approach to the Historical Study on the Reception of Olympism in Japan
研究代表者
和田　浩一（WADA KOICHI）
フェリス女学院大学・国際交流学部・准教授
研究者番号：20309438

研究成果の概要（和文）：本研究はコーパス言語学的なアプローチによって、近代オリンピック参加以前の日本におけるオリンピズムの受容に直結する雑誌記事の著者を推定した。主な成果として、1）ピエール・ド・クーベルタンの筆による著書 6 冊および雑誌記事 49 本分のコーパスを作成したこと、2）著者推定の分析に必要なデータ形式へのコーパスからの整形方法を確立したこと、3）文長と K 特性値とから上記文献の著者がクーベルタンであったとの仮説を検証したこと、が挙げられる。

研究成果の概要（英文）：This study identified an author of the magazine article which led to Japanese acceptance of Olympism before participating in the modern Olympic Games. Mainly in this study we 1) created the corpus consisted of six books and forty-nine magazine articles written by Pierre de Coubertin, 2) established the method for formatting the corpus to the data format necessary for analyses of authorship attribution, and 3) tested the hypothesis that the author above was Coubertin by using statistics of sentence length and K-characteristic.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
２０１０年度	1,800,000	540,000	2,340,000
２０１１年度	800,000	240,000	1,040,000
２０１２年度	500,000	150,000	650,000
総　計	3,100,000	930,000	4,030,000

研究分野：総合領域
科研費の分科・細目：健康・スポーツ科学、スポーツ科学
キーワード：言語学、コーパス、クーベルタン、オリンピック、オリンピズム、史料批判、著者推定

１．研究開始当初の背景

近代オリンピック草創期において、古代および近代オリンピックを知らなかった日本が、ピエール・ド・クーベルタン（1863-1937）のオリンピックの理念（オリンピズム）を、どのように受け入れかつ解釈していったのか。このような問題意識をもつオリンピズムの受容史研究を進める中で、『中学世界』の 1903 年 11 月号に掲載された鳥谷部春汀による「體育界の偉人クベルタン」（ママ）が、クーベルタンの半生を描く Mary Girard "Pierre de Coubertin : An Appreciation" (*Fortnightly Review* 74, 1903) の抄訳だったことを明らかにした。

この英文記事はオリンピズムをクーベルタンの教育学的思想として描いているが、このような評価が定着するのは、根底的な史料収集に基づいて彼の全生涯と業績とを俯瞰した Boulongne の博士論文 "La vie et l'oeuvre pédagogique de Pierre de Coubertin" (1976) が

発表されてからのことだった。著者の Girard はこれより 70 年以上前の 1903 年に、なぜクーベルタンを教育改革者として描くことができたのか。その人物像が判然としない Girard は、実はクーベルタン本人だったのではなかったのか。これらの疑問が解明され、もし「Girard =クーベルタン」ということが明らかになれば、前述の英文記事のもつ意味が変わり、ひいてはオリンピズムの歴史認識にも変化が生じる。

本研究は、オリンピズムの受容史の認識に影響を与える一つの重要な文献について、「誰」がこれを記したのかという基本的な外的史料批判を、伝統的（歴史学的）な手法ではなく、文章をコンピュータで計量的に分析するコーパス言語学的手法によって進めようとする新しい取り組みである。1)「Girard = クーベルタン」という仮説の検証に加え、2) 体育・スポーツの歴史研究におけるコーパス言語学的アプローチの有効性を検討する、という二つの性格を有することが、本研究の特徴となっている。

２．研究の目的

この研究の目的は、近代オリンピック草創期におけるオリンピズムの日本的解釈の特徴を、電子文献資料（コーパス）を活用して実証的に跡づけながら、オリンピズムの受容史研究に対するコーパス言語学的アプローチの有効性を検討することである。

具体的な課題は、以下のとおりである。

（1）嘉納治五郎（1860-1938）が国際オリンピック委員会のアジア初の委員に就任した 1909 年以前に発表された、クーベルタンの著書・雑誌記事をコーパス化する。
（2）1909 年以前の日本におけるオリンピズム受容過程の書誌学的な特徴を、具体的には、「體育界の偉人クベルタン」（1903 年）の原典 “Pierre de Coubertin : An Appreciation”（1903 年）の著者 Girard はクーベルタンだったという仮説を、コーパス言語学の多様な分析の結果によって検証する。
（3）（2）の成果を元に、オリンピズム学説史研究への応用が期待される言語学的分析手法を検討する。

３．研究の方法

（1）文献収集

クーベルタンの全著作は、概算で 6 万ページを超えると言われている。このうち、本研究で対象とする 1909 年以前に発表された著作は、著書 18 冊と雑誌記事約 600 本（13 誌）である。本研究では、未入手の著書 11 冊と雑誌記事約 300 本とを、所蔵が確認されているフランス国立図書館、ストラスブール大学図書館、オリンピック研究センター史料室で収集することにした。

（2）コーパス作成

コーパスの作成はアルバイターの補助を得ながら、次の手順で進めることにした。

①収集済みの英語文献（著書 2 冊、雑誌記事 32 本［3 誌］）のコーパス化（テキスト入力作業）を進める。その際、1）きれいとは言えないコピー等の状態、2）複数の言語が含まれている、などといった当該資料の性格に適した入力・確認・修正作業の方法を確立させる。
②「(3) データ分析」の②で決定するデータ形式に合わせたテキストデータの整形方法を確立する。この結果は、①のコーパス化作業にフィードバックする。
③①②のおよその目処がついた時点で、仏語文献のコーパス化作業を始める。

なお、コーパス作成の作業はコーパス言語学が専門の松田ではなく、体育・スポーツ史が専門の和田が担当した。これは、1)「コーパス作成」と「テキスト整形」に体育・スポーツ史における《史料》としての正確性を確保することには、和田が責任をもつべきである、2）構築したコーパスの中身を修正したり、データを追加・加工したりすることは今後、頻繁に起こりうる、という二つの理由による。本研究はこのように、研究期間終了後の研究の発展をも見据えた方法論の確立を念頭においた。

（3）データ分析

データの分析は、次の手順で進めることにした。

①K 特性値やクラスター分析、主成分分析などの多様なコーパス言語学的分析手法の中から、「スポーツ科学の分野において初」となる著者推定作業に適した手法を検討する。
②①で決定した手法に必要なデータ形式を定め、(2) のコーパス作成作業に反映させる。
③作成したコーパスと決定した分析手法とを用いて、著者の推定に必要な計量データを算出する。

（4）オリンピズムの受容史研究におけるコーパス言語学的アプローチの妥当性の検討

（2）（3）で確立した方法論やデータ分析の結果を、「Girard =クーベルタン」という仮説を検証する中間発表に落とし込む。スポーツ史学会や関連する研究会等で受けた他の研

究者からの指摘を加え、研究を総括する。

４．研究成果

以下、「３．研究の方法」で示した 4 項目に対応させながら、本研究で得られた成果を述べる。

（1）文献収集

本研究で対象とした 1909 年以前に発表されたクーベルタンの筆による著作・雑誌記事のうち、未入手のものの収集作業を、フランス国立図書館とストラスブール大学図書館（以上、フランス）、オリンピック研究センター図書館（スイス）、マインツ大学図書館、コブレンス大学図書館（以上、ドイツ）において進めた。その結果、未入手の状態だったクーベルタンの著書 11 冊と雑誌記事約 300 本のうち、著書は 5 冊を、雑誌記事は 251 本を収集することができた。

これらの過程において、*Bibliographie des oeuvres de Pierre de Coubertin*（『クーベルタン著作目録』1991）に記載のない雑誌記事 2 本（Notes Athéniennes/ Kerkyra, in : *Le monde moderne*, 1895/1897.）を発見できた。これは、アジアの研究者が新史料発掘に貢献できることを示せたという意味で、クーベルタン研究における大きな成果だった。

収集した著書・雑誌記事の概要は、次のとおりである。

［著書］計 5 冊

L'Evolution Française sous la Troisième République (Paris : Plon-Nourrit, 1896)

The Evolution of France under the Third Republic (New York/ Boston : Thomas Y. Crowell & Company, 1897)

Souvenirs d'Amérique et de Grèce (Paris : Hachette 1897)

La Chronique de France IV (Auxerre : A. Lainer, 1903)

Une Campagne de vingt-et-un ans (1887-1908) (Paris : L'Education physique, 1909)

［雑誌記事］計 251 本

La revue athlétique 1890-1891 : 19 本

Les Sports Athlétiques 1890-1897 : 80 本

L'Indépendance Belge 1899-1906 : 56 本

Revue du Pays de Caux 1902-1903 : 52 本

Revue pour les Français 1906-1909 : 44 本

（2）コーパス作成

体育・スポーツの歴史研究で扱う史料を、言葉を数量的に扱うコーパス言語学的なアプローチによって分析した先行研究は 4 点ある。19 世紀および 20 世紀のフランス体育書における運動記述を分析した清水重勇の研究（1988-1990 年度、1991-1992 年度）と、19 世紀初頭のアメリカ新体育理論における術語体系を議論した小田切毅一の研究（1990-1991 年度、1994-1996 年度）である。科学研究費補助金の助成を受けたこれらの先駆的研究は、主に体育・スポーツの領域に特徴的な語彙の種類や使用頻度に焦点を当てたもので、本研究が見据える史料批判を目指したものではなかった。いずれの研究も、キーボードからの手入力でテキストデータを作成し、Basic 等の自作プログラムを通して作った語彙リストに目を通して特徴的な語彙を取り上げるという手法をとっている。

本研究では、コーパス化作業の能率化を図るため、1）自動給紙で高速読み取り可能なスキャナーでコピーした史料を読み取り、2）その結果を OCR ソフトで Microsoft Word 形式に変換し、3）このデータを「Word → テキストエディタ」にコピー&ペーストすることでテキストファイルを作成し、4）印刷した結果と原文とを照らし合わせながら、間違い箇所をテキストエディタ上で修正する、という方法を取ることにした。クーベルタンの著書・雑誌記事にはアクセント付文字を含むフランス語が含まれるので、テキストファイルの文字コードは「Windows 1252, iso-8859-1」とした。

①まずは以下の英語文献（著書 2 冊、雑誌記事 32 本［3 誌］）のコーパス化を進めた。

［著書］

The Evolution of France under the Third Republic (New York/ Boston : Thomas Y. Crowell & Company, 1897)

France since 1814 (London : Chapman and Hall, 1900)

［雑誌記事］

Century Magazine 1896-1901： 4 本

Fortnightly Review 1897-1904 ： 16 本

American Monthly Review of Reviews 1896-1901 ： 12 本

②データ分析には 5 つの統計値、すなわち文長（一文に含まれる単語数）の平均値・標準偏差・中央値・四分位範囲、および K 特性値を用いることにした（理由等は後述）。必要となるデータ形式は「1 センテンス＋改行」である。

①で作成したコーパス（電子資料）を「1 センテンス＋改行」のデータ形式に整形する作業を進めるに当たっては、コーパス言語学やコンピュータのプログラミング言語などに関する十分な知識をもたない和田（体育・スポーツ史）が、「整形過程」を理解しながら進められる方法を確立させることにした。研究者自身が十分に理解できない複雑なプログラムにデータの整形を託すのではなく、一つ一つの作業が何

をしているのかを確実に理解して作業を進めることによって、データ（史料）の正確性を担保するためである。

整形作業の主な内容は、1）文字以外の情報を削除する、2）各センテンスの終わりに改行を入れる、3）約物（句読点や疑問符、括弧）を削除する、ことである。具体的には、テキストエディタの検索・置換機能と基本的な正規表現とを用いながら、テキストデータを史料の特徴に合わせて少しずつ整形していく方法を確立することができた。（和田浩一「体育・スポーツ史研究関連史料からテキストデータへ――コーパス言語学的アプローチによるオリンピズムの受容史研究を見据えて――」『国際交流研究』第 15 号［2013 年、pp. 245-266.］で詳述）

コーパス言語学を専門としない研究者にも理解できるテキスト整形の方法に言及する論文・報告は、皆無に等しい。したがって、本研究で確立した方法は、歴史学とコーパス言語学の学際的研究の発展に、方法論的立場から寄与できよう。

③①②の作業を通し、コーパス作成とデータ整形の方法がほぼ確立できたので、フランス語文献のコーパス化作業を始めた。対象とした文献は次のとおりである。

［著書］4 冊

L'éducation en Angleterre (Paris : Hachette, 1888)

Notes sur l'éducation publique (Paris : Hachette, 1901)

La gymnastique utilitaire (Paris : Alcan, 1905)

Pages d'histoire contemporaine (Paris : Plon-Nourrit, 1909)

［雑誌記事］17 本

L'éducation anglaise 1887/ Un programme : Le Play 1887/ L'éducation athlétique 1889/ L'éducation de la paix 1889/ Un concours litteraire entre athlètes 1891/ Exercices de sport 1894/ Père Didon 1900/ Une nouvelle formule d'éducation physique 1902/ La débrouillardise 1902/ L'éducation physique de vos fils 1902/ Jules Simon 1903/ Le Play, réformateur et socioligue 1906/ Traité d'escrime équestre 1906/ Le devoir d'un philhèllene 1906/ Les classification sportive 1907/ La pelota 1909/ Patineurs, jouez au hockey 1909

（※．以上の作業では、清水重勇・神戸大学名誉教授の協力を得た。心から感謝したい）

（3）データ分析

①著者推定を想定した具体的な分析手法について、テキストデータ分析に使用する伝統的な各種統計的手法から、近年目覚ましい発達を遂げたテキストマイニングの手法までを検討した。

文章の計量的な分析による著者推定の試みを、国内外・時代・ジャンルを問わず幅広く報告している村上征勝『文化を計る：文化計量学序説』（朝倉書店、2002 年）によれば、テキストデータの分析は単純な方法から複雑な方法まで多様にあり、これらの選択には、1）文化現象の内容や研究目的に応じた手法の検討、2）精度を含むデータの性格への配慮、3）できるだけ単純な方法での解明を試みる姿勢、が重要になるという。そして、著者の文体を理解するための具体的な情報として文長、単語長、単語の出現率、品詞の出現率、品詞の接続関係、語彙量、読点の付け方などを示している。体育・スポーツ史研究においてコーパス言語学の知見を初めて活用することから、本研究では、著者推定の方法として歴史がありかつこれまで頻用されてきた「文長」と「K 特性値」とを用いることにした。複雑な手法を用いて導き出された分析結果の意味が体育・スポーツ史研究者（和田）に理解できないものであれば、本研究のような学際的研究は成立しないと考えたからである。

なお、K 特性値とは、当該テキスト中の語彙の豊富さを示す指標であり、語彙が豊富であるほど値は小さくなるという性質をもつ。仮にデータ中ですべての単語が一度しか使われていなければ、K 特性値はゼロとなる。

②文長の平均値・標準偏差・中央値・四分位範囲、および K 特性値の 5 つの統計値を、Microsoft Excel の関数を用いて算出することにした。必要なデータ形式は、「1 センテンス＋改行」である。作成したコーパスからこのデータ形式へのテキスト整形の方法を、「（2）コーパス作成」の②で検討した。

③クーベルタン自身の手によるものであることが明らかなテキストデータを分析し、その統計的諸特徴と、Girard のテキストの諸特徴とを比較するという手順を取ることにした。もし、Girard=クーベルタンであれば、両者のテキストは②で定めた 5 つの統計値において似通った分布を示すはずである。この分析過程において、20 世紀初頭の突出したアメリカ人ジャーナリスト Albert Shaw (1857-1947) が *American Monthly Review of Reviews* に発表したクーベルタンの紹介記事（1898 年）とオリンピック関連記事（1894 年）とを、身元がはっきりしている英語を母語とする作家の文章（参考資料）として用いた。

Girard のテキストが雑誌記事であること

に合わせ、これと比較するクーベルタンのテキストも、著書を含めず、彼の英文記事から構成することにした。すなわち、*Century Magazine* から 4 本(1896-1901 年)、*Fortnightly Review* から 16 本(1897-1904 年)、*American Monthly Review of Reviews* から 12 本（1896-1901）の計 32 本である。

表 1 に Girard、Coubertin（クーベルタン）、Shaw それぞれの総文数、総単語数、文長平均値、標準偏差、中央値、四分位範囲、K 特性値を示した。文長の分布を箱髭グラフにしたのが図 1 である。

表 1　3 テキストデータにおける各種統計値

	Girard	Coubertin	Shaw
総　文　数	156	6,556	180
総単語数	4,863	172,085	5,464
文長平均値	31.173	26.248	30.356
標準偏差	17.487	16.385	16.904
中　央　値	28.0	23.0	27.5
四分位範囲	23.5	20.0	20.0
K 特性値	118.764	114.783	133.838

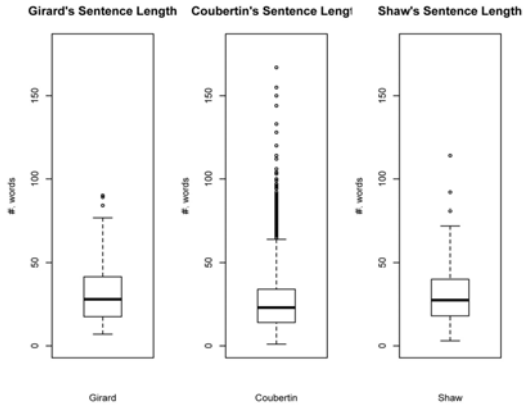

図 1　3 テキストデータにおける文長の分布（左から Girard、Coubertin、Shaw）

（4）オリンピズムの受容史研究におけるコーパス言語学的アプローチの妥当性の検討

Girard の文長平均値と中央値は、むしろ参考データであった Shaw に近く、標準偏差と四分位範囲は Coubertin と Shaw の方が Girard と Coubertin よりも近いか同一である。唯一、Girard の K 特性値のみが Coubertin に近い値を示している。これらの結果は、Girard = Coubertin とするのが困難であることを示すものとなった。

なお、ここで詳細は省くが、Coubertin のテキストをそれぞれの発表雑誌別に 3 種類に分けても、統計値に大差がないことは確認済みである。

ところで、K 特性値はテキスト量に大きく左右されることが知られている。表 1 から明らかなように、実は Coubertin のデータ量は他 2 者のそれを圧倒的に凌駕しており、これにより結果が影響された可能性が高い。これを補正するために、Coubertin のデータ全体から Girard と同数の 157 文を無作為抽出し、改めて K 特性値を算出したところ、103.865 であった。つまり、散らばりの方向が逆とは言え、ちょうど Coubertin と Shaw の真ん中に Girard が位置する形になる。補正されたデータを用いても、Girard = Coubertin とするのは難しいことが示された。

結果として、上記方法で算出した 1）文長の平均値、中央値、四分位範囲の値は Girard = Coubertin の関係性を否定し、2）K 特性値もこの関係性を明確に肯定するには至らなかった。したがって、「クーベルタンの半生を記した Girard による記事が、クーベルタン自身の筆によるものである」という仮説は、コーパス言語学の立場からは肯定されなかった。

以上の結果を発表した日本スポーツ史学会（2012 年 12 月）における質疑応答では、以下の意見がフロアーから出された。

1）フランス人クーベルタンの手による英文雑誌記事の文体の特徴が、コーパスの対象期間となった 10 年間において変化している可能性はないのか。
2）伝統的（歴史的）な史料批判とコーパス言語学的なアプローチによるそれとを総合的に検討すると、「Girard = Coubertin の可能性も否定できない」と付け加えられるのではないのか。

1）の視点は、これから展開するより精密な研究に生かす予定である。また、2）については、コーパス言語学的なアプローチによって導き出されたデータは万能ではなく、データの分析に関しては、計量的な分析とは異なる観点からの裏づけ――ここでは伝統的な史料批判――が必要であるという、体育・スポーツ史研究者の研究態度に対する貴重な示唆と思われた。オリンピズムの受容史研究におけるコーパス言語学的アプローチの妥当性は、このような複眼的な研究手法を用いる中で初めて裏づけられると結論づけたい。

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 4 件）

①和田浩一「嘉納治五郎のピエール・ド・クーベルタン宛書簡（2）―― 第一次世界大戦後からクーベルタンの IOC 会長辞任まで」『講道館柔道科学研究会紀要』第 14 輯、査読なし、2013 年、pp. 11-30.

②和田浩一「体育・スポーツ史研究関連史料からテキストデータへ――コーパス言語学的アプローチによるオリンピズムの受容史研究を見据えて――」『国際交流研究』第 15 号、査読なし、2013 年、pp. 245-266.

③松田謙次郎「日本語の攻防　ら抜き言葉」『日本語学』31(15)、査読なし、2012 年、pp. 66-75.

④和田浩一「オリンピズムという思想――新しいオリンピズムの構想への序章」『現代スポーツ評論』23、査読なし、2010 年、pp. 62-71.

〔学会発表〕（計 6 件）

①和田浩一・松田謙次郎「オリンピズムの受容史研究におけるコーパス言語学的史料批判――“Pierre” de Coubertin: An Appreciation” (*Fortnightly Review,* 1903) の作者をめぐって」、スポーツ史学会第 26 回大会（甲南大学）、2012 年 12 月 1 日

②松田謙次郎「テキスト・マイニング入門」、平成 23 年度科研基盤研究 C「教育期から労働期への移行段階における若年女性の自立と家族資本―日米比較調査―」（研究代表者竹田美知）講習会（神戸松蔭女子学院大学）、2011 年 9 月 16 日

③ 和田浩一「Pierre de Coubertin: An Appreciation (*Fortnightly Review,* August 1903) の作者は誰なのか ――記事の歴史的位置づけと内容の分析とによる作者同定の試み」、スポーツ史学会第 24 回大会（大和郡山市商工会館）、2010 年 11 月 28 日

④和田浩一「鳥谷部春汀によるクーベルタン・オリンピズムの解釈：「體育界の偉人クベルタン」『中学世界』（1903 年 11 月号）の分析」、日本体育学会第 61 回大会（中京大学）、2010 年 9 月 8 日

〔図書〕（計 5 件）

①一般財団法人嘉納治五郎記念国際スポーツ研究・交流センター「オリンピック教育を活用した青少年の健全育成に関する調査委託」報告書（和田浩一「クーベルタンから見たオリンピック教育」）、一般財団法人嘉納治五郎記念国際スポーツ研究・交流センター、2013 年、pp. 18-32.

②楠戸一彦先生退職記念論集刊行会『体育・スポーツ史の世界：大地と人と歴史との対話』（和田浩一「オリンピックの用語史：江戸後期から明治前期にかけて出版された英和辞典に注目して」）、広島：渓水社、2012 年、pp. 283-304.

③新井博・榊原浩晃編著『スポーツの歴史と文化』（和田浩一「オリンピック・ムーブメントと世界平和――ピエール・ド・クーベルタンと嘉納治五郎の教育思想を中心に」）、道和書院、2012 年、pp. 125-140.

④日比谷潤子編、松田謙次郎著『はじめて学ぶ社会言語学――ことばのバリエーションを考える』ミネルヴァ書房、2012 年、pp. 54-79.

〔その他〕

①学会発表の説明・資料などのウェブサイト上の掲載
http://homepage3.nifty.com/wadaco/index.html

②一般市民を対象とした公開講座での発表

1）和田浩一「100 年前のオリンピック：ロンドンと日本を結んだクーベルタンからの手紙」、船橋市宮本公民館「オリンピックの歴史を学ぼう」、2012 年 6 月 15 日

2）和田浩一「クーベルタンの目に映る 2012 年ロンドン・オリンピック」、Support Our Kids 実行委員会（東日本大震災被災児童自立支援プロジェクト、各国大使館後援）「イギリスホームステイ事前研修」、2012 年 8 月 9 日、国立オリンピック記念青少年総合センター

③FM ラジオ特別講演
和田浩一「未来につながるオリンピック」、TOKYO FM「明日にかける橋」、2012 年 9 月 24 日-10 月 5 日（計 10 日間）

６．研究組織

(1)研究代表者
和田　浩一（WADA KOICHI）
フェリス女学院大学・国際交流学部・准教授
研究者番号：20309438

(2)研究分担者
松田　謙次郎（MATSUDA KENJIRO）
神戸松蔭女子学院大学・文学部・教授
研究者番号：40263636